VOLUME 1

HUMAN CENTRIC

CAADRIA 2023

CEPT University

28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia

Edited by Immanuel Koh (Chair), Dagmar Reinhardt,
Mohammed Makki, Mona Khakhar, Nic Bao

Human-Centric

Proceedings of the 28th International Conference on Computer-Aided Architectural Design Research in Asia (CAADRIA 2023)

Volume 1

Edited by

Immanuel Koh
Singapore University of Technology and Design

Dagmar Reinhardt
University of Sydney

Mohammed Makki
University of Technology, Sydney

Mona Khakhar
CEPT University

Nic Bao
RMIT University

Human-Centric
28th International Conference on Computer-Aided Architectural Design Research in Asia (CAADRIA 2023)

21-23 March 2023

Hosted By:

CEPT University
Ahmedabad, INDIA

The Association for Computer-Aided Architectural Design Research in Asia (CAADRIA), Hong Kong

ISBN: 978-988-78917-9-6
ISSN: 2710-4257 (print)
2710-4265 (online)

Printed in Ahmedabad, INDIA

Foreword

The annual Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) conference provides an international community of researchers and practitioners with a venue to exchange, to discuss and to publish their latest ideas and accomplishments. These proceedings, consisting of two volumes, contain the research papers that were accepted for presentation at the 28th International CAADRIA Conference, organised by CEPT University, Ahmedabad, India. The papers are also available online at the open access cumulative database *CumInCAD* (http://papers.cumincad.org).

The proceedings are the outcome of an extensive collaborative and voluntary effort between the International Reviewing Committee and Paper Selection Committee, with the former consisting of 140 reviewers from over 30 different countries working across the globe. It marks another year of growing quality and strength in the history of CAADRIA. The papers in this publication have been selected through a two-stage double-blind peer review process. Calls for papers in July of 2022 resulted in the submission of 450 abstracts from 37 countries, of which 292 were invited for further development as full paper submission. At least 3 to 4 international reviewers were assigned to conduct a double-blind review of each of the submitted papers. Following the reviewers' recommendation and the Paper Selection Committee's evaluation, 139 papers were finally accepted in February of 2023 for their inclusion in the publication and presentation at the conference.

We congratulate the authors for their accomplishment. Our sincere thanks go out to our international reviewers, who volunteered their valuable time and effort. We also thank the organising team in CEPT University for developing the 2023 CAADRIA Conference as a strong collaborative effort between academic institutions, industry and practice. We thank Gabriel Wurzer for his relentless support with customizing the submission and review system to the needs of CAADRIA from the abstract submission all the way to production. Finally, I sincerely thank the CAADRIA community for offering me the opportunity and honour to serve as the chair of the Paper Selection Committee during the past 1 year of preparation.

PAPER SELECTION COMMITTEE

Immanuel Koh	*Singapore University of Technology and Design (Chair)*
Dagmar Reinhardt	*University of Sydney*
Mohammed Makki	*University of Technology, Sydney*
Mona Khakhar	*CEPT University*
Nic Bao	*RMIT University*

Theme: ***'Human-Centric'***

In the 21st century, the global immersion in technology has connected the world like never before, and technological impact is reaching even some of the most remote and impoverished areas of the world. As a consequence, the advancement and adoption of new technologies in all aspects of our society have profoundly changed how we design our environment to be healthy, liveable and equitable.

Computational design, simulation, analysis, fabrication, and management allow us to evaluate and forecast the performance of habitats with a deeper understanding of contexts.

Technology for social good is an agenda to articulate values, behaviours, and attitudes that focus on social impact – placing people at the focal point. In the context of a rapidly changing climate, it is has become evident that we are approaching a moment of no return; a moment in which our fate hangs in the balance. We as a species, and the societies we construct, must take immediate action and work collectively towards a brighter and more sustainable future.

CAADRIA2023 calls in particular for contributions addressing the conference theme by discussing the influence of technology on society, economy, environment, and governance pertaining to habitats where people are at the heart of the construction industry. We invite conversations and debate around key questions:

- How do different computational design approaches contribute to the design, development and policy making of complex societies?
- To what extent automation, machine intelligence and control of complex systems impact societal development?
- How does the practical and real-world applications of computational research and technologies advance broader social well-being?

The 28th Annual Conference for Computer-Aided Architectural Design Research in Asia (CAADRIA) brings together academics, researchers, and practitioners to contribute to the fields of computational design methods, instruments, and processes to innovate for humans. Contributions focusing on the Asia/Pacific context was particularly encouraged.

CAADRIA 2023 invites contributions including but not limited to the following topics:

- Artificial intelligence and machine learning in design
- Building Information Modelling
- Climate change and sustainability
- Collective, collaborative & interdisciplinary design
- Creativity, design thinking and human-computer interaction
- Digital Heritage
- Digital representation and visualization
- Generative, algorithmic & evolutionary design
- Interactive environments
- Bio-designs
- Pedagogical shifts in computational design
- Structural performance-based design and optimization
- Robotics, digital fabrication and construction
- Innovative material systems and manufacturing methods
- Theory, philosophy & methodology of computational design research
- Urban analytics, big data analysis and smart cities
- Urban modelling and simulation
- VR/AR/XR

Conference Organisers & Hosts,

Urvi Sheth	*CEPT University (Chair)*
Anjali Yagnik	*CEPT University*
Chirayu Bhatt	*CEPT University*

{https://caadria2023.org}

About CAADRIA

The *Association for Computer-Aided Architectural Design Research in Asia* (CAADRIA) promotes teaching and research in CAAD in the larger Austral-Asian and Pacific region supported by a global membership.

CAADRIA was founded in 1996 with the following objectives:

- To facilitate the dissemination of information about CAAD among Asian schools of architecture, planning, engineering, and building sciences.
- To encourage the exchange of staff, students, experience, courseware, and software among schools.
- To identify research and develop needs in CAAD education and to initiate collaboration to satisfy them.
- To promote research and teaching in CAAD that enhances creativity rather than production.

CAADRIA organizes among others an annual conference, the first of which was held in 1996 in Hong Kong. Since then, 26 conferences have been held in Australia, China, Hong Kong, India, Japan, Korea, Malaysia, New Zealand, Singapore, Taiwan, and Thailand. The annual CAADRIA conferences provide an opportunity to meet, to learn about the latest research, and to continue the discourse in the field. The 28th conference, in 2023, is hosted jointly by CEPT University, Ahmedabad. CAADRIA 2023 is held as a hybrid conference, continuing the Association's mission to bring together researchers, practitioners, and schools of the Asia-pacific region at an extended time of global Covid-19 related travel restrictions.

CAADRIA is one of the four founding organizations of the *International Journal of Architectural Computing* (IJAC), and typically co-edits one issue each year. IJAC is published by SAGE in both paper and electronic versions.

Anastasia Globa
President, CAADRIA

CAADRIA Officers

President: Anastasia Globa
University of Sydney, Australia

Secretary: Urvi Sheth
CEPT University, India

Treasurer: Hyoung-June Park
University of Hawaii, USA

Membership Officer: Jeroen van Ameijde
Chinese University of Hong Kong, Hong Kong SAR

Web Master: K. Daniel Yu
University of New South Wales, Australia

Administrative Officer: Marc Aurel Schnabel
Forum 8, Japan

Election Officer: Christiane M. Herr
Southern University of Science and Technology, China

CAADRIA 2023 Conference Committees

Hosting Institution:
CEPT University,
Ahmedabad, INDIA

Organising Committee:

Urvi Sheth	*General Conference Chair,*
Anjali Yagnik	*Co-Host*
Chirayu Bhatt	*Co-Host*
Mona Khakhar	*Local Paper Selection Chair*
Urvi Sheth	*Workshop Chair*
Jinal Shah	*Workshop Co-Host*
Urvi Sheth	*Exhibition Chair,*
Rishav Jain	*Exhibition Curatorial team*
Subin Jameel	*Exhibition Curatorial team*
Alagappan	*Exhibition Curatorial team*
Jay Thakkar	*Exhibition team*
Chandni Chhabra	*Exhibition team*
Juned Shaikh	*Exhibition team*
Namrata Dobaria	*Exhibition team*

Mahroof	*IT services*
Santosh Zokarkar	*Communication team*
Hemant Sharma	*Communication team*
Sparrow Design Firm	*Exhibition Design, Graphics Design*
Netlink	*Website Design*

Paper Selection Committee:

Immanuel Koh
Singapore University of Technology and Design *Chair*
Dagmar Reinhardt
University of Sydney
Mohammed Makki
University of Technology, Sydney
Mona Khakhar
CEPT University
Nic Bao
RMIT University

Post Graduate Students Consortium:

Sky Lo Tian Tian *Chair*
Dagmar Reinhardt
Christiane M. Herr

Sasada Prize Committee:

Tom Kvan *Chair*
Christiane M. Herr
Atsuko Karga

Young CAADRIA Award Committee:

Nic Bao *Chair*
Anastasia Globa
Immanuel Koh
Urvi Sheth
Christiane M. Herr
Sky Lo Tian Tian

International Reviewing Committee

Osaka University

Song Gao
Independent researcher

Fernando García Amen
The University of the Republic

Nicole Gardner
University of New South Wales

Nadja Gaudilliere-Jami
TU Darmstadt | CITA

John Gero
The University of North Carolina at Charlotte

Mona Ghandi
Washington State University

Anastasia Globa
The University of Sydney

Zhuoxing Gu
Tongji University

Qi Guo
Harbin Institute of Technology, Shenzhen

M. Hank Haeusler
University of New South Wales

Gilles Halin
Université de Lorraine

Yoojin Han
Yonsei University

Farahbod Heidari
Tarbiat Modares University

Christiane Herr
Southern University of Science and Technology

Pablo C Herrera
Universidad Peruana de Ciencias Aplicadas

Anca-Simona Horvath
Aalborg University

Wei Hu
Tongji University

Xiaoran Huang
North China University of Technology

Markus Hudert
Aarhus University

Aswin Indraprastha
Institut Teknologi Bandung

Mads Brath Jensen
Aalborg University

Woonseong Jeong
Chungbuk National University

Guohua Ji
Nanjing University

Sam Joyce
Singapore University of Technology and Design

Atsuko Kaga
Osaka University

Yuval Kahlon
Tokyo Institute of Technology

Mona Khakhar
CEPT University

Chin Koi Khoo
Deakin University

Daiki Kido
Kajima Corporation

Geoff Kimm
Swinburne University of Technology

Marirena Kladeftira
ETH Zurich

Christoph Klemmt
University of Cincinnati

Immanuel Koh
Singapore University of Technology and Design

Satakhun Kosavinta
King Mongkut's Institute of Technology Ladkrabang

Dennis Lagemann
ETH Zurich

Ih-Cheng Lai
Tamkang University

Christoph Langenhan
Technical University of Munich

Surapong Lertsithichai
Chulalongkorn University

Matthias Leschok

ETH Zurich

Bin Li
South China University of Technology

Andrew Li
Kyoto Institute of Technology

Chieh-Jen Lin
Tainan University of Technology

Yuezhong Liu
Nanyang Technological University

Jie Liu
Beijing University of Posts and Telecommunications

Tian Tian Lo
The Hong Kong Polytechnic University

Paul Loh
Bond University

Davide Lombardi
Xi'an Jiaotong - Liverpool University

Thorsten Lomker
Zayed University

Werner Lonsing
Independent Researcher

Shuai Lu
The University of Sydney

Yao Lu
University of Pennsylvania

Dan Luo
University of Queensland

Mohammed Makki
University of Technology Sydney

Bob Martens
TU Wien

Linda Matthews
University of Technology Sydney

Iasef Md Rian
Xi'an Jiaotong-Liverpool University

James Melsom
University of Technology Sydney

António Menezes Leitão
University of Lisbon

Saurabh Mhatre
Harvard Graduate School of Design

Alexandra Moisi
University of Innsbruck

Joy Mondal
WEsearch Lab

Walaiporn Nakapan
Parabolab

Zuhair Nasar
University of Kufa

Worawan Natephra
Mahasarakham University

Mine Özkar
Istanbul Technical University

Farzaneh Oghazian
The Pennsylvania State University

Mine Özkar
Istanbul Technical University

Stefana Parascho
Ecole Polytechnique Fédérale de Lausanne

Hyoung-June Park
University of Hawaii at Manoa

Tang Peng
Southeast University

Inês Pereira
University of Lisbon

Ahmad Rafi
Multimedia University

Balaji Rajasekaran
C.A.R.E School of Architecture

Dagmar Reinhardt
The University of Sydney

Jinmo Rhee
Carnegie Mellon University

Manuel Rodriguez Ladron De Guevara
Carnegie Mellon University

Nicolas Rogeau
Ecole Polytechnique Fédérale de Lausanne

Luís Romão
University of Lisbon

Viktoria Sandor
Universität Innsbruck

Marc Aurel Schnabel
DARA.digital

Gerhard Schubert
Technical University of Munich

Awnili Shabnam
Swinburne University of Technology

Maryam Shafiei
The University of Queensland

Milad Showkatbakhsh
Architectural Association School of Architecture

Sahar Soltani
Monash University

Djordje Stojanovic
University of Melbourne

Rudi Stouffs
National University of Singapore

Marcin Strzala
Warsaw University of Technology

Ying Yi Tan
Singapore University of Technology and Design

Wajiha Tariq Sheikh
The Chinese University of Hong Kong

Adrian Taylor
BVN

Ziyu Tong
Nanjing University

Kenneth Tracy
Singapore University of Technology and Design

Jeroen van Ameijde
The Chinese University of Hong Kong

Tomas Vivanco
PUC Chile
Tongji University

Likai Wang
Nanjing University

Shih-Yuan Wang
National Chiao Tung University

Likai Wang
National University of Singapore

Sihan Wang
China Academy of Art

Claudia Westermann
Xi'an Jiaotong-Liverpool University

Albert Wiltsche
Graz University of Technology

Chao Yan
Tongji University

Jiawei Yao
Tongji University

Rongrong Yu
University of South Australia

Yuqing Zhang
South China University of Technology

Hao Zheng
University of Pennsylvania

Xinwei Zhuang
University of California Berkeley

Keynote 1

Martha Tsigkari
Senior Partner, Head Research and Design, Foster + Partners, London
Associate Professor, UCL Bartlett, London

Martha Tsigkari is a Senior Partner and Head of the Applied R+D (ARD) group at Foster + Partners. Her background spans architecture, engineering, and computer science. She has two decades of experience working in projects of all scales and uses. She is a specialist in a wide range of areas including computational design, performance-driven design, optimisation, machine learning, interfaces & interaction, design-to-production, and fast feedback & integration. Her work incorporates the development of ML, optimisation and simulation tools, the introduction of integrated processes and the creation of physical interfaces. She has a particular interest in machine learning and has lead R&D related to the use of deep neural networks in the design process, to solve problems spanning from micromaterials to performance driven architectural layouts. She has also been investigating the use of Genetic Algorithms in the design pipeline, using them to develop solutions spaces that fit multi-objective criteria.

The White Rabbit

Foster + Partners is a global studio for architecture, urbanism, and design, rooted in sustainability. The studio integrates the skills of architecture with engineering, both structural and environmental, urbanism, interior, and industrial design, model and filmmaking, aeronautics, and many more – our collegiate working environment is similar to a compact university. These diverse skills make us capable of tackling a wide range of projects, particularly those of considerable complexity and scale. Design is at the core of everything that we do. We design buildings, spaces, and cities; we listen, we question, and we innovate.

At Foster+ Partners, we constantly strive to change the paradigm of design. To that end, technology is an intrinsic part of how we operate. We embrace change and see new technologies as a possibility, rather than a challenge. That is why we have specialist groups such as Applied Research + Development, an integrated multi-disciplinary team of architects and engineers, who are also seasoned programmers. The Applied R+D group conducts state-of-the-art research and development to solve complex design challenges, taking the latest advances out of the lab and into the hands of architects and engineers.

Using real project examples, Martha demonstrated how technology has helped us develop ideas into buildings. This was a presentation about how innovation, disruption, and uncertainty can create new possibilities for the AEC industry.

Keynote 2

Sigrid Brell-Cokcan

Founder and Head of Individualized Production (IP), RWTH Aachen University.

Current President of the Association for Robots in Architecture (RiA).

The board member of euRobotics until 2021.

Construction Robots

In recent years, she has pioneered the simple application of industrial robots for creative industries and participated in numerous international research and industrial projects. The Association for Robots in Architecture is a development partner of KUKA and Autodesk and a validated EU research institution within the FP7 program.

In 2016, Sigrid Brell-Cokcan founded the new Topic Group for Construction Robotics within euRobotics to contribute to the Multi-Annual Roadmap (MAR) for Horizon Europe. In addition, as editor-in-chief, she launched the new scientific Springer Journal Construction Robotics since 2017. Since 2022, she has been part of the writing team of the living Springer Encyclopedia of Robotics | SpringerLink on the topic of Robotics in Construction | SpringerLink.

At RWTH Aachen, her new professorship for Individualized Construction Production deals with the use of innovative machines in material and construction production. In order to enable efficient, individualized production from batch size one for the construction industry, new and user-friendly methods of human-machine interaction are being developed.

The IP chair employs researchers from various areas of robotics and building production to streamline the necessary digital workflow from initial planning to the production process and to redesign the construction site of the future via intuitive, easy-to-use interfaces.

Results on innovative developments in "haptic programming" and "cloud remote control" were presented to the public as finalists in the KUKA Innovation Award 2016 and KUKA Innovation Award on Artificial Intelligence 2021 at the Hannover Messe. This year, she has been nominated for the BAUMA Innovation Award 2022 of Research with the ROBETON project for the robot-supported controlled dismantling of concrete components. The cross-faculty Centre for Construction Robotics (CCR) on RWTH Aachen Campus was co-founded by Sigrid Brell-Cokcan in 2018 to focus on automation in construction with key industry leaders along the construction industry value chain. In 2020, Sigrid Brell-Cokcan initiated the new international consecutive Master's program "Construction & Robotics" for the Bachelor of Architecture, Computer Science, Civil Engineering, and Mechanical Engineering in order to train future engineers in the digital environment across academia for the construction industry.

Keynote 3

Neil Leach
Professor, FIU, Tongji, and EGS.

He has previously taught at other leading institutions, including the AA, Harvard GSD, Columbia GSAPP, Cornell, SCI-Arc, IaaC, and DIA. He is a co-founder of DigitalFUTURES, an academician in the Academy of Europe, and the recipient of two NASA research fellowships exploring 3D print technologies for the Moon and Mars. He has published over 40 books on architectural theory and digital design including L B Alberti, On the Art of Building in Ten Books (MIT, 1988), Rethinking Architecture (Routledge, 1997), The Anaesthetics of Architecture (MIT, 1999), Architecture and Revolution (Routledge, 1999). Designing for a Digital World (Wiley, 2002), Digital Tectonics (Wiley, 2004), Camouflage (MIT, 2006), Digital Cities (Wiley, 2009), Space Architecture (Wiley, 2014), Swarm Intelligence (TongjiUP, 2017), Computational Design (TongjiUP, 2017), Digital Fabrication (TongjiUP, 2017), Swarm Intelligence (TongjiUP, 2017), 3D Printed Body Architecture (Wiley, 2017), Architectural Intelligence (Springer, 2020) and Architecture in the Age of Artificial Intelligence: An Introduction to AI for Architects (Bloomsbury, 2022). Machine Hallucinations: Architecture and Artificial Intelligence (Wiley) will be published in May 2022. He is currently working on Interactive Futures (Birkhauser, 2023), The AI Design Revolution: How AI will Transform Architecture (Routledge, 2023), and The Death of the Architect: The Demise of the Architectural Profession in the Age of AI, (Bloomsbury, 2024)

He talked about **Artificial Intelligence in Human-Centric Architecture**.

TABLE OF CONTENTS

Artificial Intelligence and Machine Learning in Design

AI-BEWITCHED ARCHITECTURE OF HANSEL AND GRETEL

Food-to-Architecture in 2D & 3D with GANs and Diffusion Models

IMMANUEL KOH[1]
[1]*Singapore University of Technology and Design.*
[1]*immanuel_koh@sutd.edu.sg, 0000-0002-1181-1082*

Abstract. Architects such as Le Corbusier, Frank Gehry, Aldo Rossi, and Greg Lynn have implicitly turned culinary formalism into architectural formalism during their careers. How might AI assist in a similar act of bisociation (or conceptual blending)? The paper is the first to explore this food2architecture bisociation explicitly, and specifically with generative adversarial networks (GANs) such as CycleGAN and VQGAN-CLIP, and diffusion models such as OpenAI's DALL-E 2, Midjourney and DreamFusion (using Stable Diffusion). Instead of using textual input prompts to generate images of architecture only with the discipline's own vocabulary, the research merges them with the vocabulary of food, thus exploiting their potential in blending their respective conceptual and formal characteristics. While these diffusion models have recently been used by the general public to generate 2D imagery posts on various social media platforms, no existing work has conducted a detailed and systematic analysis on their exclusive capacity in bisociating food and architecture. Imagery outputs generated during two workshops involving 150 designers and non-designers are included here as illustrations. Beginning and ending the paper with the all-familiar fairy tale of the gingerbread house, the research explores the creative design bisociative affordance of today's text-to-image and text-to-3D models by turning culinary inputs into architectural outputs -- envisioning an explicitly computational version of the implicit 'food2architecture' mental models plausibly used by some of the most creative architects.

Keywords. Deep Learning, Midjourney, DALL-E 2, DreamFusion, Stable Diffusion, GANs.

1. Introduction

"Hansel, who liked the taste of the roof, tore down a great piece of it, and Gretel pushed out the whole of one round window-pane, sat down, and enjoyed herself with it." This quote in English (Grimm, 2020) from the famous German fairy tale 'Hansel and Gretel' written in 1812 by Brothers Grimm is a great example of what Arthur Koestler called 'bisociation' – the creative act of simultaneously associating two ordinarily unrelated fields in a single idea or object (Koestler, 1964). The gingerbread house in the story is the bisociation between architecture and food, one that can be inhabited from within as well as ingested from without. In fact, the act of ingestion can even be argued from both within and without when the witch was pushed into the oven and burnt to death as the story unfolded. As a somewhat cautionary tale, as in film 'The Belly of an Architect'(Greenaway, 1987), Hansel's desperation for an edible architecture can be as cancerous as it is empowering. Existing literature has mainly focused on the differences in how aesthetic experiences are triggered when encountering architecture and food (Jacobsen, 2006; Michel, 2014); while others emphasise their opposition in scale, materiality, and structures as implicit parallels (Horwitz and Singley, 2006; Legendre, 2011). However, none has incorporated AI as a subject or as an experimental tool in their investigation.

From Le Corbusier's 'Still Life' (1920) featuring culinary objects like wine bottles and plates which bisociates the layering of frontal planes with the stratification of façade planes in the 1927 Villa Stein at Garches (Krauss, 1987); to Aldo Rossi's 'Coffee and Tea Piazza' (1983) designed for Alessi which bisociates coffeepot and teapot on a tray with buildings in a piazza (Starkey, 2012); to Frank Gehry's 'Fish Lamps' (1984 & 2013) which bisociates the scales of the carp(to-be- gefilte-fish) with the broken shards of the ColorCore Formica laminates (Goldberger, 2021), or the form of the live carp which bisociates with his 1992 Olympic Fish Pavilion in Barcelona and 1997 Guggenheim Museum Bilbao; and finally to Greg Lynn's 'Folding in Architecture' which bisociates the folding in his architectural curvilinearity with the folding in culinary theory of mixtures types (Lynn, 1993), all four architects have shown us that bisociation between food and architecture is a fertile ground for the theory and practice of architectural creativity. Conversely, it echoes what Antonin Carême, founder of modern French 'grande cuisine', expressed in the early 19th century that, "Most noble of all the arts is architecture, and its greatest manifestation is the art of the pastry chef." The research is thus following a trajectory of an inherent bisociational potential between food and architecture -- a novel and important (yet neglected) area of research. The work is even more significant now given the recent rise of powerful concept-blending generative text2image AI systems being assimilated into architecture.

A key premise when conducting architecture research in generative artificial intelligence has been that the implicitly creative intuition of architects might one day be explicated computationally. In view of the discipline's apprenticeship-based historical origin and today's worship of genius starchitects' status, deep neural networks (if trained with sufficient architectural datasets) are being plausibly tasked to not only unravel the inner working of such creativity, but even democratize architectural expertise. Yet, how might one conduct research in exploring this elusive nature of architectural creativity? By focusing on a well-known aspect of creativity and situating it between food and architecture, the paper aims to interrogate and evaluate the bisociative affordance of today's deep generative models. As a means of evaluating the experimental results, the research defines 'successful' generated images as those showing a good representation of the text prompt and a good degree of freedom of interpretation. Five key parameters have been set up in the experimental framework to ensure a methodical process, namely, architectural terms, architectural styles, architects, colours, and forms. A 'food2archtecture' workshop was held twice during and after the initial research (one in mid-2022 and another in early-2023) comprising of 150 first year design students and pre-university non-design students at Singapore University of Technology and Design. A selection of their outputs is included here for evaluating the effectiveness of the proposed experimental framework.

2. Methods & Results

Five different models are identified for the experiments, namely, CycleGAN, VQGAN-CLIP, DALL-E 2, Midjourney and DreamFusion. The following sections describe the inputs and outputs for the respective models to compare their behaviours, and thus their potential use for bisociating architecture and food.

2.1. AI MODEL 1 (CYCLEGAN)

The training dataset comprises of images and texts of food and architecture. The food dataset is scraped from the websites "Great British Chef" and "Great Italian Chef", whereby each of these professionally photographed images is further curated based on criteria such as optimal image zoom level, sculptural formal features, and images with discernible themes. A total of 918 images are cropped and resized to 256 x 256 pixels and each contains accompanying text information. The architecture dataset is taken from the Harvard Dataverse, specifically, the project titled "Annotated Image Database of Architecture", which is a repository of over 35,000 architectural photographs taken worldwide and each labelled with a vast list of hierarchical categories and a series of auxiliary annotations. Among them, three criteria are used to curate the images, namely, exteriors, optimal image zoom level and sculptural formal features. The text description for each building is stored and retrieved from its accompanying Archdaily link. The final set of architecture images is of a similar quantity (1000) and resolution (256 x 256) as those in the food dataset. Figure 1 shows samples of the generated food images as outputs when images of buildings are used as inputs with a cycleGAN model trained for 500 epochs. It can be observed that buildings with distinct colours or clearly contrasting textures and landscaping are the most successful ones. For example, the pinkish metal cladding of the building in the first column of Figure 1 is translated into salmon, the metallic bracing across the façade of the building in the second column of Figure 1 is translated into grilled potatoes, and bamboo structures of the building in the tropics in the last column of Figure 1 is translated as steaks.

Figure 1. Different inputs on the top row and their CycleGAN outputs in the bottom row.

2.2. AI MODEL 2 (VQGAN-CLIP)

There are 4 VQGAN-CLIP models (Crowson et al., 2022) each pretrained with a different dataset, namely, ImageNet 16384, COCO, Gumbel 8192 and Wikiart 16384. It uses the CLIP model (Ramesh, et al., 2021) to guide a VQGAN model (Esser et al., 2021) in generating images with text prompts alone.

Figure 2 shows samples of the generated food images as outputs when two different set of 5 keywords are used, with the 'gingerbread crumbling creation fainthearted excess' for the first row, and 'swelteringly lashings freeze balsamic bowl' for the second row. These sets of keywords are extracted with TF-IDF (term frequency-inverse document frequency) from the original text of the food dataset used in the preceding cycleGAN model. Unfortunately, the results from this VQGAN-CLIP model are the least successful despite the use of 4 variants of the model.

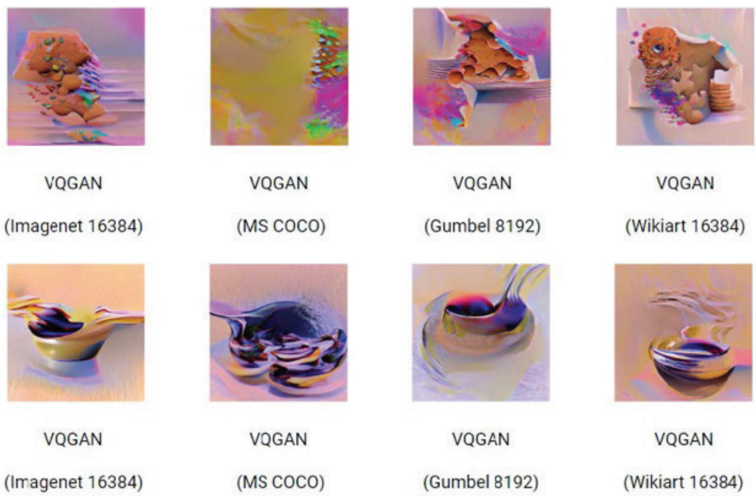


Figure 2. VQGAN-CLIP results using 4 different pretrained models shown in each column. Each row shows the comparative results from the same text prompt.

2.3. AI MODEL 3 & 4 (2D DIFFUSION MODELS)

Two main pretrained text-to-image AI models are used for this section -- Midjourney (midjourney.com) and DALL-E 2 (Ramesh et al., 2022). Generally, DALL-E 2 tends to return images that are more literal and abiding to the prompt, whereas Midjourney has a greater freedom of interpretation and expression of the prompt. To facilitate the ease of evaluating the inputs and outputs of the experiment, a limited vocabulary of words from the food and architecture domains, and the following prompt structure are used -- (Food Base + 'architecture') and (Food 1 + Building Element 1).

2.3.1. With or Without Architectural Terms

Prompt Structure: (Matcha Cake + ‘Rotunda'), (Pudding Roof), (Chocolate Columns)

The omission of the word “architecture” or any other architectural terms such as “rotunda” made a huge difference in the image generation of Midjourney. As shown in the third and fourth rows of Figure 3, when “rotunda” is omitted, Midjourney returns literal matcha cake as outputs, while DALL-E 2 still returns architectural images. DALL-E 2 is thus observed to be more robust than Midjourney in inferring from the other prompts (i.e., “roof” and “columns”) and interpreting them as architectural terms for generating architectural images.

Figure 3. Results from Midjourney (1st & 2nd columns) and DALL-E 2 (3rd & 4th Columns). The same text prompt "Matcha cake rotunda, with pudding as roof, chocolate columns" is used for all the 4 rows, with the exception that the word "Rotunda" is omitted for the 3rd and 4th rows.

2.3.2. With or Without Architectural Styles

Prompt Structure: (“Cheesecake” Architecture, Strawberry roof, Whip cream facade, “X" Architecture style)

Architectural style is explored by adding a 4th parameter (“X" Architecture style) to the original base prompt structure. For example, Figure 4 uses the same text prompt where 'X' is the only varying parameter such as 'Ornate', 'Modernist', 'Contemporary', 'Brutalist', 'Classical', and 'Renaissance'.

It is observed that the bisociation capacity of Midjourney in generating more legible architectural style might be constrained by the specificity of the food vocabulary used. For example, the materiality of “Strawberry Roof" and "Whip cream façade" might have resulted in the softness and smoothness that are highly uncharacteristic of the "Brutalist" and "Modernist" styles.

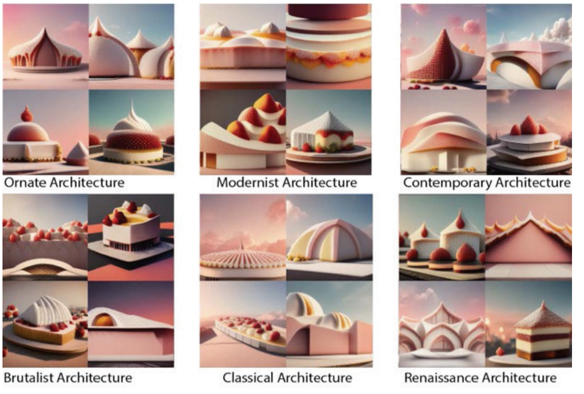


Figure 4. Results from Midjourney using the same text prompt "Cheesecake Architecture, Strawberry roof, Whip cream façade, 'X' Architecture style", where 'X' is the only varying parameter such as 'Ornate', 'Modernist', 'Contemporary', 'Brutalist', 'Classical', and 'Renaissance'.

2.3.3. With or Without Architects

Prompt Structure: ("Cheesecake" Architecture, Strawberry roof, Whip cream facade, designed by "X" Architect)

This test is conducted to identify potential bias that Midjourney might have in generating the signature architectural features of famous architects. In Figure 5, architects who are more well-known for their sculptural forms (e.g., Zaha Hadid) yielded higher accuracy in the generated images. Yet, in the case of Tadao Ando who is known for his raw concrete and basic geometrical expression, the generated images are less representative, given the red, pink, and white hues derived from the prompt's "strawberry" and "whip cream". In this context, Midjourney might be more performative in referencing architects whose works are sculptural and colourful when bisociating with food concepts.

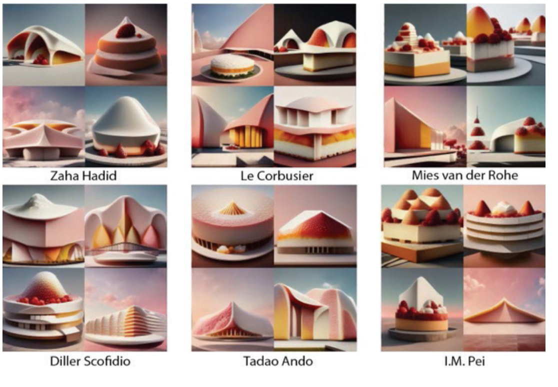


Figure 5. Results from Midjourney using the same text prompt "Cheesecake Architecture, Strawberry roof, Whip cream façade, designed by "X" Architect", where 'X' is the only varying parameter such as 'Zaha Hadid', 'Le Corbusier', 'Mies van der Rohe', 'Diller Scofidio', 'Tadao Ando', and 'I.M. Pei'.

2.3.4. With or Without Colours

Prompt Structure: (‘ ___ ' Pavilion) , (Rock Sugar Facade) , (Sugar glazed Roof)

Colour plays a huge role in changing the mood and representation of the generated image. The prompt structure above defines the first part as the colour definer, where desserts with distinct colours are used as the base. The subsequent 2 parts are fixed to ensure non-aggressive assertion of certain colours to the image but be guided by the colour definer itself. While the colour of the base food item did come across very strongly in all of the images, the rock sugar and sugar texturing influenced the form much greater than the base (“cake”, “pudding”, “pie”).

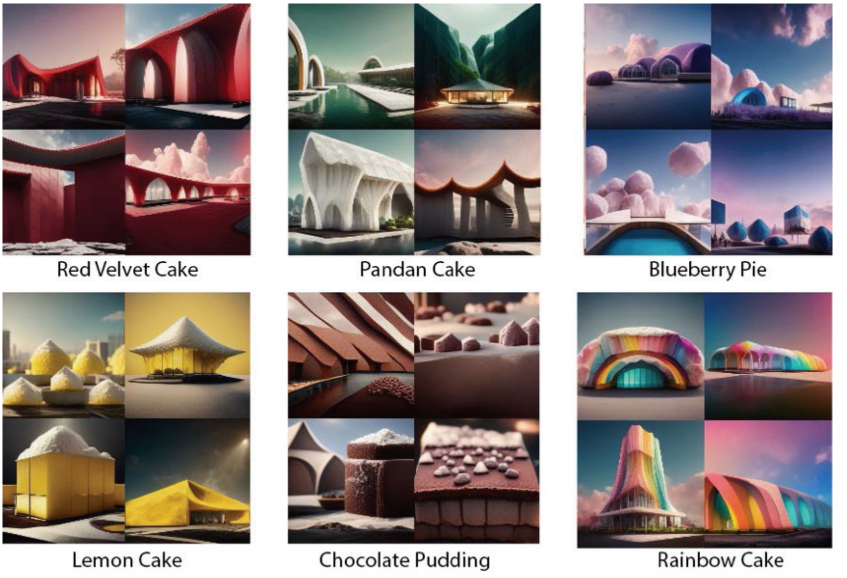


Figure 6. Results from Midjourney using the same text prompt but varying the colours via food colour associations, such as (clockwise from top-left), red velvet cake, pandan cake, blueberry pie, rainbow cake, chocolate pudding and lemon cake.

2.3.5. With or Without Forms

Prompt Structure: (“Cheesecake” Architecture, Rigid Food Sculptural Element, Sculptural Food Rigid Element)

Midjourney is tested to verify if there is a greater bias towards sculptural architectural forms or food item prompt. The food items chosen are selected to be white to prevent colour from influencing the form. For example, rigid food includes rock sugar, biscuit, white chocolate, cheesecake; sculptural food includes whipped cream, melted cheese, melted white chocolate, cream; rigid element includes column, beams, roof, wall; and sculptural element includes roof, facade, and landscaping. In Figure 7, it is observed that Midjourney will automatically generate more sculptural forms if no specification is made.

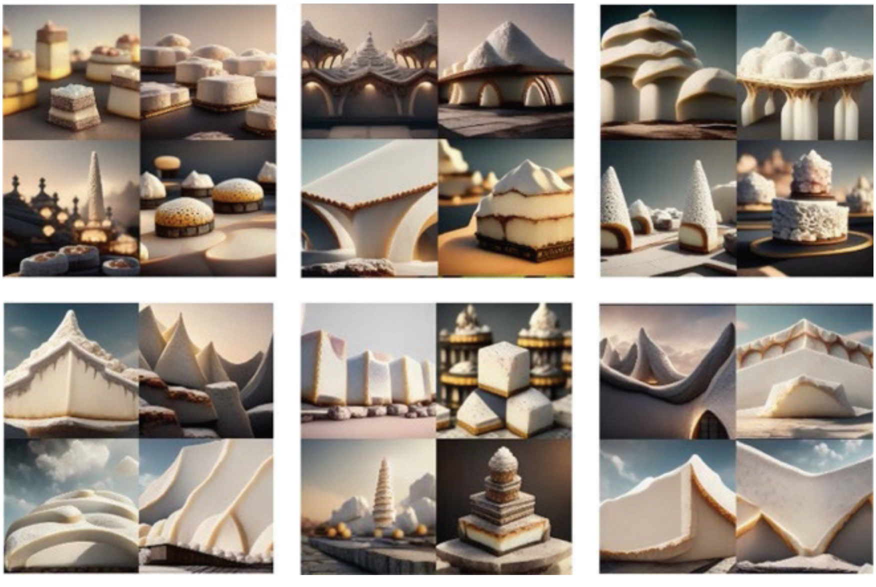


Figure 7. Results from Midjourney to decouple formal basis from food and architecture.

2.3.6. 'Hansel & Gretel' Workshop

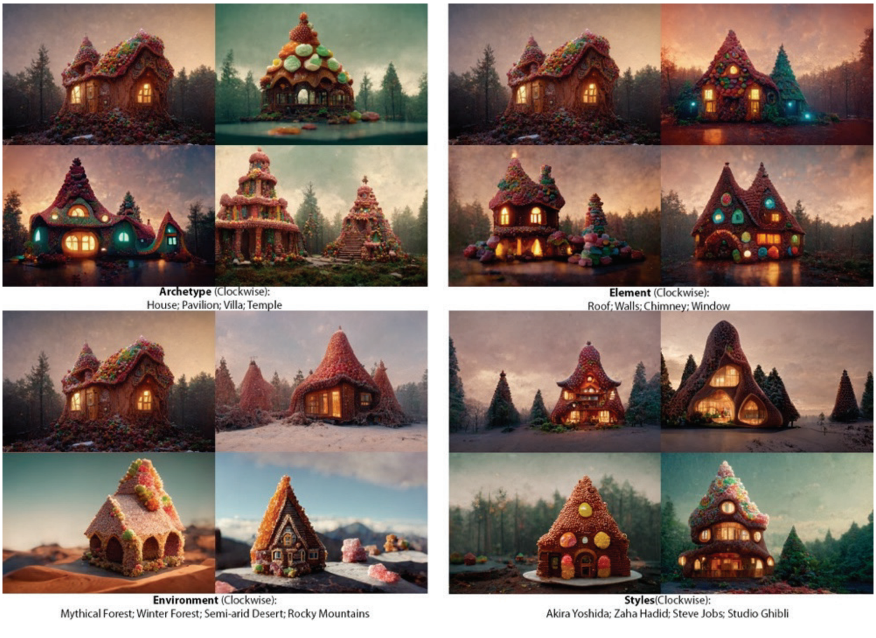


Figure 8. Results from a second workshop titled 'Hansel & Gretel' designed by the author to not only explore the affordance of Midjourney as a generative design tool, but also how such a text-to-image design process might be formulated in an educational context.

Thus far, the results are based on the outputs of the first workshop. A second workshop titled 'Hansel & Gretel' is designed with a clearly stipulated task, that is "to design an edible cottage worthy of enticing Hansel and Gretel into the witch's lair". For facilitating experimentation/discussion/evaluation in a more coherent manner, specific text prompt structures are again crafted for the participants to make their own variations. It is structured as follows: (1) Building Archetypes, (2) Building Elements, (3) Building Environments, (4) Building Styles, and (5) Standardised Formatting. An example of such a prompt might be "House made of gingerbread, sugar candy windows, in a mythical forest, designed by Studio Ghibli, hyper-realistic, octane, unreal engine render, –ar 3:2". Figure 8 shows the generated results from modifying only the first 4 parts of the same text prompt structure. The top-left array of four images are generated based on text prompt variation of building archetypes such as house, pavilion, villa and temple; the top-right array of four images are generated based on variation of building elements such as roof, wall, chimney and window; the bottom-left array of four images are generated based on variation of building environments such as mythical forest, winter forest, semi-arid desert and rocky mountains; and the bottom-right array of four images are generated based on variation of building styles such as Akira Yoshida, Zaha Hadid, Steve Jobs and Studio Ghibli.

2.4. AI MODEL 5 (3D DIFFUSION MODELS)

The last experiment explores the use of a text2image model called Stable Diffusion (Rombach et al., 2022) for generating 3D models with DreamFusion (Poole et al., 2022). Figure 9 shows the denoising sequence (from left to right) when generating a 3D volume using the DreamFusion model with the simple text prompt of "A Gingerbread House". Food features such as blobby gummy bears, colourful sugar candies, dripping caramel, icing sugar…etc., and architecture features such as windows, doors, doorframes, columns, roof, shop signature, and entrance porch…etc., are not only recognisable, but quite smoothly blended together -- convincingly looking like a gingerbread house. While the 3D bisociated outputs are considerably successful, the inference time is unfortunately long, in the range of hours, rather than seconds or minutes. This could be detrimental when appropriated in a real-time design session as compared to text-to-image models like Midjourney or DALL-E 2.

Figure 9. Denoising sequence (from left to right) when generating a 3D volume using the DreamFusion model with the simple text prompt of "A Gingerbread House".

3. Conclusion

The bisociation afforded by text-to-image diffusion models such as Midjourney has proven to be at an extremely low entry level, yet surprisingly productive, for both designers and non-designers. Unlike other models such as CycleGAN which still requires a training set, VQGAN-CLIP which is less capable of generating high-fidelity imagery, DALL-E 2 which is often too literal in its interpretation, and DreamFusion which suffers from lower resolutional outputs and long inference time, Midjourney stands out as a good candidate for generating and illustrating design bisociations between food and architecture. Based on the interaction with the two workshops' participants, users are at ease in conceiving text prompts and selecting the desired outputs among the array of outputs generated, even without a clue of Midjourney's blackbox. Future works will explore specific gastronomic properties of food with those of architecture through bisociating notions such as textures, structures, and compositions, but using non-blackbox diffusion models instead.

References

Grimm, J. (2020) Grimms' Fairy Tales: Translator by Edgar Taylor and Marian Edwardes. Amazon Digital Services LLC - KDP Print US.

Koestler, A. (1964) The act of creation. Oxford, England: Macmillan (The act of creation), p. 751.

Jacobsen, T. (2006) 'Bridging the Arts and Sciences: A Framework for the Psychology of Aesthetics', Leonardo, 39(2), pp. 155–162. Available at: https://doi.org/10.1162/leon.2006.39.2.155.

Michel, C. et al. (2014) 'A taste of Kandinsky: assessing the influence of the artistic visual presentation of food on the dining experience', p. 11.

Horwitz, J. and Singley, P. (eds) (2006) Eating Architecture. Cambridge, Mass.: The MIT Press.

Liaropoulos-Legendre, G. (2011) Pasta by Design. Thames & Hudson.

Krauss, R. (1987). Death of a Hermeneutic Phantom: Materialization of the Sign in the Work of Peter Eisenman. In P. Eisenman (Ed.), Houses of Cards (pp. 166–184). Oxford University Press.

Starkey, S. (2012). Collapse to Expand: Alessi's Tea and Coffee Piazza. https://www.acsa-arch.org/chapter/collapse-to-expand-alessis-tea-and-coffee-piazza/

Goldberger, P. (2021, June 2). Frank Gehry: Fish Lamps | Essay. Gagosian Quarterly. https://gagosian.com/quarterly/2021/06/02/essay-frank-gehry-fish-lamps/

Lynn, G. (1993). Architectural Curvilinearity: The Folded, The Pliant And The Supple. In Lynn, Greg (Ed.), Folding in Architecture (Vol. 63, pp. 24–31). Academy Editions.

Greenway, P. (1987). The Belly of an Architect. Callender Company, Mondial, Tangram Film.

Zhu, J.-Y. et al. (2020) 'Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks'. arXiv. Available at: http://arxiv.org/abs/1703.10593

Crowson, K. et al. (2022) 'VQGAN-CLIP: Open Domain Image Generation and Editing with Natural Language Guidance'. arXiv. Available at: https://doi.org/10.48550/arXiv.2204.08583.

Ramesh, A. et al. (2022) 'Hierarchical Text-Conditional Image Generation with CLIP Latents'. arXiv. Available at: https://doi.org/10.48550/arXiv.2204.06125.

Saharia, C. et al. (2022) 'Photorealistic Text-to-Image Diffusion Models with Deep Language Understanding'. arXiv. Available at: https://doi.org/10.48550/arXiv.2205.11487.

Esser, P., Rombach, R. and Ommer, B. (2021) 'Taming Transformers for High-Resolution Image Synthesis'. arXiv. Available at: https://doi.org/10.48550/arXiv.2012.09841.

Poole, B., Jain, A., Barron, J. T., & Mildenhall, B. (2022). DreamFusion: Text-to-3D using 2D Diffusion (arXiv:2209.14988). arXiv. https://doi.org/10.48550/arXiv.2209.14988.

Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). High-Resolution Image Synthesis with Latent Diffusion Models (arXiv:2112.10752). arXiv. https://doi.org/10.48550/arXiv.2112.10752

PREDICTING AMENITIES DISTRIBUTIONS FOR WORKERS FROM THE BUILT ENVIRONMENT BASED ON MACHINE LEARNING

HONGYU WAN[1], ANQI PAN[2], YANWEN XUE[3] and HAO ZHENG[4]
[1,3]Tongji University
[2]Polytechnic University of Milan
[4]City University of Hong Kong
[1]hongyuwan@tongji.edu.cn, 0000-0003-0067-8287
[2]anqipanpolimi@gmail.com, 0000-0003-0676-774X
[3]yanwenxue013@gmail.com, 0000-0003-2803-8337
[4]zhhao@alumni.upenn.edu, 0000-0001-5769-6035

Abstract. The working population has increased in cities with urbanization. Providing a supportive built environment with reasonable amenities distribution for them is becoming more important. Previous GIS-based approaches to urban planning for this issue tend to be subjective with high labour costs. This paper uses the generative adversarial network (GAN) to explore the relationship between amenities distributions and urban morphology, thus effectively predicting and visualizing the ideal amenities distributions in fast-growing cities based on the condition of well-developed megacities. In this research, we take Shanghai, one of the global cities in China with a big labour market, as the research sample. First, we use the Point of Interest (POI) data to draw the heatmap of urban amenities that support workers' daily life and collect the corresponding city maps. Then, we cut them into hundreds of image pairs as the training set and train a GAN model for predicting the future amenities distributions in other cities. To implement the model, we further collect the city maps of Jiaxing, one of the second-tier cities near Shanghai, as the testing set. Results show that our trained model can accurately predict amenities distributions for its future. The GAN-based prediction could effectively support detailed urban planning.

Keywords. Machine Learning, Big Data Analysis, Point of interest, Urban Planning, Amenities Distributions

1. Introduction

1.1. AMENITIES DISTRIBUTIONS: AN IMPORTANT FACTOR IN PLANNING

With fast urbanization in China, urban agglomeration is gradually becoming the trend in China with a massive influx of people into cities to work. Several studies have identified that urban amenities, such as shopping, restaurants, and transportation

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 19-28.

system, have a close relationship with employment distribution (Clark et al., 2002; Florida, Mellander and Stolarick, 2008; Li, Wei and Wu, 2019). Many cities have long-term plans for urban development, which entail a particular emphasis on urban amenities. In other words, ensuring a reasonable amenities distribution to meet the needs of the changing demographics is one of the important tasks in urban planning.

1.2. THE CHALLENGES OF AMENITIES PLANNING FOR WORKERS

To plan diverse businesses and services to meet prospected demands of urban life, the department of planning has utilized geographical information system (GIS) since the early 1980s (Longley and Clarke Clarke, 1996; Yeh, 1999). More specifically, Planners apply models to analyze geo-physical and socioeconomic variables, such as people behaviors, land use, existing amenities, and so on, to provide information for decision-making in planning (Stillwell and Clarke, 2004; Parry et al., 2018). However, two limitations exist in the traditional approach: 1) planning decisions are mainly subjective based on the experience of planners, because analysis of the present patterns of fluctuating factors might not able to predict the demands for the future directly; 2) data-informed planning is time-consuming and high labor-cost, usually taking years. On the other hand, the material form of the city has unusual durability, which always adapts to new economic needs and reflects the fashion of the time (Kostof, S, 1991; Kostof and Castillo, 1992). Therefore, addressing these two problems above requires exploring an efficient approach to finding out the relationship between reasonable amenities distribution and urban morphology.

1.3. NEW POTENTIALS OFFERED BY MACHINE LEARNING METHODS IN AMENITIES PLANNING

Generative Adversarial Networks (GAN), a kind of artificial intelligence algorithm, can study training examples and learn the probability distribution that generated them (Goodfellow et al., 2014). Recent studies on image-to-image translation have identified that such GAN can synthesize pictures from label maps very effectively (Isola et al., 2017). Following this investigation, many researchers have used GAN to help plan for the future by predicting certain factors in the city, such as crime rates (He and Zheng, 2021), residents' health status (Cao and Zheng, 2021), and short-term passenger flow on rails (Li et al., 2022). Therefore, machine learning can help shape a better-built environment for workers by generating reasonable amenities distributions efficiently based on a city map.

1.4. USING GAN TO SUPPORT AMENITIES PLANNING FOR WORKERS

In this paper, we suppose that well-developed big cities with large populations have a certain reference value of amenities planning for many growing cities with similar geographical locations. For example, Shanghai and Jiaxing are two cities in the Greater Shanghai metropolitan area. Shanghai is one of the top global cities in China with a big labour market with a relatively mature built environment for workers, while Jiaxing is a fast-growing city with an increasing population nearby. Therefore, we use GAN to explore the relationship between amenities distribution and urban morphology in Shanghai, and then predict the distribution of amenities in Jiaxing for the planning for the future. More specifically, we take the city map and the heatmap of amenities for workers in Shanghai to train a GAN model, and then feed the city map of Jiaxing into the model to generate the predicted heatmap of amenities.

2. Methodology

2.1. DATA COLLECTION

2.1.1. The collection of city map

To obtain a high-definition map reflecting the built environment in Shanghai, we collect information on buildings, roads, green space, and waters and then we label them in different colors from the open platform API of BDmap by python (Shen et al., 2020). Labels function as supervisory signals that guide the network in processing, extracting, and transforming visual information to achieve maximum performance on the required tasks (Figure 1).

Figure 1. The collection of city maps

2.1.2. The generation of amenities heatmap

In this research, "Amenities for workers" refers to urban businesses, services, and transportation systems that contribute to the daily life of workers. More specifically, subways and buses help them transport from home to work, convenience stores provide everyday items, and restaurants enable them to have meals. Thus, we collect the Point of Interest (POI) data of subway stations, bus stops, convenience stores, and restaurants from AMaps via API in the form of CSV files in the corresponding area of the city map (Zhang et al., 2022). Then, we use Rhino and Grasshopper to visualize the geospatial distribution of the large amount of data as a heatmap (Figure 2).

According to legal regulations and planning standard about those amenities in China, the service radius is 500 meters for subway stations and restaurants or 300 meters for bus stops and convenience stores. To reflect the service area of each amenity, we divide the research range into 25X25m grid squares, and generate points on the nearest grid intersections to their coordinates based on the latitudes and longitudes of all amenities in CSV files with four colours representing the type. We suppose that the number of one kind of amenities, which include one of four vertexes of a square in their service area, reflects the density of this kind of amenities in the area of the square, and the highest density determines the density of the square. Therefore, we label each square in the colour of the densest points and opacity based on the density. Eventually, we obtain the heatmap of four kinds of POI with a gradual change in colours.

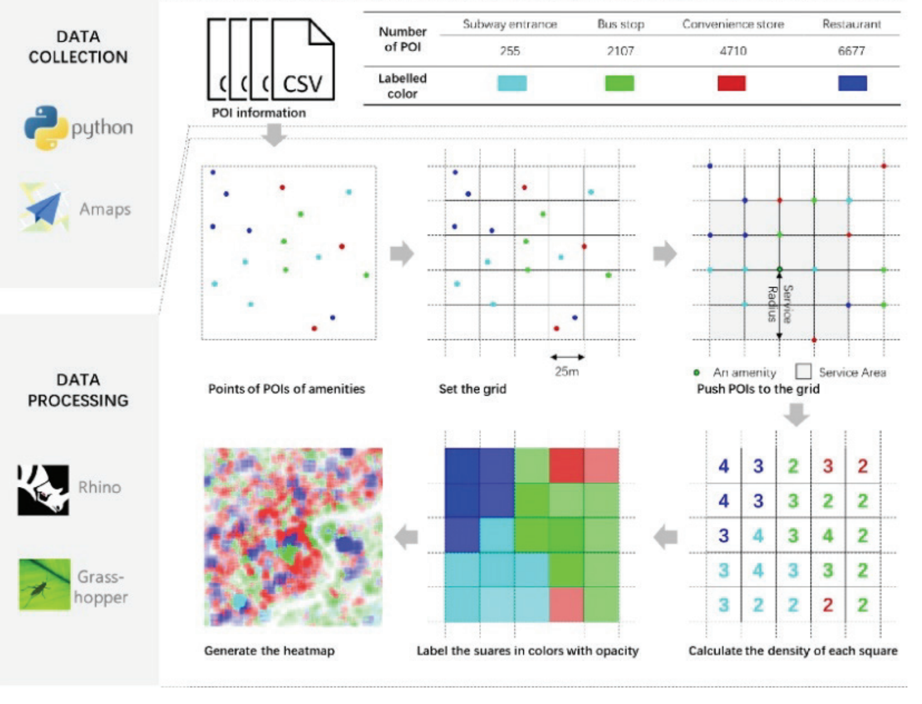


Figure 2. The generation of amenities heatmap

2.2. GENERATIVE ADVERSARIAL NETWORK

In this paper, we use Pix2PixHD (Wang et al., 2018), an image-based Generative Adversarial Network (GAN) model, to give an amenity-distribution prediction based on a given map of the city area. The GAN takes city maps as input and commuting amenities heat maps as output and trains the network to recognize the relationship between two sets of images.

The GAN consists of two neural networks, Generator (G) and Discriminator (D). G takes the city map as input (Train_A in Figure 3) to generate predicted amenities distribution (Predicted_B in Figure 3). D takes real or generated amenities distribution as input and attempts to distinguish the generated amenities distribution from the real one. G and D are trained alternately and compete with each other. D rates the authenticity of the two sets of image data and outputs probability values. D minimizes the cross-entropy of the actual and the expected output value while G maximizes the probability that the generated data is true by adjusting the parameters based on the feedback from D. The training process gradually improves the accuracy of G and D. When D eventually cannot distinguish between the real image and the generated image, G is considered to be optimal.

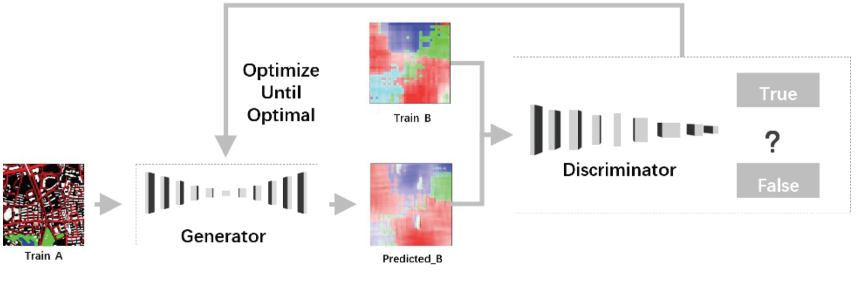


Figure 3. The principles of the GAN model

2.3. TRAINING PROCESS

First, we resize the city map and corresponding heatmap of amenities into 5120X5120 to prepare the training set and testing set. For the testing set as representatives of different regions of the city, we selected one vertical column and one horizontal column with a width of 512 pixels from the middle of the entire data area and cut them into small images with a resolution of 512X512 using Python Image Library (PIL). In terms of the training set, we cut the rest four parts of the map into small pieces of the same size by shifting 128 pixels after each cut. (Figure 4) This operation avoids missing the data on the edge of each small image and provides enough sets of uniform and suitable sizes for machine learning. After processing the heatmap in the same way, we obtain 900 pairs of images as the training set and 19 pairs as the testing set.

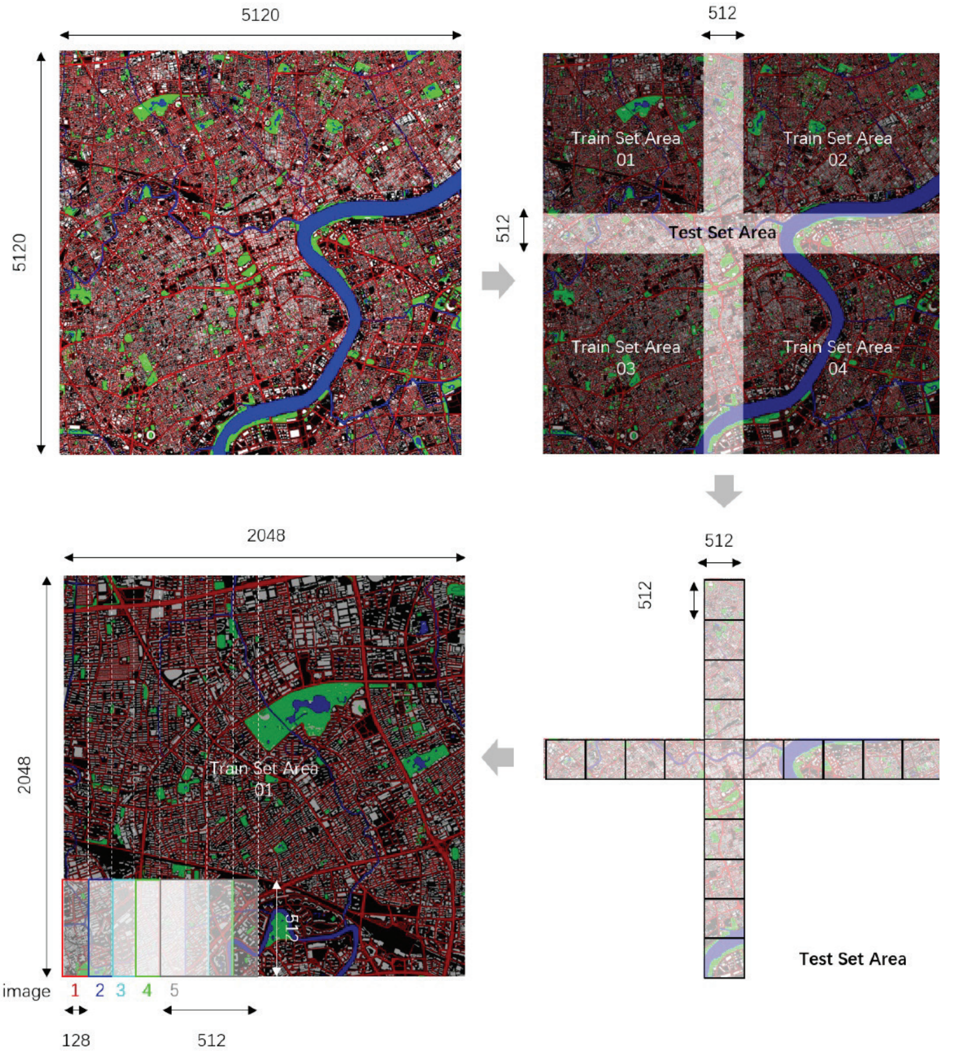


Figure 4. Preparation of training set and testing set

Then, high-resolution image-to-image translation with conditional adversarial networks is used as the main operating algorithm. For better convergence, we use the constant training gradient in the first 280 epochs of training, and the gradient decreases in the next 120 epochs, which is conducive to more precise fitting of data when the training is nearly finished. The saved temporary results are recorded on the monitoring panel (Figure 5). When the model is not fully trained, the accuracy of the predicted facility distribution in the generated image is low. Later, at about 115 epochs, the accuracy of generated images is improved to a certain extent, but some concentrated facilities in the areas cannot be predicted by the neural network. After 400 epochs of training, the synthesized image performs accurately, with the facility heatmap showing a clear pattern corresponding with different types of functional areas. Therefore, we decided to stop training after 400 epochs and store the prediction model at this point as the final model. The loss value of the generator and discriminator are recorded during the training process Figure 6 compares the loss values of the generator and discriminator

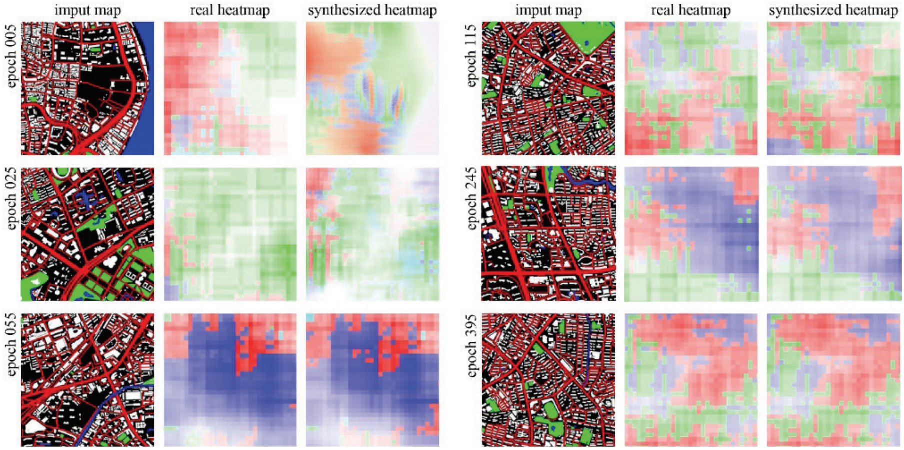


Figure 5. Synthesized results for each training epoch

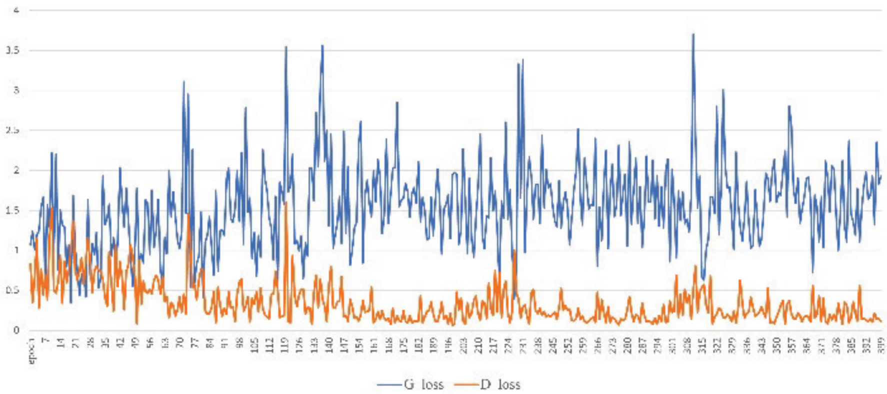


Figure 6. Generator loss (G_LOSS) and Discriminator loss (D_LOSS) during training

3. Results

3.1. ACCURACY OF THE TRAINING SET

To evaluate the validity of the trained model in practical prediction, we implement the numerical value of the FID (Fréchet Inception Distance), which indicates the diversity and quality of the generated images (Heusel et al., 2017). The lower the FID score is, the closer the generated and the real images are, and the better the validity of the model is.

In this paper, we calculated the FID scores between the real and synthesized images, and between the real and random images. First, we use PIL to generate two random sets as the reference, with the same number of images as the training and the testing sets respectively in the uniform size of 512x512. As Figure 7 shows, the FID score of the trained images is 11.2, much lower than the random set (479.4), while the FID score of the tested images is a little bit higher (207) but still much lower than the random one.

The result indicates the feasibility of the model, which has a certain ability to predict the distribution of amenities for workers. But it is far from perfect. We speculate that the relationship between the convenience of commuters and urban morphology is still complicated and disturbed by many factors.

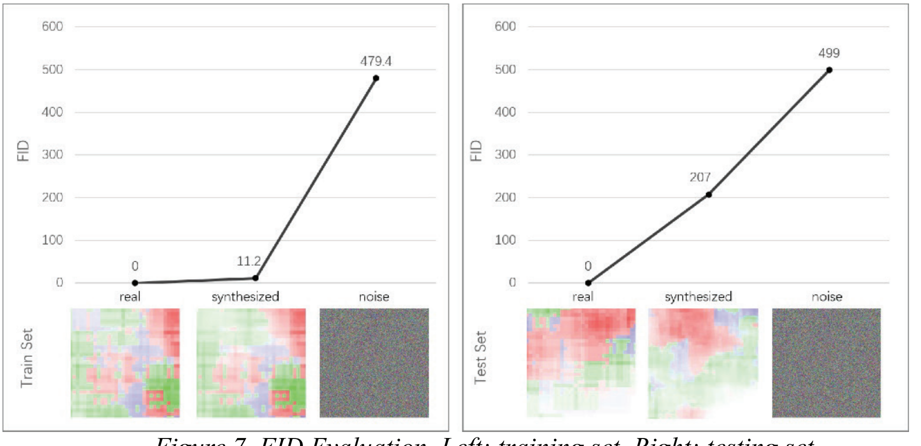


Figure 7. FID Evaluation. Left: training set. Right: testing set

3.2. PREDICTION OF AMENITIES DISTRIBUTION

We implement the model on Jiaxing, the representative of the second-tier cities around Shanghai. As the population increases, Jiaxing is carrying out a lot of urbanization construction to provide enough urban amenities for newly arrived labour force. Because of the similar geographical locations, using the model to predict amenities distributions based on Shanghai might offer a significant reference for planning in Jiaxing.

According to the direct observation and comparison, we find that Jiaxing should add more convenience stores and restaurants and reduce some bus stops in the future to cater for a larger population of workers. To be more specific, we choose city map images of similar urban texture in Jiaxing and Shanghai, and then find some rules behind the amenities distributions and city map learned by the model from Shanghai (Figure 8). Two typical areas show the capability of the model to advise on amenities distribution based on underlying rules in Shanghai. One has large buildings with empty land, where industrial parks are usually located. As predicted, convenience stores increase around buildings and restaurants increase on empty land, which is identified by the similar area of Shanghai. The other has small and dense buildings along a river, which is usually a residential district. According to the prediction, convenience stores increased into a stretch along the river, similar to the distribution by the river of

Shanghai.

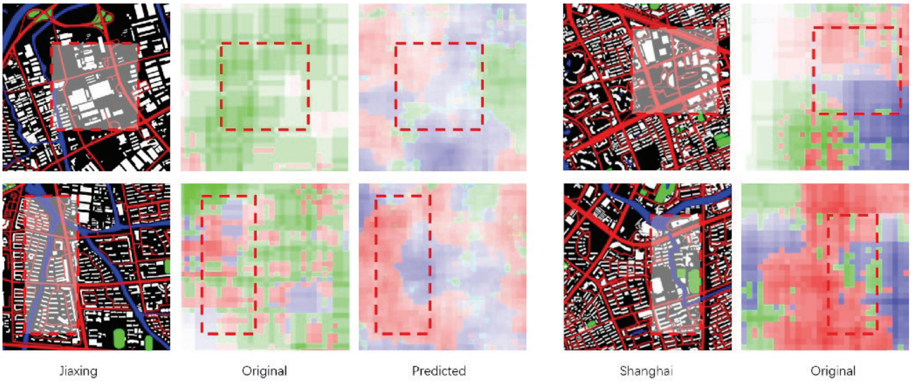


Figure 8. Generated results for predicting facility heatmap in Jiaxing

4. Conclusion

Under rapid urbanization, an increasing number of workers flood into cities to work, which leads to the importance on the planning of amenities for them. The machine learning model in the research with big data brings new potential to urban planning for them. It can not only reveal the principles of inter-relationship between the urban morphology and other city factors, but also help planners to make quick and reasonable decisions in fast-growing cities by providing evidence in the megacities nearby. Compared with previous methods, this approach avoids subjective evaluation and saves labour and time. Taking the amenities for example, the results show the ability of the method to predict the distribution of amenities in urban areas, thus guiding appropriate interventions in future planning.

Despite the contributions, the research has some limitations. First, we overlooked some dimensionality of the context, such as the time of data collection, the difference between city form and economics of two cities. Some inexplicable result of prediction indicates the sample of the city map in Shanghai may not cover all the urban textures in Jiaxing. Second, the predicted heatmap of amenities with low resolution fails to support planning more specifically. Our subsequent research will aim to use more samples to train more accurate machine learning models and add details to the heatmap.

References

Cao, S., & Zheng, H. (2021). A Poi-based machine learning method for predicting residents' health status. Proceedings of the 2021 DigitalFUTURES, 139–147. https://doi.org/10.1007/978-981-16-5983-6_13

Clark, T. N., Lloyd, R., Wong, K. K., & Jain, P. (2002). Amenities drive urban growth. *Journal of Urban Affairs,* 24(5), 493–515.https://doi.org/10.1111/1467-9906.00134

Florida, R., Mellander, C., & Stolarick, K. (2008). Inside the black box of regional development—human capital, the creative class and tolerance. *Journal of economic geography,* 8(5), 615-649. https://doi.org/10.1093/jeg/lbn023

Goodfellow, I., et al.(2020). Generative Adversarial Networks. Commun. ACM, 63(11), 139–144. doi:10.1145/3422622

He, J., & Zheng, H. (2021). Prediction of crime rate in urban neighborhoods based on machine learning. *Engineering Applications of Artificial Intelligence,* 106, 104460. https://doi.org/10.1016/j.engappai.2021.104460

Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., & Hochreiter, S. (2017). Gans trained by a two time-scale update rule converge to a local nash equilibrium. *Advances in neural information processing systems,* 30. https://doi.org/10.48550/arXiv.1706.08500

Isola, P., Zhu, J. Y., Zhou, T., & Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1125-1134). https://doi.org/10.1109/CVPR.2017.632

Kostof, S. (1991). *The city shaped* (pp. 9-39). Little, Brown and Company. https://doi.org/10.1080/03612759.1992.9950662

Kostof, S., & Castillo, G. (1992). *The city assembled: the elements of urban form through history.* Thames and Hudson

Li, H., Wei, Y. D., & Wu, Y. (2019). Urban amenity, human capital and employment distribution in Shanghai. *Habitat International,* 91, 102025. https://doi.org/10.1016/j.habitatint.2019.102025

Li, H., Zhang, J., Yang, L., Qi, J., & Gao, Z. (2022). Graph-GAN: A spatial-temporal neural network for short-term passenger flow prediction in urban rail transit systems. *arXiv preprint arXiv:2202.06727.*https://doi.org/10.48550/arXiv.2202.06727

Longley, P. A., & Clarke, G. (Eds.). (1996). *GIS for business and service planning*. John Wiley & Sons.

Parry, J. A., Ganaie, S. A., & Bhat, M. S. (2018). GIS based land suitability analysis using AHP model for urban services planning in Srinagar and Jammu urban centers of J&K, India. *Journal of Urban Management,* 7(2), 46-56. https://doi.org/10.1016/j.jum.2018.05.002

Shen, J., Liu, C., Ren, Y., & Zheng, H. (2020). Machine learning assisted urban filling. CAADRIA 2020 (pp. 681-690). https://doi.org/10.52842/conf.caadria.2020.2.679

Stillwell, J., & Clarke, G. (Eds.). (2004). *Applied GIS and spatial analysis* (pp. 254-255). Chichester: Wiley. ISBN: 978-0-470-84409-0

Yeh, A. G. (1999). Urban planning and GIS. *Geographical information systems,* 2(877-888), 1.

Zhang, Y., Zhang, Q., Zhao, Y., Deng, Y., & Zheng, H. (2022). Urban spatial risk prediction and optimization analysis of POI based on deep learning from the perspective of an epidemic. *International Journal of Applied Earth Observation and Geoinformation,* 112, 102942. https://doi.org/10.1016/j.jag.2022.102942

ELEMENTAL SABOTAGE

Diffusing Functional Morphologies

DONGYUN KIM[1], LLOYD SUKGYO LEE[2] and HANJUN KIM[3]
[1]*TWO Platforms Inc.*
[2]*Volume 64*
[3]*SoomeenHahm Design Ltd.*
[1]*archclojure@gmail.com, 0000-0003-1193-3845*
[2]*lsg.architecture@gmail.com, 0000-0001-6650-2829*
[3]*hanjun.kim@soomeenhahm.com, 0000-0002-4793-4818*

Abstract. What modernism has failed discuss is the sheer diversity of forms that can come with a singular function. Their order led to the systematisation of standardised architectural elements that are now widely used around the world. However, functions of architecture have manifested into a multitude of forms across different cultures. For instance, the roof of the Korean traditional architectural type, Hanok, can be easily differentiated from the Victorian gable roof. The function-form relationship in architecture has long been investigated, but there is a lack of objective tool or technique that directly contributes to categorisation or diversification of existing architectural forms. With the advent of novel Machine Learning techniques, especially Generative Adversarial Networks (GANs), architectural forms can now be traced with objective methodologies such as a probabilistic-based model. This paper seeks to increase architectural diversity, borrowing the objectivity of state-of-the-art machine learning techniques in three dimensions such as 3D GAN.

Keywords. Generative Adversarial Networks, 3D GAN, Morphology, Semiology, Form-finding.

1. Introduction

*For the clarity of the overall concept, the paper defines the following terms: *Function-Form*: Form in which an architectural function manifest *Functional Element*: Element that embodies a specific architectural function

What the modernist mantra "form ever follows function" failed to imply is the sheer diversity of forms that can come with a singular function. The pervasive modernist order led to the systematisation of standardised architectural elements that are now widely used around the world, eventually yielding look-alike buildings that can be considered contextless and universal. These generic buildings reinforce standards around how architectural functions manifest into a limited number of forms. Such

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 29-38.

architectural homogenisation converges diverse local architectural possibilities into a singular set of standardised forms. The generic side-hung window, which fails to connote any cultural meaning or proposes diverse forms that can arise from it, replaced Deul-uh-yeul-gae door (들어열개), Japanese Shoji (障子) or Louvred windows of the Mediterranean region. After defining architectural homogenisation as the prevalence of a limited number of forms that architectural functions manifest in (or function-forms), we attempted to put forth a design methodology to diversify homogenised function-forms utilising an existing machine learning technology. In order to yield a successful outcome of the newly proposed design methodology, it is important to address the following two questions: a) why is the diversity of function-forms important? and b) why is Machine Learning relevant in this methodology?

1.1. WHY IS THE DIVERSITY OF FUNCTION-FORMS IMPORTANT?

Diversity, at large, can be said to contribute to resilience. In ecological studies, it is promptly understood that biodiversity contributes to better resilience of ecosystem functions (Oliver et al., 2015). In Social-Ecological Systems (SES), Response Diversity (RD) of species to environmental changes is known to contribute to the resilience of ecosystems (Leslie & McCabe, 2013). Furthermore, work-unit diversity is said to help the resilience of the organisational framework in business practice (Duchek et al., 2019). As highlighted above, the diversity of constituent elements of a system often leads to resilience, where resilience can be understood as the ability of a system to withstand changes or difficulties. Architecture can, too, be understood within this framework. Baroque architecture can be identified with a common set of rules and elements that compose it. However, as seen with the Sicilian Baroque and Earthquake baroque in the Philippines, the formal changes in their constituent elements provided both social (Puleo, 2014) and structural resilience (Soliman, 2019). The buttresses of Paoay chruch in the Philippines, for example, still denote the function of support while bearing significant difference in shape and proportion to other buttresses that are found in the European continent. Though the concept of diversity has a general and wide scope, it becomes clear through the said examples, that creating different variations of forms from a common set of rules could contribute to diversity, and, furthermore, resilience. In this regard, the necessity for the diversity of function-forms in architecture can be sparingly justified.

1.2. WHY IS MACHINE LEARNING RELEVANT IN THIS METHODOLOGY?

Umberto Eco sheds light on the design process through A Theory of Semiotics, “Designer cannot make a new form functional without the support of existing processes of codification.” (Eco, 1986, p.178). In other words, newly designed forms must expand upon the logic of forms that already exist. This process bears a striking resemblance to how machine learning models produce output: common attributes among data entries manifest within newly generated designs, continuing the legacy of the existing. This similarity between the two processes creates a potent link between the existing understanding of design process within architecture and the machine learning technology, elucidating possibilities of utilising the technology in a new way that is more aligned with the existing discourse.

2. Methodology

According to Eco, "the form of the object must, besides making the function possible, denote that function clearly enough to make it practicable as well as desirable" (Neil Leach (Edt), 2014). While Eco couples the 'denoted' function of an object with its essential formal manifestation, he fails to further explore the diversity of function-form relationship beyond suggesting the changes to an object's connotative variations. With the advent of Artificial Intelligence (AI) and increased computational resources such as GPU and cloud, statistical approaches can be used to examine and enhance Eco's principle of codification. One of the main machine learning frameworks, Generative Adversarial Networks (GANs) (Goodfellow et al., 2014), bear major similarities to Eco's idea of denotation. GAN is a generative modelling approach that employs machine learning methods like convolutional neural networks. Generated output from GAN is made of different learned features from the original dataset extracted from an abstract space called latent space. This implies that GAN shows a sufficient ability to create and amplify diversity of the input dataset through reconfiguration.

The generator and the discriminator are the two parts of GANs that create a zero-sum game. The generator is trained to learn the distribution of input data and generate fake data, while the discriminator distinguishes fake data from input data with probabilistic calculations. If a set of images is used as training data, the probability distribution of colour at pixel coordinates informs the generator to create increasingly better fake data until the discriminator cannot distinguish fake data from 'real' input data. This back-and-forth dynamics between the two GAN parts build an abstract multidimensional space called 'latent space', where learned attributes of the input data exist as a continuous high-dimension vector format. This continuity of latent space is essential in widening the scope of the input dataset that is being trained. It allows for intermediate correlation within discrete training dataset to be calculated and any missing link between dataset entries to be filled. GAN's applicability is not limited to 2D images. GAN can analyse the probability distribution of voxels, instead of pixels, for 3D models. Recognising this, we attempted to utilize 3D GAN (Wu et al., 2016) to analyse various function-forms and obtain diverse design alternatives to internationally homogeneous function-forms.

In 1) Introduction, it was concluded that analysis of forms that denote a specific function in different local contexts could provide clues to generating variations of an internationally homogeneous function-form. With 3D GAN, targeted local function-forms and homogeneous function-forms can be put into a single dataset and analysed. Following this, the continual latent space within the 3D GAN model can project missing links between local and internationally homogeneous function-forms, amplifying the existing entry data and illicit forms that the homogeneous function-form morph into.

3. Experiment

Eco's principle of codification from A Theory of Semiotics and the insight into GAN allowed us to recognize 3D GAN's viability as a design tool. Based on entry data, the latent space in a 3D GAN model will be able to generate a diversity of forms that are “most probable.” In other words, 3D GAN can act as a generative tool that diversifies the entry data using a statistical approach. Though a degree of subjectivity cannot be entirely avoided, unlike Critical Regionalism, generated design output from 3D GAN is derived using data-based probability distribution.

Original definition	Redefinition by students
[Britannica] canopy, in architecture, a projecting hood or cover suspended over an altar, statue, or niche. [Designing Buildings - The construction Wiki] A canopy is an overhead roof structure that has open sides. Canopies are typically intended to provide shelter from the rainor sun, but may also be used for decorative purposes, or to give emphasis to a route or part of a building. During the Renaissance to become a baldachin, a fixed structure supported on pillars that was common in baroque architecture. Today,canopies may be independent of other structures, or may projects out from a structure such as a building, typically providing shelter at an entrance. It may be supported by the building it is attached to, and / or ground mountings cables, stanchions or upright support posts. [Law Insider] Canopy means a rigid multi-sided structure covered with fabric, metal, or other material and supported by a building at one or more points or extremities and by columns or posts embedded in the ground at other points or extremities. It may be illuminated by means of internal or external sources. [Planning law definition of canopy] According to planning laws today a canopy is defined as a light roof-like structure, supported by the exterior wall of a building and on columns or wholly on columns, consisting of a fixed or movable frame covered with approved cloth, plastic or metal, extending over entrance doorways only, with the purpose of providing protection from sun and rain and embellishment of the façade.	Team [Circulation]: Canopy is an assembly of overlapping architectural elements that create a sheltered space. Team [Temporary Encounters]: something that creates a shaded exterior shelter below for a temporary stay, unbound by more than one surrounding surface wall Team [Moving Parts]: An overhead projection with the primary function to cover and provide shelter, rather than to enclose
Remarks: Function / Symbolic Function / Element or material / Design or shape / Composition	

Table 1. Redefinition example of canopy (Encyclopedia Britannica (*Canopy | Architecture*, n.d.) / Designing Buildings – The construction Wiki (*Canopy*, n.d.) / Law Insider (*Canopy Sample Clauses*, n.d.))

With the drawn conclusion, we decided to carry out an experiment during a workshop with Social Algorithms Research Group. The two main objectives of the experiment were: a) to establish a design method that can diversify internationally homogenising function-forms; and b) to test the viability of 3D GAN as a design tool that aligns with the existing design theory and comprehension. In the preparation phase, the tutors of the workshop identified four architectural elements that embody a particular function (functional element), which were: opening, floor, support, and canopy. The idea was to identify elements that incorporate a specific function, which, with data processing through GAN, will inform various function-forms. As it is essential to create a dataset that disturbs singular and homogeneous ways functions manifest into forms, we asked students to define each functional element to be used as new parameters to define a dataset to incorporate both local (Korean) and generic (internationally homogeneous) examples (Table 1). The overall workflow of the experiment was as follows: 1) students redefined assigned functional elements; 2) students collected references of their definition and constructed 3D models for input datasets; 3) the voxelised model datasets were used to train 3D GAN to elicit patterns within each dataset 4) selected generated 3D GAN samples 5) analysed the selected 3D models for new prototypes for architectural design applications.

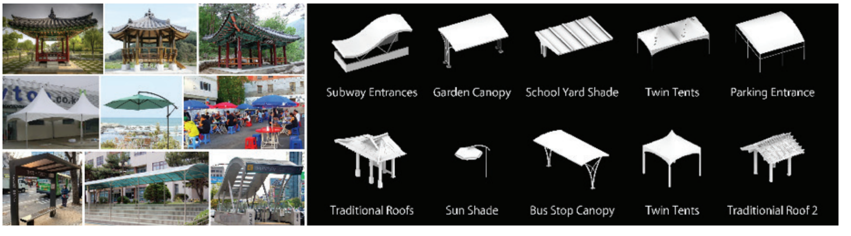


Figure 1. Reference images and 3D model examples of canopy

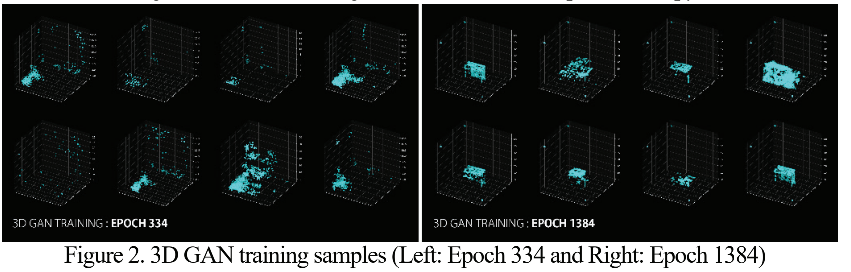


Figure 2. 3D GAN training samples (Left: Epoch 334 and Right: Epoch 1384)

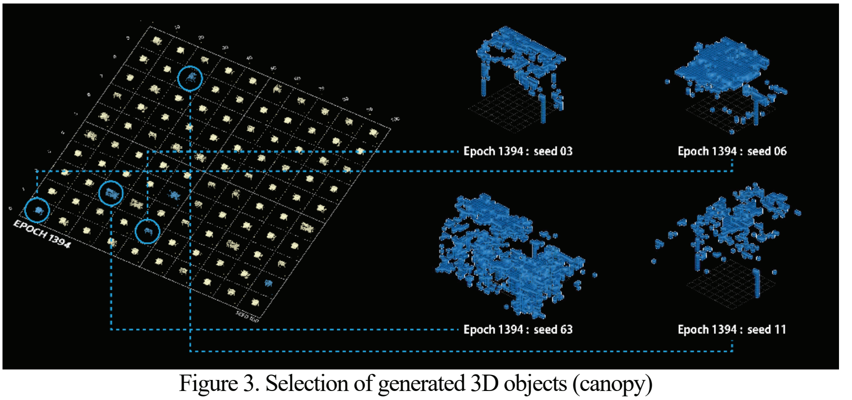


Figure 3. Selection of generated 3D objects (canopy)

During the experiment, team 'Temporary Encounters' proposed that the definition of "canopy" to encompass a wide range of structures that includes traditional Hanok roofs with supporting pillars, as well as other modern canopies that can be considered more generic, such as apartment building entrances, sunshade umbrellas, bus stops, subway entrances, etc. In accordance with this redefinition that now addresses local context, the students collected references from traditional and modern Korean architecture to refine curated 3D model data (Figure 1). The processed and categorised 3D models were then put into 3D GAN for morphological generation.

The amount of training data was kept at approximately 100 3D models per an element, and the training details are as follows: GPU v100x1; Average training time 3 hours; Voxel size 64x64x64; Epochs 2000; Batch size 16; Generator lr = 0.0025; Discrimiator lr = 0.00001; Dimension of latent vector 200.

In the training process, 3D GAN sequentially learns the probability of the presence of all input data forms in voxels, which refer to the probabilistic distribution of the small cubes that constitute the model space. The learned probability includes the overall form and the relationship between adjacent voxels (Figure 2 and Figure 3). After completing the training, students interpreted the generated 3D models and chose some of the output as design elements in new architectural compositions. For example, different generated forms from the same functional element dataset were aggregated, and then assembled with other elements to form a new spatial configuration that departs from the homogeneous architectural function-forms (Figure 5).

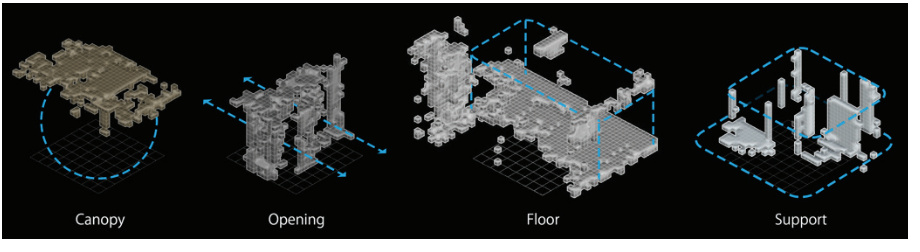


Figure 4. Outcome samples of canopy, opening, floor, and support from 3D GAN

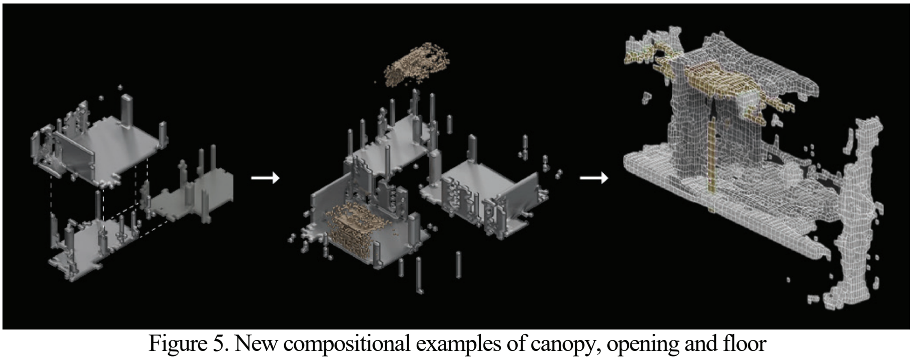


Figure 5. New compositional examples of canopy, opening and floor
(Floor + Floor / Floor + Canopy / Canopy + Opening + Floor)

4. Result and Analysis

The alternative redefinition of functional elements allowed for extraordinary Machine Learning outcomes with significant potential. The outcome was a low-resolution voxel shape with compositions of rough pieces due to the similarly low-resolution dataset and small sample size, which resulted in a wide degree of scattering, making it challenging to directly convert the result into a practical proposal. Nevertheless, each voxelised shape exhibits distinguishable spatial qualities in accordance with the new definitions by students. Generated shapes can be interpreted as canopies to 'temporary shelter', openings to 'connected passage', and floor to 'mediated territoriality' (Figure 4 and Table 2).

The original dataset included not only architectural elements with the function under the scope but also other parts that are directly connected to those elements. These connected elements were often identified and included through the parameters new definitions have set out. For example, canopies mostly include structural frames or supporting elements that hold their forms while openings are always created between two or more elements. These elements connected to functional elements often provided intermediary parts in the generated output that can be used as 'buffers' when combining different output into new architectural compositions. This design process signifies a necessity for 'intermediary buffers' that can be overlapped between different targeted design elements under the scope of the experiment. In transitioning to a design phase from GAN training, we also noticed that students exerted subjective biases while selecting output samples to be taken into the design process. Most students intentionally picked samples that fall within the redefinition set out in the initial phase of the experiment, while avoiding seeds that can be perceived as internationally homogeneous function-forms (Table 2).

Redefined element	Redefined function	Application finding
Canopy: something that creates a shaded exterior shelter below for a temporary stay, unbounded by more than one surrounding surface wall Floor: empty and flat surfaces that mediate indoor and outdoor spaces Opening: an open space between elements and elements, a passage that light and scenery can go through	Temporary: creating shelter with a surface above Mediate: generate new territoriality where people can stay Passage: a passage that light and scenery can go through	Pavilion: - a space that forms a temporary relationship between the user and the space, - a space that challenges the surrounding physical boundaries

Table 2. Redefined elements and functions for a new application

Although additional refinement is required to apply the findings from this experiment to practical architectural design, it elucidates diverse possibilities for a new way of processing architectural data and design generation. Though a degree of subjectivity cannot be avoided entirely with 3D GAN, especially when building dataset parameters and selecting design output, it provides a useful tool that is based on data and probability. This approach provides an alternative to a design practice, as observed through the examples of Critical Regionalism which was largely based on subjective interpretation of 'aesthetics'.

5. Conclusion

In the paper, we propose a new design method to diversify globally homogenising function-forms using Machine Learning process. During the experiment, we redefined architectural elements to allow for a dataset that fits our agenda to include data entries that can deviate from the understanding of homogeneous architectural elements; further curated and redefined the dataset; and trained and generated results which exhibit both generic (internationally homogeneous) and diversified (locality-driven) architectural features. This process led to positive findings that fulfil both of our initial objectives for the experiment, which were: a) to establish a design method that can diversify internationally homogenizing function-forms; and b) to test the viability of 3D GAN as a design tool that aligns with the existing design theory and comprehension.

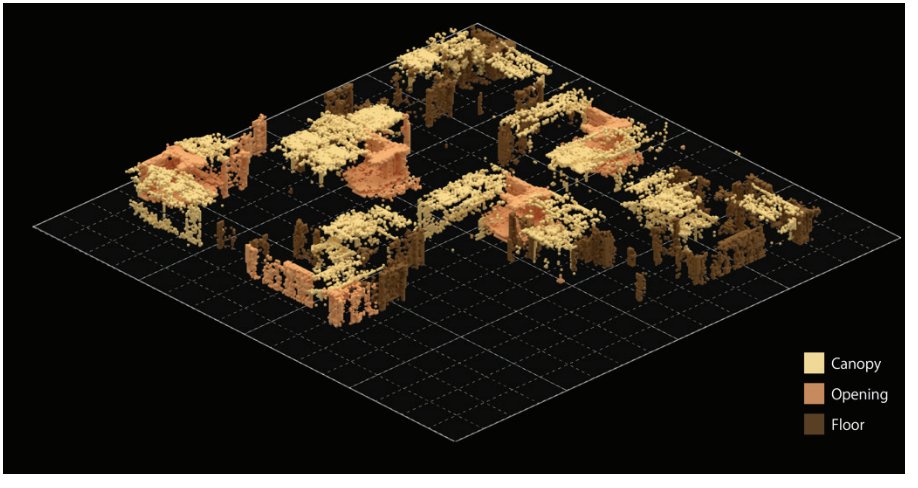


Figure 6. New composition system

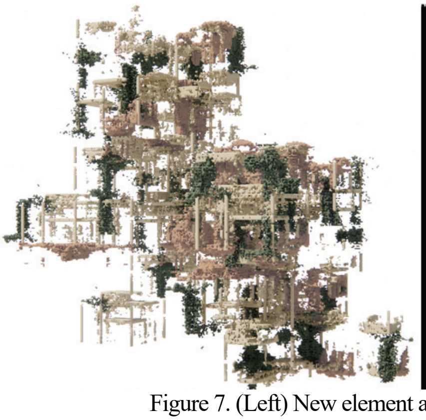


Figure 7. (Left) New element application by team 'Resonance' and (Right) Latent walk of new composition systems

Figures display newly aggregated compositions where generated elements from 3D GAN interact (Figure 6 and Figure 7). The constituent elements of these compositions, while they deviate from the globally homogeneous norms around architectural elements, hint new spatial understanding and aesthetics. The compositions signify a new possibility of architectural diversity while reflecting more localised contexts. 3D GAN also shows how new technologies can be used to understand existing architectural typologies and frameworks through Machine Learning while proposing to utilise data as fruitful design resources to unveil patterns and system. This experiment reveals the possibility of its application beyond element-scale. For example, spatial properties of different urban areas can be analysed in conjunction with a data parameter like climate or cultural zone. Furthermore, examining function-form relationship through Machine Learning in urban scale can be used as a tool to infer the correlation between the form and typology in a more objective manner. This new data-driven design approach expands the role of a designer to include curatorial decision-making. The experiment also elucidated a new comprehension where Machine Learning processes can be understood in conjunction with existing theory of design codification, especially one by Umberto Eco. This is significant as it provides ways emerging technologies can be considered and experimented within the existing architectural discourse.

However, several limitations were observed throughout the experiment. Firstly, it is not easy to clearly define the relationship between data parameters with morphological data entries. For instance, the definition of each functional element was not clear enough that students often had to use their subjective judgement when creating the input dataset for 3D GAN. Secondly, limited computing resources does not allow for machine learning processes that include high-resolution voxel models. Lastly, there are very limited number of architecture-centred datasets, forcing designers to often manually generate 3D data.

Acknowledgements

This research is a result of a collaborative effort during SocialAlgorithms9.0 Workshop (AA Visiting School Seoul 2022), which is further developed by the authors after the workshop. The workshop and its findings would have been made impossible without the help of the director, Jooeun Sung and Soomeen Hahm, who showed unfaltering support with timely advice and organisational cooperation. We would also like to send our gratitude to the following participants whose works have inspired tutors and other guests' critics.

Elemental Sabotage: Diffusing functional morphologies, Unit 4
Tutors: Dongyun Kim, Lloyd Sukgyo Lee, and Hanjun Kim
Participating students:
Team Circulation (Hyejin Jeon, Kitae Kim, Jiaxim Sun),
Team Moving Parts (Eddy Choi, Rayoung Kim, Isabel Solin),
Team Resonance (Minji Kim, Yueyao Li, Hyelim Yu),
Team Temporary Encounters (Jieun Kim, Sun Q Kim, Sol Ah Yoo)

References

Canopy. (n.d.). Www.designingbuildings.co.uk. https://www.designingbuildings.co.uk/wiki/Canopy

Canopy. (n.d.). Encyclopedia Britannica. https://www.britannica.com/technology/canopy-architecture

Canopy Sample Clauses. (n.d.). Law Insider. Retrieved February 5, 2023, from https://www.lawinsider.com/clause/canopy

Duchek, S., Raetze, S., & Scheuch, I. (2019). The role of diversity in organizational resilience: a theoretical framework. Business Research, 13(2), 387–423. https://doi.org/10.1007/s40685-019-0084-8

Eco, U. (1986). 2. Function and Sign: Semiotics of Architecture. The City and the Sign, 55–86. https://doi.org/10.7312/gott93206-004

Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative Adversarial Networks. ArXiv.org. https://arxiv.org/abs/1406.2661

Leslie, P., & McCabe, J. T. (2013). Response Diversity and Resilience in Social-Ecological Systems. Current Anthropology, 54(2), 114–143. https://doi.org/10.1086/669563

Neil Leach (Edt). (2014). Rethinking Architecture| A reader in cultural theory. https://marywoodarchtheory.files.wordpress.com/2013/10/function-and-the-sign-the-semiotics-of-arch_u-eco.pdf

Oliver, T. H., Heard, M. S., Isaac, N. J. B., Roy, D. B., Procter, D., Eigenbrod, F., Freckleton, R., Hector, A., Orme, C. D. L., Petchey, O. L., Proença, V., Raffaelli, D., Suttle, K. B., Mace, G. M., Martín-López, B., Woodcock, B. A., & Bullock, J. M. (2015). Biodiversity and Resilience of Ecosystem Functions. Trends in Ecology & Evolution, 30(11), 673–684. https://doi.org/10.1016/j.tree.2015.08.009

Puleo, T. J. (2014). The Sicilian Baroque: Reconciling Postquake Tensions. Environment and Planning A: Economy and Space, 46(11), 2552–2568. https://doi.org/10.1068/a46295

Soliman, E. M. (2019). Blessed Fate: The Study on how Heritage Churches resist natural calamities. IOP Conference Series: Earth and Environmental Science, 294, 012063. https://doi.org/10.1088/1755-1315/294/1/012063

Wu, J., Zhang, C., Xue, T., Freeman, W., & Tenenbaum, J. (2016). Learning a Probabilistic Latent Space of Object Shapes via 3D GenerativeAdversarial Modeling.

AI AND TYPOLOGY

3D Point Cloud Generative Adversarial Network Based on Tree Structured Graph Convolutions (TreeGAN) and Architectural Design

XIANG YIN[1]
[1]*University of Michigan.*
[1]*yinxiang@umich.edu, 0000-0002-4624-8437*

Abstract. The paper discusses a novel design approach that applies artificial intelligence as an auxiliary tool throughout typology research and architectural design. The method attempts to utilize neural network as a research tool to detect and identify features of a typical architectural type within the specific society context and demonstrate its potential for regional design under the theme of human centric. Typology classification, computational vision, and human-machine collaboration are entwined throughout machine learning and architectural design. The paper aims to demonstrate the ability of 3D Point Cloud Generative Adversarial Network Based on Tree Structured Graph Convolutions (TreeGAN) to study the inherent principle and characteristic of an architectural type and its potential to provide possible design inspirations based on the typological formation principles concluded by Deep Learning. The article exhibits the key result generated by TreeGAN in a specific architecture type—churches, as the prototype of a design method and conducts a project in Manhattan.

Keywords. Architecture Typology, Artificial Intelligence, Machine Learning, TreeGAN, Human-machine Collaboration.

1. Introduction

1.1. ARCHITECTURE TYPOLOGY

The dictionary definition of typology is 'a classification according to general type'. The origin of architecture typology raised from the 18th century when architects conducted classified processes in the continuous and integrated architecture system. The Primitive Hut (Laugier, 1755) presented the basic elements and components of an architecture and discussed a prototype of a typical building. The idea of 'type' was firstly theorized in the history of architecture by Quatremére de Quincy. His analysis of architectural precedents went beyond the limit of classic architectural configuration and defined 'type' as a pre-existent origin, a distinct 'mother tongue' in architecture (Guney, 2007). No matter how the external architectural features change, the architecture’s fundamental frame and geometric nature is permanent. 'Type' is the permanent and

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 39-48.

elementary logic principle that represents the inherent characteristics of architecture before form (Rossi, 1966).

Anton Schweighofer assumed that the construction purpose of typology is discovering new things which already exist, rather than inventing new things (Kühn, 2000), which means the inherent characteristics and principles of the buildings could be extracted by analyzing previous architectural projects and could be applied to new inventions. Typology, if grasped as an act that has an ideological background, is an organizing tool to define characteristics between different architectural features within specific built environments, and the ideal typology model may have more characteristics than all the examples that are included and analyzed (Gregotti, 1985). Thus, typology is a conceptual summary of common architectural characteristics. The typology study has typical approaches such as abstraction, reduction and schematization, and the process conceptualizes abstract principles of a typical type of architecture by identifying, clarifying, and comparing examples (Ayyıldız, 2017). Subsequently, the principles and conclusions generated from the typological analysis could presuppose design methods, since typology is also a research methodology. Architects usually search for a main principle to apply for the spatial arrangements and configuration formation when they have an interpretation of the design assignment (de Jong and Engel, 2002). This is a recurring process that demonstrates principles for the design and indicates the use of a typological solution (Figure 1).

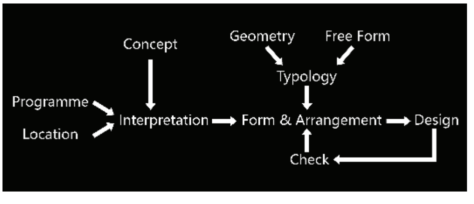


Figure 1. Phases and influences in design process

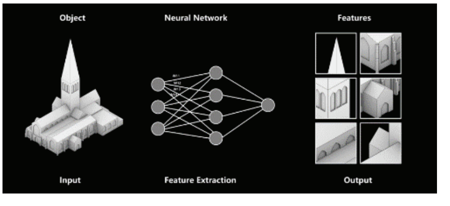


Figure 2. Deep Learning-feature extraction

1.2. MACHINE LEARNING

With the development of feature recognition ability of artificial intelligence (AI), the typology research and application process can be improved more efficiently and productively via Deep Learning. Deep Learning can learn a specific architectural type based on an extensive database through feature extraction (Figure 2). The exterior features of a specific architectural type are detectable and measurable when AI analyzes the input architectural models in the database. The neural networks could clearly recognize the input objects and extract valuable features (Dara, 2018). Then the neural network clarifies the extracted architectural features, and finally the characteristics and principles of the specific typology are summarized with the output presentation of point cloud models. Also, the output results could give architects novel inspirations during the recurring design process. Using AI as an auxiliary tool during the human-machine collaborated design process, the paper aims to demonstrate (1) the ability of AI to study and understand the inherent principle and characteristic of the specific configuration of an architectural type by analyzing input architectural models and (2) its potential to provide typic models of a certain architectural type based on the typological formation concluded by Deep Learning.

1.3. HUMAN CENTRIC

The development of an architectural type involves both needs and expectation of beauty; a particular type indicates the form and the way of life, based on the specific society context (Rossi, 1966). The culture and history form the city identity, however, in the era of globalization and industrialization, the cities expand quickly and many of the city identities become vague. Corresponding to the human centric theme, the typology research method excavates the valuable identities within the built environment and the culture of local society. The typological analysis in architecture makes it possible to keep the continuity of valuable architectural identities and transfer the elements of a city's history to its future architecture design (Ayyıldız, 2017). Also, the application of AI in typology research helps architects understand the society context and formation logic more efficiently, leading to the production of human-centric architecture.

2. Type Selection

To reveal the method and process, a certain type of architecture needs to be selected and studied in this paper. Based on the elements and characteristics extracted and analyzed from the samples, the architectural ‘type’ could be a form type, a structure type, or a function type (de Jong and Engel, 2002). The characteristics of a form type are only related to the form, such as round, square and triangle. The structure type includes the properties of separations and connections, for example, basilica and peripteros. The function type includes external formation within the functional concern, which is most suitable in typological research for the classification. Therefore, this paper would pick ‘Church’ as a functional architectural type to study.

Churches are one of the most typical types of architecture for worship services and other religious activities. The churches contain different kinds of styles and symbols in history. The church design is driven by cultural sympathy, conventional function, and distinctive architecture materials. Churches symbolize artistic achievements of various ages and represent the continuous feature variation in history. For instance, the Gothic Church is famous for its stained glass, pointed arches, ribbed vaults, and flying buttress; with the development of modern techniques, high-tech churches are more spatially flexible, such as asymmetric plans. Although the cruciform plan and numerous iconic elements of ancient churches have a certain historical contingency, the spatial and functional characteristics are what any church needs to have: the sublimeness of space, the mystery feeling of religion and the uniqueness of configuration (Li, 2011). Even though the features and styles of churches are continuously changing in history, the basic logic and principle of the space and function are consistent, such as the main worship space and essential artifacts to underlie spiritual themes.

The paper has seen the philosophy of AI and architectural typology share the resembling comparative and cognitive methodologies of objects and brings up the features of churches as a starting point into the AI typology research. The 'type' of churches could be identified and summarized by AI and serve to be a unique machinery representation after AI is trained with big data.

3. Method

The paper will rely on a specific site in Manhattan to show how a project is delivered by Deep Learning on typology. To start with, it is necessary to give the readers a background knowledge about existing churches in Manhattan. Although Gothic churches have long been admired in the area, the interior infrastructure is crumbling. Thus, the old churches' leadership started to strike a deal with the developers to demolish old buildings and use the land to construct some other types of architecture for profit, like residential buildings. The target site is the Norfolk Street Baptist Church, which was built in 1848. It still has all the excellent Gothic Revival touches of a mid-19th century church. In 2012, the congregation committee asked the landmarks commission to tear down the synagogue and sell the land to developers. But citizens are unsatisfied with the reconstruction projects because they may destroy the historic urban texture of New York. Thus, a new church needs to be designed to substitute the old structures in the same place. Meanwhile the new church should respond to the regional culture and extend the values to the future.

3.1. INITIAL DOCUMENTATION

The first step is to set up a model database for Deep Learning. To establish the church database and relate the project to local historic texture, the author investigates the existing churches in Manhattan and creates the massing models in Rhino to build an initial model database. 232 churches in Manhattan are documented but the total number of the church samples is inadequate to conduct a successful machine learning. So, more churches in the surrounding area of Manhattan are added to the database and finally there are 2000 churches in the database. The churches are modelled with simplified geometric features for better identification by AI (Figure 3).

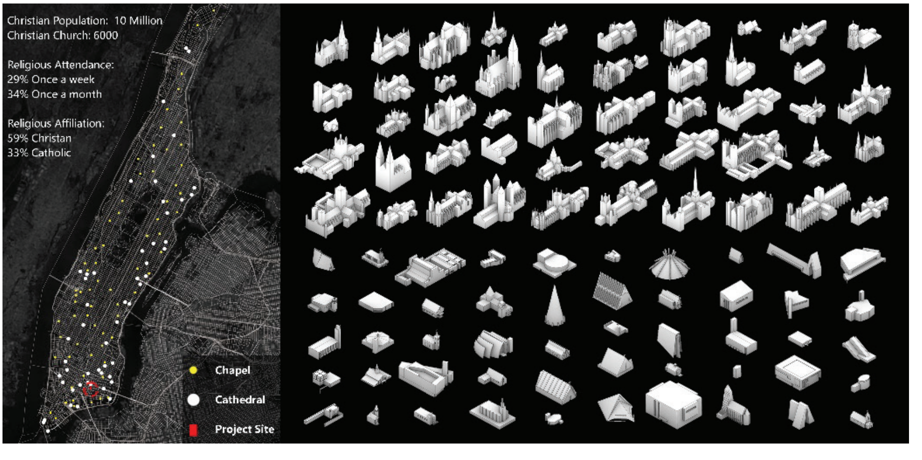


Figure 3. Left: Churches in Manhattan. Right: Part of the documented church samples.

In addition, 35 of the 232 investigated churches (15%) in Manhattan are contemporary churches, their spatial arrangement and feature characteristics are obviously different from classic churches because of the movement of modernism. For a better result, the ratio of contemporary churches is the same in the extended database.

3.2. TREEGAN

Recently, plenty of approaches are related to machine learning and architectural 3D model generation. However, it is still difficult to apply these techniques and methods throughout the architectural design because of the limitation of model resolution and spatial information loss. Related works like training 3D Graph Convolutional Neural Networks (GCNN) to generate models by inputting parameters after Deep Learning, but the results heavily rely on how the models are labelled, which makes it difficult to capture features objectively and comprehensively (del Campo et al., 2020). Other methods like 3D model generation by training a series of plan and section images could acquire a high model resolution and enough feature details, but the process remains complicated, and the results are unpredictable (Zhang and Blasetti, 2020).

Different from other neural networks, 3D Point Cloud Generative Adversarial Network Based on Tree Structured Graph Convolutions (TreeGAN) could generate a diverse range of point clouds in an unsupervised manner without any prior knowledge. TreeGAN can learn a latent space of features from the given input 3D models and generate semantic parts of objects more accurately and efficiently (Shu et al., 2019). The church models in the database are firstly exported as point cloud PTS files in CloudCompare and input into TreeGAN. To keep a reasonable training time, the point cloud files of the input churches cannot be too complex, with a maximum of 2,500 points each file. In TreeGAN, all the church input models are ancestors with different feature spaces and documented in different layers. All the different feature spaces containing the information from previous layers are combined effectively, for the current point to generate the next point in the best mapping way (Figure 4). So, the features of 2,000 churches are extracted, combined, and summarized by AI. In addition, compared with the conventional learning loop in other GANs, the graph convolution in TreeGAN has a fully connected layer of ancestor points in the mapping loop, so TreeGAN could deal with complicated features and spatial information of the churches, and makes the point distribution in results more accurate and exhaustive. Finally, the new church models are generated and exported in point clouds. Through the process of Deep Learning, the characteristics and formation principles of the church models are studied by TreeGan and the 'type' of church is detected by AI.

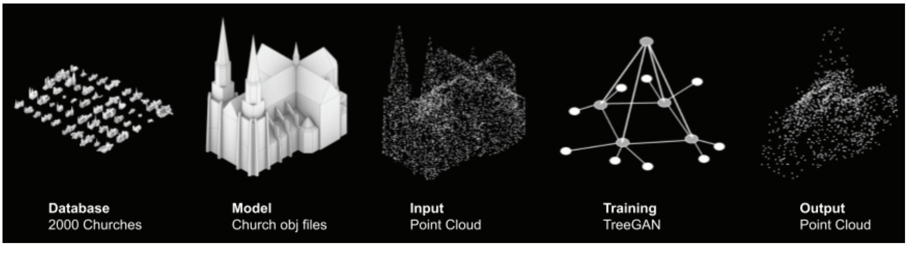


Figure 4. TreeGAN Deep Learning with church database.

TreeGAN could output various point cloud models with different training parameters, such as training epoch and seed. The number of the epochs is a hyperparameter that indicates the total number of iterations of all the data in one cycle for training the input model. Seed helps get predictable and repeatable results during multiple executions of the training. To have a variation and comparison of the results, the author set different epochs and different seeds to output. Each epoch with a selected seed has 6 output samples (PC 1~6). Some of the results show lots of typical features of churches. Among all the output samples, the model generated with epoch 100/ seed 100 shows the most ideal result, since it displays a typical and balanced point cloud distribution of church features: a high tower, cross-like plan shape, reasonable dimension, and diversified interior spaces (Figure 5).

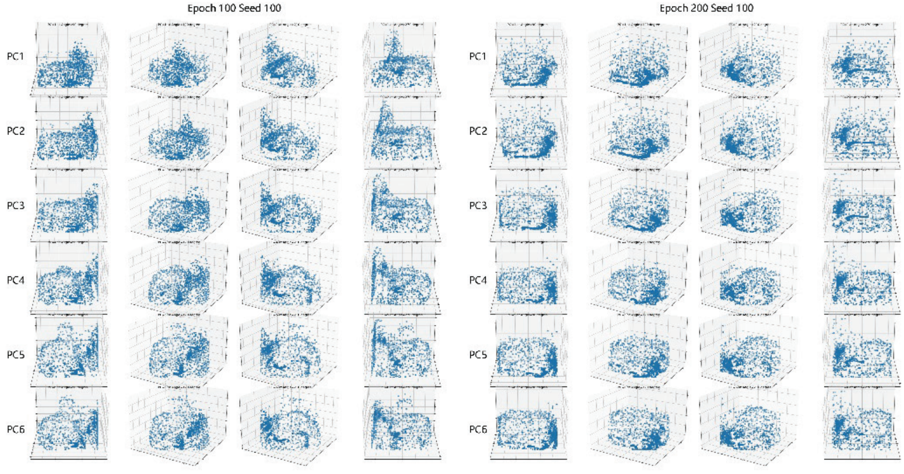
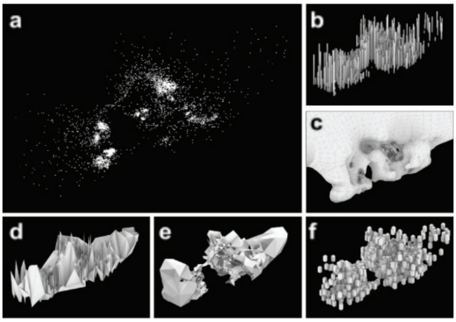


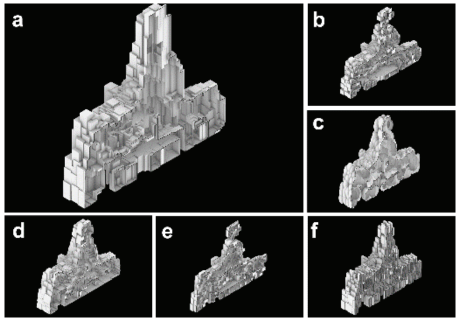
Figure 5. TreeGAN Training Results (Output Point Cloud) at 90/120/210/270 Degree Angle.

3.3. HUMAN-MACHINE COLLABORATION

3.3.1. Grasshopper Voxelization

Machine automatic generating is a bottom-up process; in order to convert the point cloud model into a spatial reproduction, it needs to collaborate with architects to make top-down decisions. Several transformation methods are tested to reconstruct mesh surfaces based on the generated point clouds, such as vertical pipes, voxelization, Topo Delaunay meshes, Poisson mesh reconstruction, and mesh-lab surface reconstruction (Figure 6). Among all methods, Voxel conversion displays a more potential architectural spatial recompositing of occupancy. Based on the voxelization approach, different options of voxels are tested in Grasshopper, such as rotating cubes, slabs, L shape blocks, and ellipsoids, with different voxel dimensions according to several epochs and seeds from point clouds results (Figure 7). The process operates a comparative spatial quality test of these volumes, and an appropriate space prototype is selected for further design. Meanwhile, the influence of voxel size on interior space is investigated. The final voxelization model consists of 10 feet both width and length voxels but their height varies from 20 feet to 50 feet depending on the point position in the model (from basement to tower). By gradually adjusting the dimension of voxels, the voxelized prototype is reaching a balanced condition and presents obvious architectural features of a church.





Left: Figure 6. Mesh Generation with Point Cloud. (a) Output Point Cloud (Epoch 100/Seed 100). (b) Vertical Pipes. (c) Poisson Mesh Reconstruction. (d) Topo Delaunay Meshes. (e) Mesh-lab Surface Reconstruction. (f) voxelization.
Right: Figure 7. Voxel Exploration. (a) Voxel: Height Variable(20~50ft) Cuboid. (b) Voxel: Cube (10x10x10ft). (c) Voxel: Ellipsoid. (d) Voxel: Trapezoid. (e) Voxel: L-Shaped. (f) Voxel: Slab.

3.3.2. Interior Reformation

The voxelized model has the embryonic form of interior space, but the interior components are too disordered for an ideal occupancy. To complete the final church model, the last step is to manually modify and adjust interior spaces based on the prototype generated by AI. The cluttered part of voxels is simplified and some supporting components, such as slabs, walls, and stairs, are added into the model. Meanwhile the initial spatial sequence and arrangement as well as some interior features are kept. Thus, the functional spaces are improved, and the church's rituality is emphasized (Figure 8).

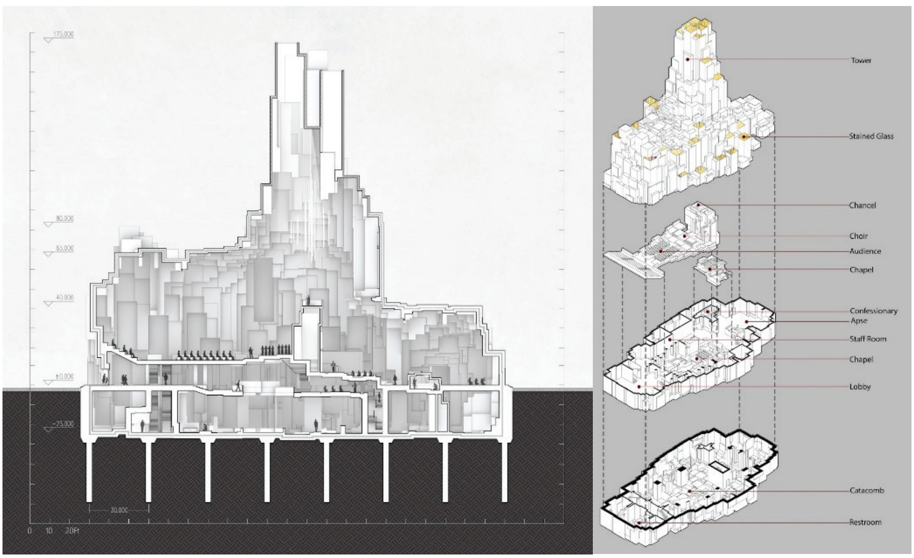


Figure 8. Left: Church section. Right: Spatial arrangement.

4. Result

4.1. ARCHITECTURAL DRAWINGS

Figure 9 shows the plans of the designed church. The final church has five parts: tower, main worship space, auxiliary chapels, catacomb, and other affiliated infrastructures. The second floor is the main worship space, includes a main nave, an auditorium, a choir, a center chancel, and an auxiliary chapel, which manifest a sense of rituality and verticality for the ceremonial opening. The center chapel is the visual concentration and the sunlight coming down from the tower will illuminate the choir and the stained glass, creating the atmosphere of worshipfulness. On the first floor, there is another auxiliary chapel and some confessionary rooms, which provide a private worship experience. The basement contains a catacomb for city usage. The church is then completely designed based on Deep Learning and its feature representation.

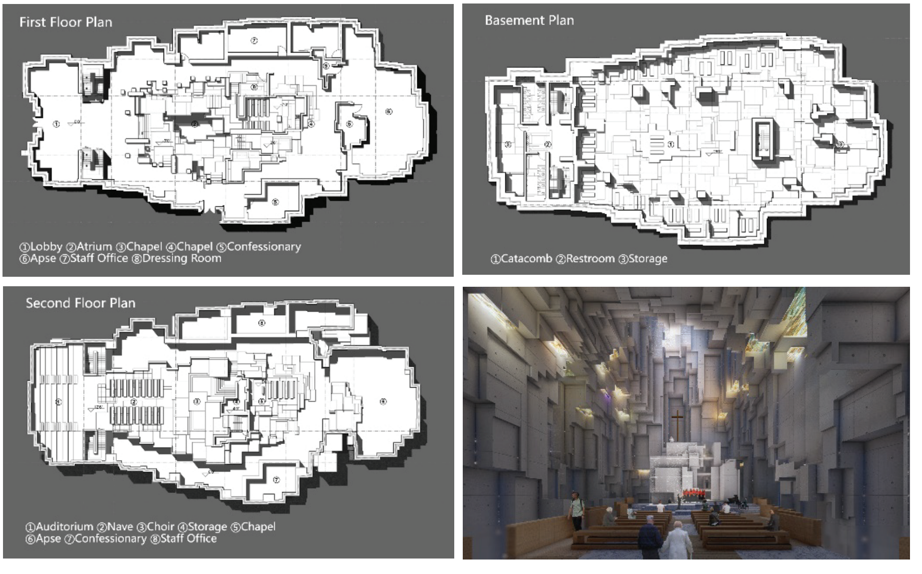


Figure 9. Plans and interior rendering.

4.2. DELIRIOUS CHURCH

The blooming of technologies and the diversity of human behaviours created a special urban context of Manhattan: the culture of congestion (Koolhaas, 1978). The completely artificial urban discourse continuously fulfils the city with new contents. In this urban environment, every building is so unique and isolated within the blocks by the grid-like streets, but in the meantime the artifacts are related and strengthening each other. Under the city background, the AI church carries on the obvious identities of local religious architecture, meanwhile the application of novel techniques results in an avant design that breaks through the traditional spatial arrangement and configuration, activating the diversified society context of Manhattan in the new digital era.

5. Discussion

The generated church model is a prototype to apply artificial intelligence throughout typological research and design. The paper thrives the traditional research method with machine logics and computational version, leads to a human-machine collaborated design process. The achievement of this paper is (1) proving AI could detect and identify the architectural configurations and extract the characteristics of a specific architectural type during the research period, (2) demonstrating AI could apply the studied information to provide models as typological solutions during the design period, and (3) merging local built environment and cultural identities into design process with the help of AI in typological research and design.

The use of TreeGAN as an analysis approach and an interdisciplinary design method in architecture is a new topic to explore. Nowadays, scarce examples dabble in the field of TreeGAN in architecture. But TreeGAN is gradually used in image generation with hierarchy (Zhang et al., 2021) and syntax-aware sequence generation (Liu et al., 2018). It is exciting that a connection is established between the existing architectural theory and AI techniques, and re-exhibit architectural history and conventional typology theory into contemporary architectural design with the machine learning assistance of TreeGAN. Meanwhile, the future work is still needed according to the presented method. Due to the computer calculation ability and the training time, the input point cloud models could not contain enough points to provide a most detailed machine learning, which brings problems to the mesh generation of the output point cloud model with inadequate points. In addition, the interior spaces and exterior features are both formed by inputting the simplified exterior point clouds into TreeGAN, so the interior spaces are heavily influenced by exterior features. This causes the functional arrangement is partially disordered during automatic generation and needs human's effort to reform interior space. Lastly, the whole machine learning process is based on exterior architectural features, the basic formation logic and principle of a typology cannot be fully studied by AI through this bottom-up process, although TreeGAN could detect and extract characteristics of architectural features. Further research could explore more detailed and automatic 3D Deep Learning methods that includes both exterior features and interior spatial information, to reduce human interaction and achieve more objective results.

6. Conclusion

The paper starts with Aldo Rossi's definition of 'type': 'I would define the concept of type as something that is permanent and complex, a logical principle that is prior to form and that constitutes it' (Rossi, 1965). The word 'type', tracing back from Ancient Greek, is the meaning of 'model', the first essence of objects in a semiotic way. Similarly, Architectural typology deploys a conventional epistemological philosophy that reveals 'type' as the root of items, which has been a long term hidden in architectural history when people pursue a technology-advanced concept and society. The philosophy of AI and architectural typology share the resembling comparative and cognitive methodologies of origins of objects, and this paper brings up the history of architectural typology and features of churches as a starting point into the AI and typology research. This practicing method reveals the architectural characteristics under the regional culture context and extracts the unique attributes out of the hybrid architectural condition. With the development of GANs and computation, the feature recognition and reproduction techniques of Deep Learning will thrive the traditional architectural theories and provide more potential design possibilities in the future.

References

Ayyıldız, S., Ertürk, F., Durak, S. & Dülger, A. (2017) Importance of typological analysis in architecture for cultural continuity: an example from Kocaeli (Turkey). In *2017 IOP Conference Series: Materials Science and Engineering* (pp.245-260). IOP publishing.

De Jong, T. & Engel, H. (2002). Typological research. In T. M. de Jong, D. J. M. van der Voordt (Eds.), *Ways to study and research* (pp.103-107). DUP Science.

Del Campo, M., Manninger, S. & Carlson, A. (2020). 3D graph convolutional neural networks in architecture design. In *35th International Conference on The Association for Computer Aided Design in Architecture: Distributed Proximities, ACADIA 2020* (pp.688-696). The Association for Computer Aided Design in Architecture (ACADIA).

Dara, S. & Tumma, P. (2018). Feature extraction by using deep learning: A survey. In *2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA)* (pp.1795-1801). Institute of Electrical and Electronic Engineers.

Gregotti, V. (1985). The grounds of typology. *Casabella, 509-510 (1985), p. 4-8.*

Guney, Y.I. (2007). Type and typology in architectural discourse. *Journal of Balıkesir University Institute of Science and Technology,* (9).

Koolhass, R. (1978). How perfect perfection can be, In *Delirious New York: a retroactive manifesto for Manhattan* (pp.161-230). Oxford University Press.

Kühn, C. (Ed.). (2000). Anton Schweighofer - the silent radical: buildings, projects, concepts. Vienna: Springer, ISBN 3-211-83343-9.

Laugier, M. A. (1755). An essay on architecture. London: T. Osbourne and Shipton.

Li, S. (2011). Study on the mode church in Germany. Central Academy of Fine Arts, China.

Liu, X., Kong, X., Liu, L., & Chiang, K. (2018). TreeGAN: syntax-aware sequence generation with generative adversarial networks. In *2018 IEEE International Conference on Data Mining (ICDM)* (pp.1140-1145).

RossI, A. (1965). The structure of urban artifacts, In *The architecture of the city* (pp.35-46). Cambridge: MIT Press.

Shu, D.W., Park, S.W., & Kwon, J. (2019). 3D point cloud generative adversarial network based on tree structured graph convolutions. In *2019 IEEE/CVF International Conference on Computer Vision (ICCV)* (pp.3858-3867).

Zhang, H. & Blasetti, E. (2020). 3D architectural form style transfer through machine learning. In *25th International Conference on Computer-Aided Architectural Design Research in Asia: RE: Anthropocene, CAADRIA 2020* (pp. 659-668). The Association for Computer-Aided Architectural Design Research in Asia (CAADRIA).

Zhang, R., Mou, L., & Xie, P. (2020). TreeGAN: incorporating class hierarchy into image generation. ArXiv. https://doi.org/10.48550/arXiv.2009.07734

COMMUNICATION WITH DETROIT

Machine Learning in Open Source Community Housing Design

MINGRUI JIANG[1] and CHENYI CAI[2*]
[1]*GSAPP, Columbia University, New York, USA.*
[2]*Department of Architecture, ETH Zurich, Switzerland.*
[2]*School of Architecture, Southeast University, Nanjing, China.*
[1]mj3092@columbia.edu, https://orcid.org/0000-0003-2603-7584
[2]cai@arch.ethz.ch, https://orcid.org/0000-0002-6249-1816
[*]*corresponding author*

Abstract. Traditional pre-design investigation includes conceptual studies, site analysis, and programming processes to analyze the site and design for users. Instead, designers and architects should consider users' ideas and their actual usage of space, which are recorded and reflected on the social media platform. To introduce more citizens' voices in the design and learn more about people's expression of Detroit city and its housing, we propose to involve the machine learning analysis in the earlier stage of the housing project using users' reflections from social media to support the conceptual design. This paper introduces a novel design framework that deals with the lacking public programs in Detroit using an online data clustering platform and demonstrates a conceptual open-source community housing design according to related findings. This framework incorporates data collection from the Twitter platform, implementation of clustering for user-oriented programs, and design applications based on the findings. Our research demonstrates an efficient and flexible approach to the open-source community housing project.

Keywords. Machine Learning, Decision Making, Social Media, User-oriented Design, Open Source Community

1. Introduction

With the advancement of new technologies, architectural design methods have constantly been evolving to bring a better environment for all. The tools to design the functions and construct the forms of buildings have been rapidly developed. Researchers have applied generative design approaches (Soddu, 2020), annealed neural networks (Yeh, 2006), fractal algorithms (Wen et al., 2010), self-connectivity models (Rian and Asayama, 2016), and so on for the computer-aided architectural design. However, the existing computational modeling methods mainly focus on the indicator or rule-based design generation, which sometimes causes a gap between people's preferences and housing functions, making architecture less user-friendly. Indeed, volumetric up-to-date straightforward, and incisive data published on the social media platform can strongly support the community housing project by encouraging public dialogue with the end users. Observing the lack of social media involvement in design decision-making and the shortage of computational design in early-stage design strategies, we introduced social media and machine learning to explore the potential of human-centric community housing design.

Compared with traditional design approaches, machine learning methods provide more opportunities for architecture and urban design. ML-based computational design has the advantage of integrating massive user data analysis and offers a higher capacity of data, more efficiency, and more holistic reliable results for design decision-making. Furthermore, clustering analysis helps analyze and visualize general patterns of the data. It instructs urban designers to query for and study the user data clusters closely linked with the design task to learn the knowledge from the existing environment and to provoke questions and ideas from the investigation.

Our research introduces machine learning for community housing design using Detroit as a case study. Image data collection from Twitter and the clustering algorithm technique in the machine learning process support our preliminary user research. We then visualized the clustering outcome and navigated into the pattern of Detroit city impression on social media, which further directed our design about open-source community housing. Noticing the lack of public activities in Detroit from the clustering analysis, we developed strategies for public events in community housing. We set up an open-source platform to allow users to communicate, interact, and share the community space based on the human-centered design (HCD) approach.

2. Related Works

In the last decades, computation-based approaches in design, including parameter-based generative design and machine learning, appeared and soon became popular tools among designers and architects due to improved efficiency. Computational Design (CD) challenges and renovates earlier architectural design conventions and praxis (Rocker, 2006). It often combines with knowledge from other expertise fields to enhance traditional workflow and explore new research methods, generating new terms, including parametric design, generative design, and algorithmic design.

In the practical field of architecture and design, artificial intelligence (AI) is widely applied. Foster and Partners built the Applied Research and Development (ARD) team, exploring the great potential of AI to shorten the analysis time and save costs. Moreover, AI technology empowers community design and profoundly influences society. It invites the local community to manage data, provide the agency to promote AI research and deal with local social needs. Compared with traditional community design, AI-powered systems support the analysis of more significant amounts of data in a shorter time. They can also timely interact with citizens as community infrastructure, adapted to long-term social changes. The co-creation and collaboration between AI and community members reduce repetitive work and result in faster decisions in community design. Hsu et al. (2022) applied AI in Community Citizen Science (CCS) framework to establish an AI system to assist citizens in collecting air pollution evidence and using the data to take action and gain knowledge.

Although computational technology has improved human use, comfort, and recreation in buildings, technology often takes a leading role in the decision-making process and excludes the human aspects. Therefore, Norman (2013, p.8) put forward the HCD approach, which first considers human needs, capabilities, and behaviors, then designs based on these considerations. Machine learning no longer objectively evaluates the massive data but portrays the assessment from personal perspectives. Regarding social media, Twitter's high-volume and real-time textual and imagery data have already been extensively studied for event detection and sentiment analysis (Hou et al., 2021). In urban study and architecture, Alvarez-Marin and Ochoa (2020) used thousands of geotagged satellite and perspective images from various urban cultures to identify and predict personal preferences in urban spaces via machine learning. Besides, Cai et al. (2021) proposed a hybrid data-driven approach to retrieving multi-dimensional cases via the clustering algorithm, reflecting the human-centric potential of machine learning in recognizing user preferences and needs.

The gaps we see from the literature review are as follows: 1) The preliminary user investigation of community design still needs a more explorable and comprehensive data analytic tool for design decision-making. 2) Practices and big data/artificial intelligence applications on social media data focus more on the analysis outcome, which remains in the stage of data visualization while lacking the bi-directional feedback from the program and the designers in the navigation process. 3) In the early design phase, the formulation of design strategies has relatively little connection with the user needs, which might be found via social media, and design concepts, which could be inspired by information abundance.

Our study proposes a practical framework of a platform that automatically crawls Twitter data via the program developed with webpage sparse technique and machine learning, and offers the spectrum of images of Detroit. It displays the overall pattern and zoom-in clusters for users to meticulously view the content regarding people, food, outdoor landscape and so on in each cluster, which not only helps designers to navigate according to specific design tasks, goals, and information they are interested but also establishes a platform for dialogue between designers and users. Taking community design as an instance and driven by our interest in Detroit's public activities as designers, we investigated the public activity cluster in the program navigation and found the lack of a series of public activities such as food sharing, childcare, community gardening, workshop activities, watching movies, and fitness. This finding later became the original motivation of our schematic design, the base of design decision-making, and a solid basis for introducing interactive platform function in the scheme.

The following sections propose several methods for an open-source community design with an integrative workflow, focusing on public activity strategies. It composites methods for data collection from social media, data analysis and visualization by machine learning, data navigation for social activity pattern discovery, design decision-making, and a framework that ties the components together.

3. Materials and Methods

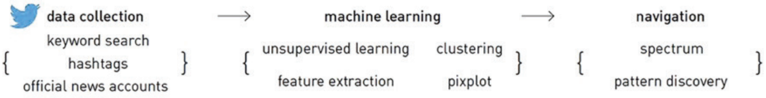


Figure 1. Workflow Process of communication with Detroit

To make design decisions based on Detroit's unique social pattern, we carried out communication with people in Detroit by figuring out urban clusters via Artificial Intelligence. As shown in Figure 1, the workflow process is divided into three steps: data collection, machine learning, and navigation.

Firstly, we collected data from keyword searches, hashtags, and official news accounts on the popular social media platform - Twitter. Secondly, through the machining learning step, we applied PixPlot to conduct the unsupervised learning to extract the feature and cluster the data since cluster analysis helps us focus on the data we are concerned about that is more related to the design task. In the last navigation step, a spectrum with ten colors allowed us to navigate among the clusters generated from machine learning and zoom into the specific cluster to discover the pattern of Detroit and learn more about local users, which in our case, is the public activity for the community design. Based on the findings, we provided needed space and an online open-source platform for Detroit residents to communicate and collaborate.

3.1. DATA COLLECTION

Considering Twitter's worldwide popularity and large number of active users, we selected Twitter as the social media platform for further research. Compared with text, images contain more information and are suitable for visualization. Furthermore, images with the relevant textual content of one tweet can convey high-quality and valuable information. Imagery provides a more accurate visual explanation for the texts in the tweets.

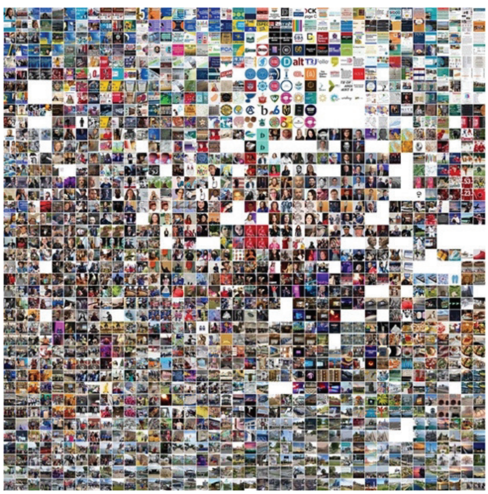


Figure 2. Collected Detroit data from Twitter

Integrated text and image for machine learning is a fascinating idea, but due to the time, technique limitation, and advantages of imagery over text in Twitter mentioned above, we take imagery as a starting point in this experiment. Therefore, we used Mathematica to exact the Twitter image data from the official Detroit news accounts, keyword searches, and hashtags such as "Detroit", "Detroit Community" and so on, as reflected in Figure 2. We then gathered over three thousand images from different Twitter users about Detroit and its community together for further cluster analysis.

3.2. MACHINE LEARNING-PROGRAM REFINING

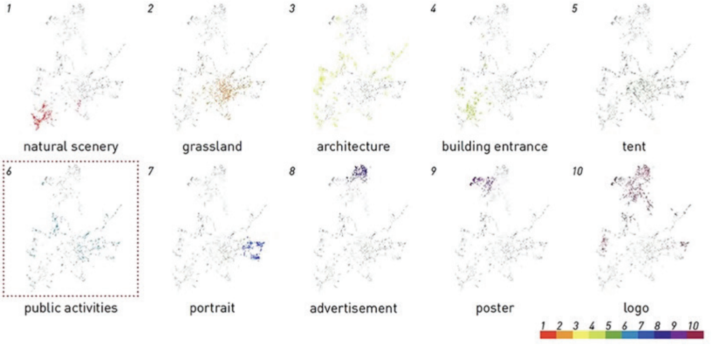


Figure 3. Clusters of data generated by PixPlot

Through unsupervised machine learning: PixPlot, simplifies the numerous and complicated Detroit image data and presents the distribution of different characteristics in a two- dimensional projection using a custom WebGL viewer. We input features from a pre-trained model: GoogleLeNet Inception v3, which is configured in the PixPlot package. As shown in Figure 3, the computing result from PixPlot contains ten clusters of Twitter images, which are visualized as a spectrum with ten colours. Clusters reflect the users' impressions of Detroit and help the further analysis of the community. The more concentrated the clusters are, the more notable characteristics they share.

Among all the ten clusters categorized by PixPlot, we selected the cluster related to Detroit urban communities based on our interest in public activities. The "public activities" group takes up a small proportion of Detroit's media data and demonstrates an unobvious distribution pattern. The "logo" cluster takes up the most portion. It has the most concentrated distribution, which depicts Twitter users' collective impression of Detroit with many logos. The patterns of the "natural scenery", "grassland", "tent", "portrait", "advertisement", "poster", and "logo" are clear to figure out, while the data of "architecture", "building entrance", and "public activities" are sparsely scattered. Its unclear distribution characteristic indicates that Detroit requires more public activities in the communities to gather people together.

3.3. DESIGN TASK

Known as the typical example of the Rust Belt city, the communities in Detroit are looking for ways to work together and revitalize vacant lots after experiencing a drastic economic decline because of deindustrialization, which caused fewer high-paying jobs and high poverty rates. The decay of land and abandoned buildings make Detroit gradually become a ghost town and lose its vitality. The developers' visions drive some new constructions without realizing people's needs, which later will become vacant buildings, and the situation reaches a deadlock. Therefore, the reuse of vacant lots becomes one of the significant tasks in Detroit.

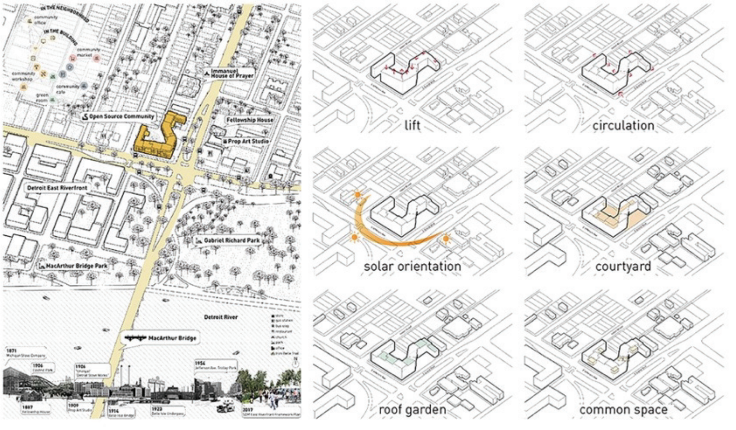


Figure 4. Site Analysis and Massing Generation

The site located at the intersection of East Jefferson Avenue and East Grand Boulevard, shown in Figure 4, is one of those vacant lots. There exist only a few public activities despite being surrounded by several historical buildings. Thus, we decided to dive into the "public activity" cluster, intending to find out what people like to do, what they do more, and what they do less in Detroit and vitalize civic life on the site. The massing fits the zoning regulation, surrounding building height, and solar orientation. Its serpentine shape naturally forms interior and exterior courtyards for public activities. Additionally, common space provides gathering places for community users and creates a collaborative shared lifestyle.

3.4. NAVIGATION ACCORDING TO THE DESIGN TASK

Zooming into the "public activities" cluster generated by PixPlot shown in Figure 5, plenty of public activities already exist, especially sporty ones including cycling, jogging, football, rugby, skating, and outdoor yoga in Detroit. People also enjoy significant outdoor gathering events such as music festivals, performances, and parades. They are eager to participate in collective activities and establish relationships with each other.

Although the "public activities" cluster data vividly showcased the vibrant civic life in Detroit through images uploaded by Twitter users, we found the lack of a series of communal community activities behind it, such as cooking class, spa, food sharing, mixology class, community gardening, bowling, movie nights, dancing, hiking, childcare, handicraft making, workshop activities, yoga, fitness, and art studio. Detroit citizens are more active in engaging with the city than the community. With our findings of fifteen lacking public activities from the clustering result, our design group went to Detroit to conduct the field research. We were impressed by the emergence of community gardens and farms such as Little Detroit Community, Hubbard Farms Community Garden, and Cadillac Urban Gardens on Merritt. We also found that the city of Detroit is pushing forward the Neighborhood Beautification Program, which provides funds for community gardens, public space improvement, and clean-up activities. Combining the clustering result, the need for community gardens, and our group's observation of existing public activities in Detroit, we selected ten proposed activities for our community housing design to activate the community's civic life. Rather than design from the designers' perspective, we envision a new human-centered potential for the community housing design from the end users' views by introducing the activities they want to do on the social media platform.

Figure 5. Existing public activities and Potential public activities in Detroit

4. Design

According to Statistical Atlas (2018), families and households in Detroit are divided into five categories: married, single female, single male, one-person, and non-family. One-person households are the most prominent household type, taking up 38.9%. The second largest one is single female families, and the smallest group is other non-family households. In the Islandview neighborhood, one-person households dramatically increase to 63%. As the second largest group, 18.9% of single-mom families live in this neighborhood. Compared with single-mom families, single-dad is the smallest group. According to the literature, we learned that a co-living pattern would help one-person groups in the aspects of promoting social interaction, creating a sense of community, and engendering mutual support (Izuhara et al., 2022). Collaborative housing also assists one-person groups in satisfying their emotional needs, resolving single moms' babysitting demands, and promoting future community sustainability.

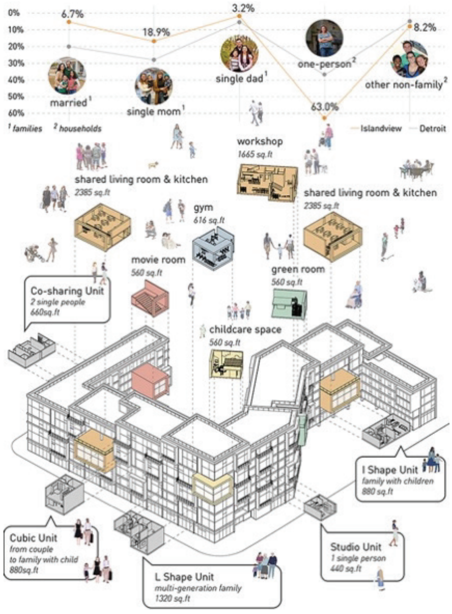
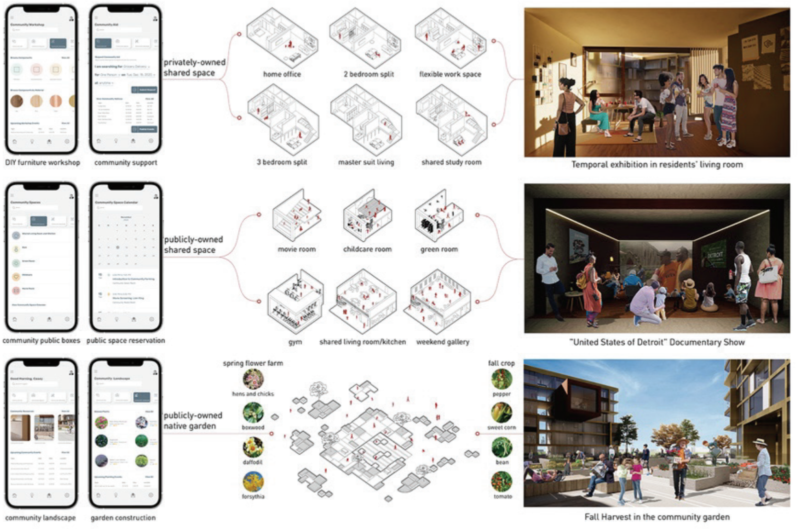


Figure 6. The design of public space in the open-source community

Based on the potential public activities in the Navigation result, communal spaces with diverse public activities are input inside the housing apartment to different types of residents shown in Figure 6. The five different housing types, co-sharing units, cubic units, L shape units, studio units, and I shape units are designed for the five main household types of the Islandview neighborhood. Residents, as the users and managers of the shared spaces in the community housing, collectively use the shared living room and kitchen, movie room, gym, workshop, and green room to interact with other occupants. For instance, for food sharing: one-person households might receive meals from their neighbors in the shared kitchen. For childcare: single moms would leave their children in the childcare space, and the elderly in the housing can take care of them. In addition, for community gardening, residents can plant vegetables together and enjoy the seasonal harvest. The diversity and inclusiveness of the human-centric design promote the connection between people and create a better living environment for them.

The Open Source Community provides physical public and private shared space and builds a collaborative network between residents through the online app. Figure 7 indicates residents' interaction and communication with each other on the app platform. Residents can publish posts about their living room status, especially when they decide to hold a party. Users are invited to either be the initiator or participants of the shared space activities by sharing their private space or reserving the public boxes. They select their interested public boxes and reserve the public-owned shared space in order to avoid time conflicts. They can check the public spaces' schedule and learn about each space's status and activities at any time inside the community. Community members with expertise in gardening are welcome to discuss their planting experience with others for the community landscape in the courtyard. Users in the open-source community platform can view the plant species, construct the garden, and join the planting and harvesting festivals in the community. The platform establishes an invisible bond between the community users and reflects the user's needs.

Figure 7. Interactive community open-source online app platform



5. Conclusion

This open-source community design provides a useful design framework for the community housing design process, assisting architects and designers in understanding the personal views about the city and user needs on social account platforms. It, in turn, helps designers design public spaces in community housing and enrich the social life of residents through the open-source online platform, which turns the community housing design from data-driven to human-centric.

Starting from Twitter, one of the most popular social platforms, we collected image information about Detroit city through keyword searches, hashtags, and official news accounts. We then used machine learning to perform feature extraction and cluster analysis via PixPlot and zoom into the "Public Activities" cluster in the navigation step according to the design task. Taking the lacking public activities in Detroit into consideration, we designed a series of public spaces in community housing and developed an open-source community online platform to help residents communicate and collaborate.

The artificial intelligence analysis approach offers an effective and efficient way of communicating with Detroit. In contrast to the traditional community housing design, its user-oriented approach breaks the top-down computational design thinking. Our work demonstrates how traditional architectural design can be bettered to connect people and shape a healthier, liveable, and equitable environment for residents with advanced new computational technology. In terms of further application, it can be the reference for other community housing projects, helping the developers and stakeholders to better learn about what people want inside the housing and design for users.

Future work will concentrate on improving the data collection technology and adapting to more application scenarios. Inputting more source data in the data collection procedure, such as gathering textual data from Twitter, would vastly enrich the clustering analysis. Furthermore, other social media platforms, including Facebook and Instagram, have the potential to amplify the data collection attributes, effectively improving the data's diversity and inclusiveness.

Acknowledgements

We would like to thank Zishen Wen, who advised the visualization work, and Huiting Du, Zhuzhen Zhang, Casey Ash, who participated in the modeling of the building.

References

Alvarez-Marin, D., & Ochoa, K. S. (2020). Indexical Cities: Articulating Personal Models of Urban Preference with Geotagged Data (arXiv:2001.10615). arXiv. https://doi.org/10.48550/arXiv.2001.10615

Cai, C., Guo, Z., Zhang, B., Wang, X., Li, B., & Tang, P. (2021). Urban Morphological Feature Extraction and Multi-Dimensional Similarity Analysis Based on Deep Learning Approaches. Sustainability, 13(12), Article 12. https://doi.org/10.3390/su13126859

Hou, Q., Han, M., Qu, F., & He, J. S. (2021). Understanding social media beyond text: A reliable practice on Twitter. Computational Social Networks, 8(1), 4. https://doi.org/10.1186/s40649-021-00088-x

Hsu, Y. C., Huang, T., Verma, H., Mauri, A., Nourbakhsh, I., & Bozzon, A. (2022). Empowering local communities using artificial intelligence. Patterns (New York, N.Y.), 3(3), 100449. https://doi.org/10.1016/j.patter.2022.100449

Izuhara, M., West, K., Hudson, J., Fernandez Arrigoitia, M., & Scanlon, K. (2022). Collaborative housing communities through the COVID-19 pandemic: Rethinking governance and mutuality. Housing Studies, 38, 1–19. https://doi.org/10.1080/02673037.2022.2077919

Norman, D.A. (2013). The design of everyday things. The MIT Pr ess.

Rian, I. M., & Asayama, S. (2016). Computational Design of a nature-inspired architectural structure using the concepts of self-similar and random fractals. Automation in Construction, 66, 43–58. https://doi.org/10.1016/j.autcon.2016.03.010

Rocker, I.M. (2006). When code matters. Architectural Design, 76(4), 16–25. https://doi.org/10.1002/ad.289

Soddu, C. (2020). GENERATIVE CITY DESIGN, Aleatority and Urban Species, Unique, Unrepeatable and Recognizable Identity, like in Nature DOMUS ARGENIA ISBN 9788896610411.

The Demographic Statistical Atlas of the United States (2018). Retrieved December 9, 2022, from https://statisticalatlas.com/place/Michigan/Detroit/Household-Income

Wen, W., Hong, L., & Xueqiang, M. (2010). Application of fractals in architectural shape design. 2010 IEEE 2nd Symposium on Web Society, pp. 185–190. https://doi.org/10.1109/SWS.2010.5607455

Yeh, I.-C. (2006). Architectural layout optimization using annealed neural network. Automation in Construction, 15(4), 531–539. https://doi.org/10.1016/j.autcon.2005.07.002

PREDICTION AND OPTIMISATION OF THE TYPICAL AIRPORT TERMINAL CORRIDOR FAÇADE SHADING USING INTEGRATED MACHINE LEARNING AND EVOLUTIONARY ALGORITHMS

YINYI SHI[1], JINHUI CHEN[2], GUANGZHI HU[3] and CUNYUAN WANG[4]
[1]Guangdong Architectural Design and Research Institute Co., Ltd.
[2] Beijing Institute of Architectural Design South China Centre.
[3]Arcplus Group PLC East China Architectural Design and Research Institute Co., Ltd.
[4]Guangzhou City Construction and Development Co., Ltd.
[1]shiyinyi_work@163.com, 0000-0002-5305-1333
[2]chenjinhui1996@qq.com, 0000-0001-5675-4125
[3]406336056@qq.com, 0000-0002-7234-3219
[4]2426452800@qq.com, 0000-0002-0529-7566

Abstract. For airport terminal buildings, the use of large glass curtain walls is beneficial to introduce daylight, but it also tends to cause excessive partial illumination, uncomfortable glare, and an increase in energy load. An appropriate façade shading design is important to help improve indoor environment performance and comfort. In this research, integrated machine learning algorithms were used to train a total of 2187 data samples of louvre shading and film shading to build a performance prediction model for adjustable façade shading of the typical terminal corridor, enabling rapid simulation and optimisation. The results show that the optimisation objectives are more closely related to the adjustable shading components and their façade areas. Besides, the optimal solutions for film and louvre shading are more influenced by the energy and light environment indicators respectively. Different target weights have a different impact on the selection of preferred solutions. This research further enriches the parametric model library and provides a reference for future design decisions and performance evaluation of airport terminal façade shading.

Keywords. Airport Terminal Corridor, Façade Shading, Machine Learning, Multi-objective Optimisation, Performance Prediction, SDG9

1. Introduction

With the continuous improvement of the public transport system, the construction of airport terminal buildings in China has gradually increased. By the end of the 14th Five-Year Plan for Civil Aviation, the number of civil airports in China will reach

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 59-68. © 2023 and published by the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA), Hong Kong.

270 (Yan Song, 2022). The environmental comfort and energy efficiency of large public transport buildings such as airport terminals are receiving increasing attention in their design process, and the evaluation, optimisation and prediction of their light environment and energy load have important potential in terms of green building and low-carbon. As the main walk-through and waiting space for passengers before boarding, the façade of the terminal corridor usually organizes with large glass curtain walls, so an appropriate design of the façade shading is important for indoor environment creation.

Previous research showed that performance simulation and algorithm-driven automated performance feedback have become common design tools, which compare multiple solutions and perform exhaustive simulations and optimisation of each solution's parameters directly according to their algorithmic preferences, requiring significant time and computational resources before the optimisation results converge (Carlucci, S et al., 2015). In recent years, with the use of machine learning algorithms in architecture, the training and prediction of existing data can help designers find out the impact of different indicators to guide their decisions, thus enabling artificial intelligence to assist in building design (Kalyanam, R. and Hoffmann, S., 2021). However, most of the previous research had focused more on small to medium-scale building spaces such as educational spaces and offices or a type of indicators such as light and energy respectively (Hainan Yan, Ke Yan and Guohua Ji, 2022; Liping Wang and Michael J. Witte, 2022). Few researchers have taken large-scale spaces such as airport terminals with year-round combined performance into account. To address the above shortcomings, this research combines machine learning and multi-objective optimisation algorithms to construct a predictive model for façade shading design and performance of typical corridors in terminal buildings, which can quickly predict and evaluate performance and provide a reference for the refined design of the terminal building.

2. Methodology

2.1. OVERVIEW OF THE WORKFLOW

Figure 1 illustrates the workflow of this research, which contains the following four parts. Firstly, a parametric model library of the typical corridor and façade shading was established in the Rhino and Grasshopper platform after analysing and determining the spatial volumes and façade shading types of constructed cases. In the second stage, the light environment indicators and cooling load of the typical parametric model were batch simulated using the Ladybug and Honeybee performance simulation plug-in, while the model parameters (such as volume length, width and height dimensions, room angle, shading component dimensions, etc.) were written correlated with the simulation indicators. In the third stage, Python and integrated machine learning algorithms were used to correlate and train the collected database, to derive the correlation between different parameters and indicators and the magnitude of influence, which helped build a prediction model. Finally, the evolutionary algorithm in Wallacei plug-in was used to iteratively find the optimal solution, then compared the comprehensive scores under different weight assignments to filter out the optimal solution respectively.

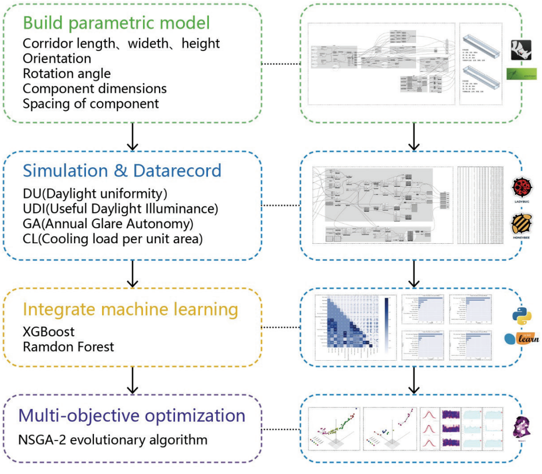


Figure 1. Overview of the workflow

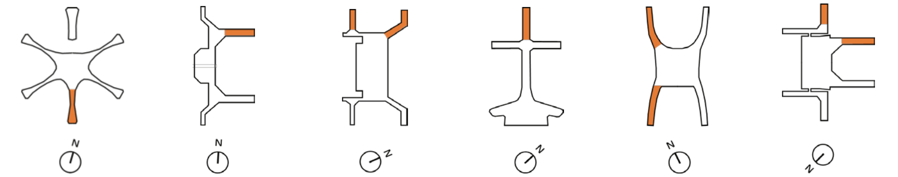
2.2. CASE STUDY AND PARAMETRIC MODEL

Table 1 shows the plan configuration and corridor pattern of some Chinese airports in main cities. The volume parameters of airport corridors are in the range of 200-300m in length, 30-50m in width and 18-30m in height with a variety of orientations, which are used to determine the volume parameters of a typical corridor in this research. The plan configuration of these airport terminal corridors is usually rectangular or quasi-rectangular, while horizontal louvres or films are mainly used as façade shading components. Figure 2 shows the parametric modelling logic of a typical terminal corridor under two types of shading with a total of 2187 parametric models respectively.

Combined with the experience of design, the impact of parameters such as room surface area, room volume, adjustable façade area and area of shading elements derived from the basic parameters should be taken into account. Ultimately, the length, width and height of a typical corridor, as well as the values of the louvre and film, are defined as the basic parameters of the direct control parameterisation model, and 5 and 6 derived parameters were added for film and louvre shading types respectively. The basic and derived parameters will be used in the analysis in subsequent chapters together.

Table1. Airport terminal configurations and corridors in main cities of China

Beijing	Shanghai	Guangzhou	Shenzhen	Chongqing	Xi'an



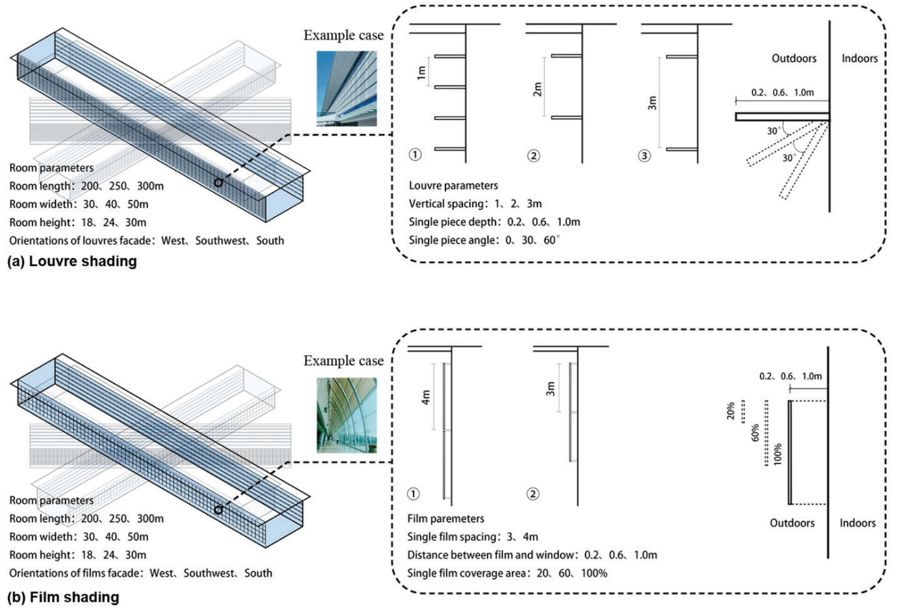


Figure 2. Parametric modelling under louvre and film shading

2.3. LIGHT ENVIRONMENT AND ENERGY SIMULATION

2.3.1. Simulation indicators

All-year-round usage needs and the flexibility of the functional spaces are the main features of a terminal corridor, this research combines relevant norms and standards while selecting a combination of static and dynamic light environment indicators as well as the annual total cooling load per unit area as the simulation indicators. The specific indicators are described as follows.

- Daylight Uniformity (DU) - An indicator based on the Daylight Factor (DF), the value of which is equal to the ratio of the minimum to the average DF value among all measured points throughout the year.
- Useful Daylight Illuminance (UDI) - The percentage of hours of effective daylight throughout the year, which indicates the percentage of hours that the measured point illuminance values within the target range. The effective illuminance constraint range in this research is 450-2000lux.
- Glare Autonomy (GA) - The percentage of hours in a specific range of Daylight Glare Probability (DGP) values for each measured point on the indoor working surface throughout the year, with a DGP constraint range of less than 0.4 in this research.
- Annual Cooling Load Per Unit Area (CL) - Indicates the cooling load per unit area of the overall space.

2.3.2. Performance simulation and dataset collection

The light environment and energy load simulation were performed in the Ladybug and Honeybee plug-in based on the Rhino and Grasshopper platform, while the external climate environment was selected from the EPW file of typical meteorological data in Guangzhou city of China. Table 2 shows the specific simulation settings of other model materials. It is worth noting that the critical temperature for air conditioning and cooling is 26 degrees Celsius, and the density of the corridor's pedestrian flow is entered according to one person per 20 ㎡ referring to engineering specifications, while the simulation period covers 8760 hours throughout the year.

After the basic setup, batch simulations are carried out with the Ladybug fly battery. The data for the different models and the corresponding indicator simulation results are recorded and saved with further access to the CPython tool.

Table2. Light environment and energy load simulation settings

Light environment simulation settings				Energy simulation settings
Material	Reflectance	Transmittance	Colour	Curtain wall: U 0.43 SHGC 0.26
Glass curtain	-	0.57*0.6*0.9[a]	Light blue-grey	Dbl Ref-B-H Clr 6mm/13mm Air
Metal roof	0.85	-	Light grey	Roof: Typical Uninsulated Metal
Marble floor	0.6	-	Grey	Building Roof (U=6.11W/㎡•k)
Louvres	0.5	-	Light grey	Cooling temperature: ⩾26° C
Films	0.75	0.15	White	Personnel load: 1 person/20 ㎡

a. 0.57 indicates the transmittance of the glass, and 0.6 and 0.9 indicate the light-blocking reduction factor and loss factor for alloy windows.

2.4. CORRELATION ANALYSIS AND MACHINE LEARNING

In this research, correlation analysis of 13 and 12 parameters with 4 performance indicators under louvre and film shading were conducted to the relationship respectively before the machine learning process. XGBoost and Random Forest integrated algorithms were selected in the machine learning process, which stands for the typical representatives of Bagging and Boosting algorithms that can be used to solve prediction problems for unstructured data (A. Paul, D.P. et al., 2018; M. Saha, S et al., 2021). The machine learning used in this research belongs to supervised learning, so 85% of the collected data were used as the training set and 15% as the verification set, and the fitted model is achieved well through hyperparameter tuning.

2.5. MULTI-OBJECTIVE OPTIMISATION AND WEIGHTS COMPARISON

Wallacei engine, based on the NSGA-2 evolutionary algorithm, was selected as the multi-objective optimisation plug-in, which helped realise the capability of iterative optimisation seeking with machine learning prediction model data, storing the data and obtaining some preferred solutions for designers' reference. In the meantime, Wallacei cluster analysis and a comprehensive score calculation function had been used to compare optimal solutions under different types and target weight assignments according to relevant research methods (Wei Xue, 2011; John A. Rice, 2006). The calculation functions are shown as follows.

The combined score calculation function can be represented as:

$$Score_{overall} = \omega_{UDI} Score_{UDI} + \omega_{DU} Score_{DU} + \omega_{GA} Score_{GA} + \omega_{CL} Score_{CL}$$

The individual scores of UDI, DU and GA can be calculated in the same way. The following equation represents the score of UDI.

$$Score_{UDI} = \left(1 - \frac{UDI - UDI_{min}}{UDI_{max} - UDI_{min}}\right)^2$$

The individual scores of CL can be represented as the following equation.

$$Score_{CL} = \left(\frac{CL - CL_{min}}{CL_{max} - CL_{min}}\right)^2$$

Where ω denotes the target weight, $\omega_{UDI} + \omega_{DU} + \omega_{GA} + \omega_{CL} = 1$.

3. Result and Discussion

3.1. CORRELATION ANALYSIS

According to the statistical test, the parameters and the optimisation objectives were not normally distributed, so the Spearman correlation coefficient was selected for the correlation analysis.

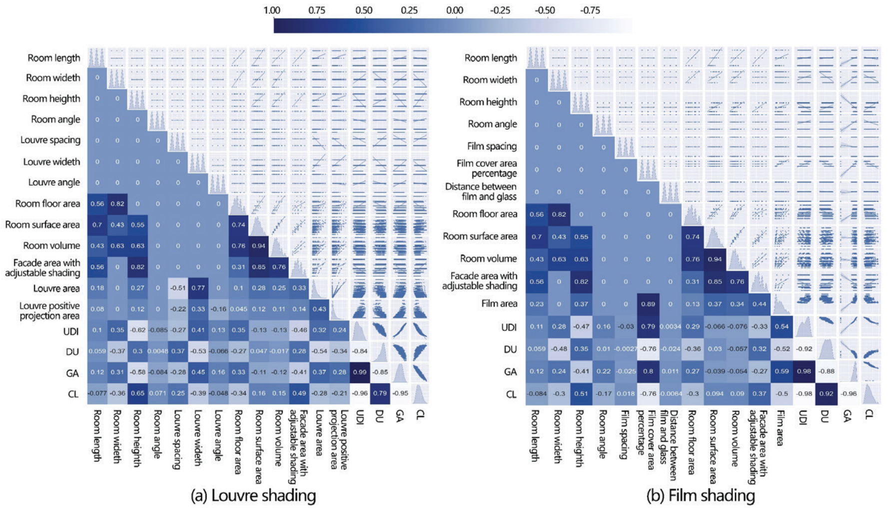


Figure 3. Correlation analysis matrix between parameters and the optimization objectives

As can be seen from Figure 3, the correlation trends of the parameters with UDI and GA are similar, and the same trend characteristics can be found in DU and CL in general. Under the louvre shading type, the correlation between the corridor volume height, adjustable louvre shading façade area, louvre width and UDI, GA, and CL are higher than others, while for DU, only the correlation coefficient of louvre area is greater than 0.5 in absolute value. Under the type of film shading, the correlation between the simulation indicators and the film cover area percentage is higher, with the absolute value of the correlation coefficient close to 0.8, followed by a certain correlation with the height and width of the room volume, the total area of film and the area of the adjustable shading façade, while the correlation with the other indicators is weaker.

3.2. MACHINE LEARNING AND PREDICTION MODEL

In this research, the fit perturbation R^2 and the root mean squared error RMSE were selected as the evaluation index of the machine learning fitting effect, and the smaller RMSE and larger R^2 portend a better prediction model. As can be seen from Table 3, the R^2 of both algorithms exceeded 0.95, reaching a maximum of 0.98, and the RMSE was in the range of 0.051-4.299, which indicate a better fitting effect. Therefore, the XGBoost algorithm training results are selected as an example and carried out feature importance analysis for louvre shading and film shading types respectively. Figure 4 shows that the film cover area percentage has the greatest influence on the prediction results of all optimisation objectives, and together with the total area of film, the volume height of the corridor space and the area of the adjustable shading façade under the film shading type. Besides, the adjustable shading façade area and the shading louvre area have a greater influence on the prediction results under the louvre shading type. All of them need to be included as independent variables during the multi-objective prediction process.

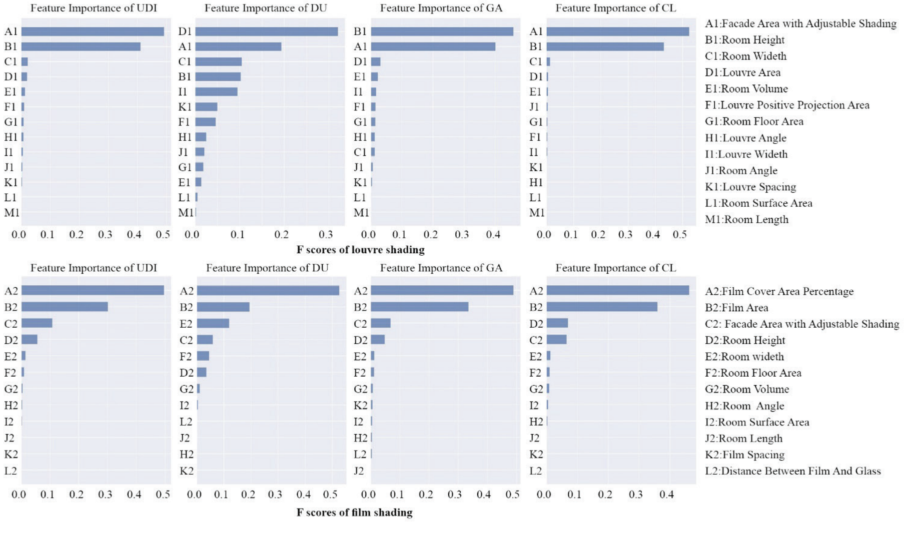


Figure 4. Results of the feature importance analysis of different optimisation objectives

Table3. Comparison of Prediction Effectiveness between Ramdon Forest and XGBoost Algorithms

Algorithms and evaluation indexes			UDI	DU	GA	CL
Film shading	XGBoost	R^2	0.987	0.977	0.972	0.984
		RMSE	1.648	0.051	1.261	2.981
	Ramdon Forest	R^2	0.988	0.976	0.976	0.987
		RMSE	1.514	0.052	1.186	2.878
Louvre shading	XGBoost	R^2	0.977	0.961	0.968	0.976
		RMSE	0.157	0.200	0.185	3.942
	Ramdon Forest	R^2	0.974	0.957	0.968	0.971
		RMSE	0.168	0.211	0.185	4.299

3.3. CLUSTER ANALYSIS AND SOLUTION SELECTION UNDER DIFFERENT WEIGHTS

3.3.1. Multi-objective Optimisation and Cluster Analysis

The multi-objective optimisation results gradually stabilised after 20 generations and reached a convergence trend. As shown in Figure 5, the 100th generation of results was selected for cluster analysis, and the 100 frontier solutions of the last generation obtained from the optimisation were divided into five clusters.

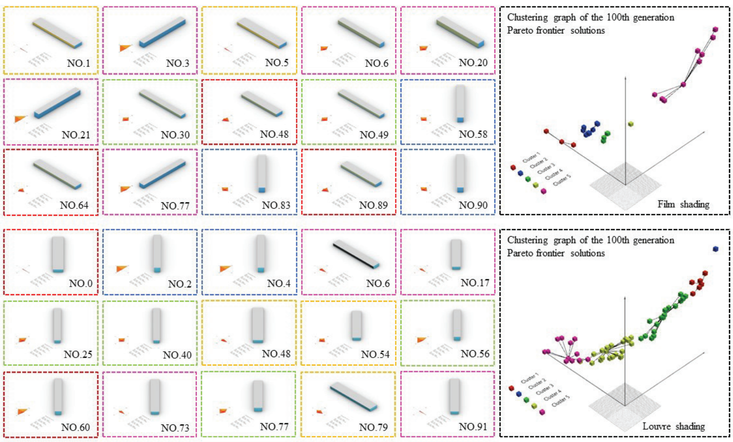


Figure 5. Spatial clustering of 100th generation Pareto front solutions with typical cases

The visualisation of the model characteristics shows that the preferred solutions cover most of the situations under the film shading type, indicating that the indoor light environment can be improved and energy-saving goals achieved by adjusting the parameters of the shading film for different volumes and orientations. But in the case of the louvre shading, the preferred solutions are mostly for the southwest and south orientation, indicating that adjusting the parameters of the shading louvre alone has a smaller contribution to the light environment to energy load than the south and southwest orientations.

3.3.2. Comprehensive preferred solutions and weight comparison

As shown in Table 4, the optimal solution and the values of each optimisation objective are divided with a weight of 0.25. By comparing the overall scores of different clusters, we can see that the typical optimal solution is NO.23 in cluster 2 under the type of film shading and NO.49 in cluster 4 under the type of louvre shading. Both of which show a greater optimisation effect.

In addition, as can be seen from Table 5, this research further compares the impact of different weights on the preferred solutions to help match the flexible needs in the design stage. Different weights have an impact on the optimal solution.

Table4. Comparison of Prediction Effectiveness of Random Forest and XGBoost Algorithms

Cluster types		Individual	UDI (%)	DU	GA (%)	CL (W/㎡•k)	$Score_{overall}$
Film shading	Cluster1	NO.82	80.61	0.64	90.73	151.50	0.0966
	Cluster2	NO.23	78.23	0.68	91.97	159.07	0.0763
	Cluster3	NO.30	69.29	0.74	87.34	157.41	0.0864
	Cluster4	NO.16	85.04	0.60	94.75	137.77	0.1125
	Cluster5	NO.71	57.16	0.79	81.91	173.28	0.1863
	Maximum value		86.76	0.85	95.10	236.60	-
	Minimum value		37.47	0.47	65.41	130.09	
Louvre shading	Cluster1	NO.60	45.32	0.82	71.52	198.18	0.4183
	Cluster2	NO.2	37.48	0.85	64.20	229.24	0.7500
	Cluster3	NO.9	62.60	0.74	84.24	168.09	0.1130
	Cluster4	NO.49	72.52	0.69	87.97	151.11	0.0560
	Cluster5	NO.73	76.78	0.61	90.27	148.71	0.0865
	Maximum value		82.06	0.85	92.91	229.24	-
	Minimum value		37.48	0.42	64.20	138.13	

Table 5. Comparison of the combined scores of the preferred solutions under different weights

Type	Weight	Individual	$Score_{overall}$
Film shading	ω 1：ω_{UDI}=0.25, ω_{DU}=0.25, ω_{GA}=0.25, ω_{CL}=0.25	NO.23	0.0763
	ω 2：ω_{UDI}=0.3, ω_{DU}=0.3, ω_{GA}=0.3, ω_{CL}=0.1	NO.23	0.0768
	ω 3：ω_{UDI}=0.15, ω_{DU}=0.15, ω_{GA}=0.15, ω_{CL}=0.55	NO.16	0.0696
Louvre shading	ω 1：ω_{UDI}=0.25, ω_{DU}=0.25, ω_{GA}=0.25, ω_{CL}=0.25	NO.49	0.0560
	ω 2：ω_{UDI}=0.3, ω_{DU}=0.3, ω_{GA}=0.3, ω_{CL}=0.1	NO.46	0.0574
	ω 3：ω_{UDI}=0.15, ω_{DU}=0.15, ω_{GA}=0.15, ω_{CL}=0.55	NO.49	0.0417

4. Conclusion

This research uses integrated machine learning algorithms to compare and analyse the impact between the model parameters and optimisation indicators of the terminal corridor under two types of shading and then combines multi-objective optimisation algorithms with a comprehensive score under different weights to select solutions. The main conclusions are shown as follows.

- The film, louvre area and adjustable façade area have the most significant degree of impact on the light environment and energy load indicators of the corridor.
- The influence of film parameters on the light environment and energy load of the corridor is more pronounced than that of louvres.
- Energy load and light environment have a greater impact on film shading and louvre shading respectively.

Research workflow and results enable rapid feedback on the design and optimisation process of typical terminal corridor shading, which has implications for sustainable green buildings. However, it is worth noting that the prediction model constructed in this research is based on performance simulation values, and there is a certain degree of error, which should be combined with field measurement to improve the prediction accuracy during the subsequent research. At the same time, The input parameters considered in this research are focused on the building itself, so the climate situation and some intelligent control technology will be added to provide a better reference for similar designs in subsequent research.

References

A. Paul., D.P. Mukherjee., P. Das., A. Gangopadhyay., A.R. Chintha., & S. Kundu. (2018). Improved random forest for classification. *IEEE Transactions on Image Processing,* 27 (8), 4012-4024. https://doi.org/10.1109/TIP.2018.2834830.

Carlucci, S., Cattarin, G., Causone, F., & Pagliano, L. (2015). Multi-objective optimization of a nearly zero-energy building based on thermal and visual discomfort minimization using a non-dominated sorting genetic algorithm (NSGA-II). *Energy and Buildings*, 104, 378 – 394.https://doi.org/10.1016/j.enbuild.2015.06.064.

Yan, H., Yan, K., & Ji, G. (2022). Optimization and prediction in the early design stage of office buildings using genetic and XGBoost algorithms. Building and Environment, 218, 109081. https://.doi.org/10.1016/j.buildenv.2022.109081

Rice, J.A. (2006). *Mathematical Statistics and Data Analysis.* Machinery Industry Press.

Kalyanam, R., & Hoffmann, S. (2021). A Reinforcement Learning-Based Approach to Automate the Electrochromic Glass and to Enhance the Visual Comfort. Applied Sciences, 11(15), 6949. MDPI AG. https://doi.org/10.3390/app11156949

Wang, L., & Witte, M. J. (2022). Integrating building energy simulation with a machine learning algorithm for evaluating indoor living walls' impacts on cooling energy use in commercial buildings. Energy and Buildings, 272, 112322. https://doi.org/10.1016/j.enbuild.2022.112322.

Saha, M., Nayak, S., Mohanty, N., Baral, V., & Rout, I. (2021). Preterm delivery prediction using gradient boosting algorithms. Communication and Intelligent Systems, 59–68. https://doi.org/10.1007/978-981-16-1089-9_6

Wei, X. (2011). *Statistical Analysis and SPSS Application.* People's University of China Press.

Yan, S. (2022).The 14th Five-Year Plan for the Development of Civil Aviation. *Civil Aviation Administration of China (CAAC)*. Retrieved November 16, 2022, from http://www.gov.cn/xinwen/2022-01/07/content_5667004.htm.

RESEARCH ON ARCHITECTURAL SKETCH TO SCHEME IMAGE BASED ON CONTEXT ENCODER

YUQIAN LI[1] and WEIGUO XU[2]
[1,2] TSINGHUA UNIVERSITY, Beijing, China
[1]liyuqian17@mails.tsinghua.edu.cn
[2]xwg@mail.tsinghua.edu.cn

Abstract. Architects are used to hand drawing sketches to express the architectural creation intention. To present these abstract sketches, architects and teams need to convert sketches into architectural scheme images, which requires a lot of time and labour. Deep learning may have the potential to improve the efficiency of this work. The common sketch-to-image generation is based on Generative Adversarial Network (GAN), and the research of edge-to-image has made a big progress. But these methods require strict alignment of data pairs, which is difficult to achieve. Zhu et al. proposed the loss of Cycle-Consistent, which solved the problem that pairs of data sets are difficult to collect. However, most of the image translation methods require strict alignment between image data pairs, which can be achieved only for the edge mapping extracted from the image; but the sketch is very different from the edge. Due to the abstractness and fuzziness of the sketch, any simple distortion cannot complete the task of providing pixel-level alignment between the sketch and image; And image translation is the transfer of image features such as colour and texture. The original image has a strong constraint on the generated image, which makes the original structure of the image impossible to be changed. By image inpainting, we address this topic using a joint image completion approach with Context-Encoder, where the architectural sketch provides the image context for generating the scheme images. This setting has two advantages: first, the joint images can avoid the complexity of cross modal problems and the strict alignment of the data pairs as image-to-image translation; second, because of the weak constraint, the outputs have greater freedom, which perhaps can generate more imaginative results. The Context-Encoder generates scheme images on the data sets of general architectural sketches. The results present that the applicability of the completion method is better than that of the method of image translation. And scheme images that is different from the original architectural sketch contours have been generated.

Keywords. Sketch, Building Scheme Image, Image Completion, Context-Encoder.

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 69-78.

1. Introduction

In the process of promoting architectural design concepts, to make abstract sketches understood by others and promote team communication and cooperation, architects will convert sketches into 3D models in computers to make the concepts more intuitive. But this process needs a lot of time and labour. If we use AI to directly convert architectural sketches into clear and accurate architectural scheme images it will be more effective.

Sketch-based image generation is a classical cross modal problem in sketch research. The generation method of sketch to image involves sketch to image, sketch and image to image, sketch/edge, and colour to image. With the development of Generative Adversarial Network (GAN) (Goodfellow,2014), Conditional Generative Adversarial Network (CGAN) (Mirza,2014) method has been tried to solve the problem of image-to-image translation (sketch is also a type of image). Some progress has been made in the research of edge image to image, such as Pix2pix (Isola,2016), Pix2pixHD (Wang,2017), CycleGAN (Zhu,2017), etc. But these works require strict alignment between image data pairs, which can be achieved for edge mapping extracted from images. However, there is a huge gap between the sketch and the edge map. When the sketch and image are used as data pairs, the above requirements are seriously violated. The existing research have evaluated the adaptability of the above methods in sketch and image data pairs, and the results are not satisfactory. It can be imagined that any simple distortion cannot provide pixel level alignment between the sketch and the image due to the abstractness and fuzziness of the sketch.

In fact, image-to-image translation transfers the colour, texture, and other features of the image. However, for architectural sketches, geometric relations, that is, computer understanding of the shape semantics of buildings, are more important. In the existing methods of edge image to image translation, the edge image has a strong constraint on the generated image, which makes the structure of the edge image cannot be changed. However, if this constraint is turned into a guidance, that means, a weak constraint, then the output edge may not necessarily follow the input edge.

Influenced by the field of image inpainting, this paper attempts to apply the method of image inpainting to the task of sketch to image translation. Image inpainting is a problem between image editing and image generation. It aims to reconstruct the unknown information of an image in a present way, to restore the image to the maximum extent. Its essence is to build some known information into complete information. High quality image inpainting requires not only reasonable semantics of the generated content, but also clear and real image texture.

The image inpainting task generally includes two sub tasks -- the removal of some parts of the image and the image completion. In this paper, the method of image inpainting is used to translate sketch to image, which only involves the part of image completion.

Image completion needs to analyse, understand, and inpaint the incomplete image based on human vision and cognitive rules, to meet the requirements of visual reality, excessive naturalness and compound context semantics. The task of generating an architectural scheme drawing from an architectural sketch is very similar to the image completion task. The generation of architectural scheme drawing is based on the sketch as the context. The computer needs to translate the sketch based on fully analysing and understanding the semantics of the sketch. Therefore, the sketch and the corresponding natural image can be integrated into a new joint image, and the image pair input in the general sense of image translation can be transformed into the joint image of the sketch and the natural image. The missing part that needs to be completed is the part of the natural image.

As shown in the fig. 1, in the general image inpainting task, for an image to be completion, the algorithm will make a reasonable infer to fill the mask part according to the content around it; while in this task, the original image is a composite one of a sketch image and a scheme image, where the position of the scheme image is the masked part, and the algorithm also needs to take the content around the mask, i.e., the sketch, for the masked part, i.e., the scheme image. So the algorithm inpaints the masked part, i.e., the scheme image, with the content of the masked surrounding, i.e., the sketch.

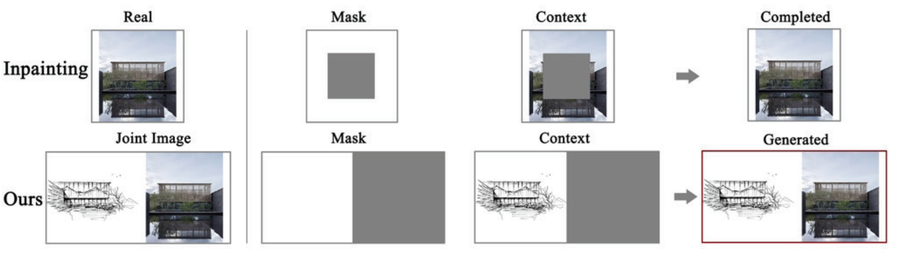


Figure 1. Image generation posed as image completion.

2. Related Works

2.1. IMAGE-TO-IMAGE TRANSLATION

The common sketch-to-image generation is based on Generative Adversarial Network (GAN), and the research of edge-to-image has made a big progress: Isola et al. proposed a unified framework to solve the problem of image translation (Isola,2016); Wang et al. improved the model structure to generate high-resolution images (Wang,2017). The above require strict alignment of data pairs, which is difficult to achieve. Zhu et al. proposed the loss of Cycle-Consistent (Zhu,2017), which solved the problem that pairs of data sets are difficult to collect.

Also, these methods are tried to be applied to solve the architectural sketch-to-scheme drawing(Chan,2020)(Li,2022), which have been successful. However, most of the image translation methods require strict alignment between image data pairs, which can be achieved only for the edge mapping extracted from the image; but the sketch is very different from the edge. Due to the abstractness and fuzziness of the sketch, any simple distortion cannot complete the task of providing pixel-level alignment between the sketch and image; And image translation is the transfer of image features such as color and texture. The original image has a strong constraint on the generated image, which makes the original structure of the image impossible to be changed.

2.2. IMAGE INPAINTING

Existing image inpainting techniques can be divided into two different groups: traditional and deep learning-based methods.

Traditional image inpainting techniques include three groups of different image inpainting algorithms: first, based on diffusion (Bertalmio,2020), image is reconstructed by solving partial differential equations; The second group is based on patch method (Darabi,2012), using texture synthesis to inpaint the missing part of the image; The third group is based on the method of convolution filtering (Hadhoud,2008), using convolution operators to reconstruct images.

When the missing area is very small, the diffusion-based technology can produce accurate results, but with the increase of the size of the missing area, the diffusion process will make the image become blurred. The patch-based technology provides better performance when filling in the missing large areas. However, such methods rely on low-level features, ignore visual semantics, and fill in missing areas, and are ineffective in inpainting complex structures, resulting in poor visual quality of the image and inability to generate new objects that cannot be found in the source image. The algorithm based on convolution filter provides a simple, fast, and accurate method to inpaint small regions. However, with the increase of the area to be repaired, the output image becomes more blurred.

The above traditional technologies cannot generate new objects that cannot be found in the source image, and they cannot produce semantically consistent results. This greatly limits traditional technologies and makes them unable to reconstruct complex structures. To solve this problem, the model needs to obtain information from external sources.

The Deep Learning techniques fulfil this purpose by training a model to recognize the mapping between images and their labels. The model is trained using external data (one or multiple datasets). Once the training is completed it proceeds to reconstruct the image. First approaches with DL provided faster and more precise ways for image reconstruction however, it still introduced some blurriness and completely lacked a semantic understanding of the image. With the introduction of Context Encoders (Pathak,2016), DL techniques where able to produce coherent images, reconstruct complex structures and produce novel objects, while maintaining local consistency. The blurriness on the results is produced by the models being ineffective in copying information from distant spatial locations. This was addressed by combining a DL approach with a patch-based technique, providing a new purpose for patch-based techniques that were under the radar for some time. (Rojas,2020)

In this paper, because architectural sketch is the original basis for the generative of architectural scheme images, there is a context semantic relationship between the sketch and the scheme images. Therefore, we attempt to implement the translation task from architectural sketch to architectural scheme images by using the image inpainting method that introduces context encoder.

3. Method

In this paper, we choose the most classical image inpainting method to study the generative of architectural sketches to images. In 2016, Pathak et al. proposed an image inpainting network called Context Encoder, which applies unsupervised feature learning driven by pixel prediction based on context to large hole image inpaint. The Context Encoder can understand the image semantics to a certain extent and predict the pixels according to the information around the area to generate new content. (Pathak, 2016)

The model architecture is shown in the fig.2. The overall architecture is a simple encoder-decoder. The encoder extracts the feature representation of the input image, and the decoder gradually enlarges the compressed feature image to restore the size of the original image. Since the convolution layer cannot directly connect all the positions in a specific feature map, the encoder composed of convolution layers cannot directly transmit information from one corner of the feature map to another. A method based on step 1 convolution for band group full connection layer to propagate information across channels is proposed as an intermediate connection between encoder and decoder to propagate information in the activity of each feature graph.

Context Encoder uses Reconstruction Loss and Adversarial Loss to deal with the continuity in context and multiple patterns in output. The Reconstruction Loss is responsible for capturing the overall structure of the inpainted area and its consistency with the surrounding visible area. The Adversarial Loss makes the prediction of the inpainted area look real. The best repair effect can be achieved by keeping the balance between the two.

The Context Encoder can understand the image semantics to a certain extent and predict the pixels according to the information around the holes to generate new content. At that time, it was a very cutting-edge image restoration technology. It lays a foundation for the follow-up research.

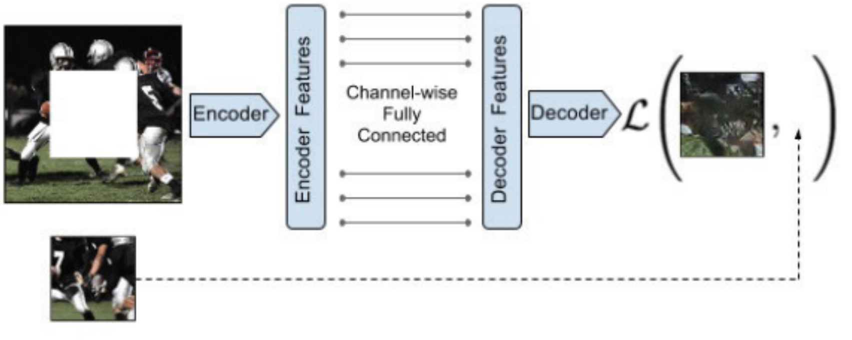


Figure 2. the Context Encoder architecture with an Encoder-Decoder Network

4. Experiment

4.1. DATASET

Training data. First, 500 pairs of data were widely collected from the Internet, that is, the architect's hand-drawn sketches and the building scheme images corresponding to the sketches; Then, 530 pairs of image data are combined, as shown in figure 3, that is, each of the sketches and each of the corresponding scheme images are combined into a new joint image as an input data. Finally, 500 input images for model training are obtained.

Test data. Collect 32 architects' hand drawn sketches from the internet as the test data of the model.

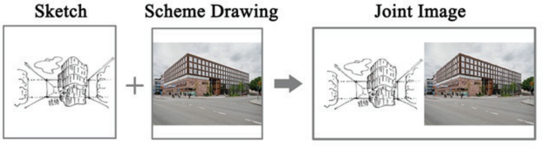


Figure 3. Training data preparation

4.2. TRAINING PROCESS AND TEST PROCESS

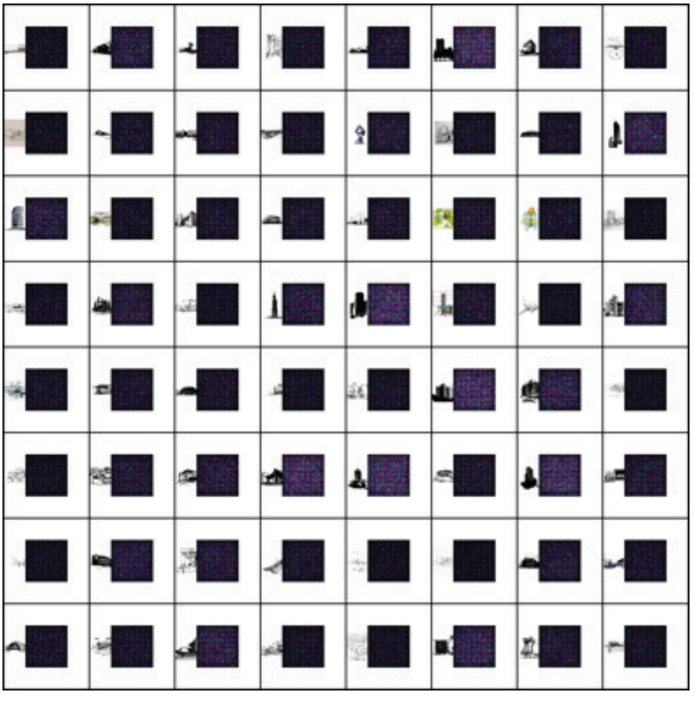


Figure 4. Some Completions in the Training process

500 joint images are used as input data to train the Context-Encoder architecture. Set the learning rate to 0.002 and the number of iterations to 5000. After the model training, 32 test data were used to test the model.

Fig. 4 shows some completions in the training process. As the set at the beginning, the algorithm make the scheme image's part in the joint image as a mask and try to infer it according to the context information of the sketch part.

4.3. RESULTS

Figure 5. Test Results

As we can see in figure 5:

1) It is obvious from the generated results that the image completion model has successfully generated the architectural scheme images

2) In the generation results, the generative of the background and foreground of the scheme drawing is clearly distinguished. The background is mainly blue sky, and the foreground is part of the building.

3) The image completion model has a strong ability to understand the semantics of sketches.

First, we have achieved a good generative of different forms of buildings. As shown in the figure, the generation of tower buildings and buildings with relatively horizontal development can be clearly different.

Secondly, the combination of multiple buildings can also be realized in the scheme images.

Finally, the most important thing is that the generated architectural scheme images can break away from the limitations of the original sketch structure and generate new forms based on the unchanged basic conception.

4) However, we also see some limitations from the results:

The first is a low-resolution output, which is related to the Context Encoder model itself.

Secondly, because the training belongs to small sample training, when there are relatively irregular sketches in the test data, the results are not abundant.

5. Conclusion

The abstractness of architectural sketch is different from the ordinary edge image. The general image translation algorithm can only realize the transfer of simple texture, colour, and other features. It cannot get the ideal architectural scheme drawing. Inspired by the image inpainting task, this paper applies the image completion process, to the generation of architectural sketches to scheme images. Image completion needs to analyse, understand, and inpaint the incomplete image based on human vision and cognitive rules, to meet the requirements of visual reality, excessive naturalness, and compound context semantics. By understanding the semantic information of the sketch, the model can realize the translation of the real and natural scheme drawing. Here, the sketch becomes a guide for the generative of the scheme drawing, and the output scheme drawing can have edges that are not completely consistent with the sketch.

In this paper, the classical algorithm--Context Encoder is used to successfully generate the sketch to scheme drawing, which proves the effectiveness of image completion in the translation of architectural sketches to scheme images.

References

Bertalmio, M., Sapiro, G., Caselles, V., & Ballester, C. (2000). Image inpainting. Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques - SIGGRAPH '00. https://doi.org/10.1145/344779.344972

Chan, Y. H., & Spaeth, B. (2020). Architectural visualisation with conditional generative adversarial networks (cgan). - what machines read in architectural sketches. Proceedings of the 38th International Conference on Education and Research in Computer Aided Architectural Design in Europe (ECAADe),Volume 2. https://doi.org/10.52842/conf.ecaade.2020.2.299

Darabi, S., Shechtman, E., Barnes, C., Goldman, D. B., & Sen, P. (2012). Image melding: Combining Inconsistent Images using Patch-based Synthesis. ACM Transactions on Graphics, 31(4), 1–10. https://doi.org/10.1145/2185520.2185578

Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Bengio, Y. (2014). Generative Adversarial Nets. Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2, 2672–2680. Presented at the Montreal, Canada. Cambridge, MA, USA: MIT Press.

Hadhoud, M.M., Moustafa, K. A.,& Shenoda, S. Z.(2008). Digital images inpainting using modified convolution based method. Proceedings of SPIE - The International Society for Optical Engineering, 1(1). https://doi.org/10.1117/12.816704

Li, Y., & Xu, W. (2021). Using Cyclegan to achieve the sketch recognition process of sketch-based modeling. Proceedings of the 2021 DigitalFUTURES, 26–34. https://doi.org/10.1007/978-981-16-5983-6_3

Mirza, M., & Osindero, S.. (2014). Conditional Generative Adversarial Nets. https://doi.org/10.48550/arXiv.1411.1784.

P. Isola, J. -Y. Zhu, T. Zhou and A. A. Efros (2017). "Image-to-Image Translation with Conditional Adversarial Networks," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA. pp. 5967-5976. https://doi.org/10.1109/CVPR.2017.632.

Pathak, D., Krahenbuhl, P., Donahue, J. , Darrell, T., & Efros, A.A. (2016). Context Encoders: Feature Learning by Inpainting. IEEE. https://doi.org/10.1109/CVPR.2016.278.

Rojas, D. J., Fernandes, B. J., & Fernandes, S. M. (2020). A review on image inpainting techniques and datasets. 2020 33rd SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI). https://doi.org/10.1109/sibgrapi51738.2020.00040

Wang, T.-C., Liu, M.-Y., Zhu, J.-Y., Tao, A., Kautz, J., & Catanzaro, B. (2017). High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs. https://doi.org/10.48550/ARXIV.1711.11585

Werner, L and Koering, D (2020)Anthropologic: Architecture and Fabrication in the cognitive age - Proceedings of the 38th eCAADe Conference - Volume 2, TU Berlin, Berlin, Germany, 16-18 September 2020, pp. 299-308. https://doi.org/10.14279/depositonce-10568

Zhu, J.-Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks. 2017 IEEE International Conference on Computer Vision (ICCV), 2242–2251. https://doi.org/10.1109/ICCV.2017.244

‘;TEXT-TO-GARDEN

Generating Traditional Chinese Garden Design From Text-Descriptions at Scale With Multimodal Machine Learning

SIYUAN ZHANG[1†], YIFAN LI[2†], SIQI ZHANG[3*], XUEXIN HE[4*], RUNJIA TIAN[5]
[1] *Huazhong University of Science and Technology University.*
[2,3,4]*Southeast University.*
[5]*Harvard University.*
*† and * indicates equal contribution to this work*
[1]*siyuanzhang@hust.edu.cn, 0000-0002-2550-3507*
[2]*yifanli0919@gmail.com, 0000-0003-0955-6198*
[3]*skylariki7@gmail.com, 0000-0002-4857-0753*
[4]*3070501583@qq.com, 0000-0001-7989-9300*
[5]*runjiatian@gmail.com,0000-0002-5983-9754*

Abstract. Generations of Chinese architects and architectural scholars have endeavoured to restore the design of demolished gardens by speculating from documentation about Traditional Chinese Garden in the form of text. In previous works, these speculative methods are case-by-case, non-extrapolatable, and usually require comparative studies from various documentation. We propose a novel approach to restore Traditional Chinese Garden designs with multimodal machine learning. More specifically, we created the first Traditional Chinese Garden dataset with paired text descriptions and images. Then we fit DALL·E, a state-of-the-art text-to-image generation model on the dataset, to learn the mapping from a text description of a garden to the actual garden design. During experiments, various data representation and labelling techniques are studied and iterated to improve generation fidelity. In addition, we added a post-processing step to convert the semantic segmentation map to a 3D garden parametrically, closing the loop for an end-to-end generation pipeline. We see the promising potential of applying our methodology to the restoration of other heritage in built environments that face similar dilemmas.

Keywords. Multimodal Machine Learning, Text-to-image Generation, Traditional Chinese Garden, Digital Heritage

1. Introduction

Restoration of Traditional Chinese Garden is vital to the world’s architectural and landscape heritage. According to the City of Suzhou, more than half of the gardens have been demolished throughout history (City of Suzhou, 2022). However, for gardens that are lost throughout history, despite the abundance of text descriptions only a small proportion of them include corresponding picture material resources, making the restoration of Traditional Chinese Garden a challenging task. Previous restoration methods rely heavily on domain knowledge in architecture design, where requiring architects to provide analogical and speculative designs from text descriptions (Hall et al., 2021). Such approach is case-by-case, hard-to-extrapolate and are not scalable to other restoration tasks. Computational approach such as photogrammetry (Brown, A. et al, 2020) and deep learning (Cai, C. and Li, B,2021) have provided new approaches that generates and restores Traditional Chinese Gardens at scale but restricted by visual image scarcity.

On the other hand, Traditional Chinese Garden are architectural spaces, allowing us to tackle the challenge from design perspective. Essentially, design is a multi-model endeavour. Various modalities, such as text descriptions(language) and drawing images(visual) may represent the same design artifact conceptually (Ladron et al., 2022). Recent research in multi-modal machine learning, in particular attension-based image-to-text generation models such as Diffusion (Dhariwal et al., 2021) and DALL·E (Ramesh, A. et al. ,2021) address the gap by learning a joint-embedding of text and images. These models are capable of zero-shot generalize to held out artifacts.

Therefore, we see significant potential of addressing the challenge of Traditional Chinese Garden restoration and generation with recent advancement in multi-modal machine learning, especially text-to-image generation model.

In this study, we propose a novel method that generates Traditional Chinese Garden at scale from text descriptions. Through the study of literature documentation about Traditional Chinese Gardens in the form of text, and a limited number of existing gardens, we synthesized a paired text and image dataset of these Gardens. We then fit DALL·E model on the dataset. We curate the dataset in six groups to search for optimal text and image data representation and found the model with optimal performance on test set. This multimodal approach also provides an alternative for the restoration of other architectural cultural heritage conservation.

2. Related Work

2.1. DIGITAL RESTORATION OF TRADITIONAL CHINESE GARDEN

Existing method to restore Traditional Chinese Garden site plans requires domain knowledge about architecture design, landscape and landscape history intensively. And are usually achieved by speculative designs with the help of architects and landscape architects (Zhao et al., 2022). A typical workflow includes site restoration, topography restoration, landscape site planning restoration and eventually architectural floor plan restoration (Zou et al., 2021) Such approach usually takes more than a year to restore one single garden.

Digital technologies such as 3D laser scanning, photogrammetric reconstruction (Amorim Côrtes. et al, 2007) and Virtual Reality (Brown, A. et al, 2020) provides an efficient alternative for Traditional Chinese Garden restoration at scale. Deep

Generative Models using Machine Learning and Deep Neural Network have been used extensively to generate and restore Traditional Chinese Gardens. Cai et al. presents a method that generates of architectural layout with certain features from Traditional Chinese Gardens (Cai et al., 2022). Liu et al. proposed a method for conditionally generating Traditional Chinese Garden layout plans using Pix2Pix, an image-to-image generation Neural Network from line drawing to colour segmentation based on programs (Liu et al., 2021).

Although such approaches provide insights to heritage restoration and generation at scale, they rely on abundant image data. Little research has been done on the generation and restoration with scarce of image data or from natural language or text descriptions. Moreover, these approaches have not tapped into the potential of the abundance of existing written text and literature documentation about Traditional Chinese Gardens.

2.2. MULTI-MODAL MACHINE LEARNING: TEXT-TO-IMAGE GENERATION

For the past decades, researchers have been actively looking for methods to bridge design intents expressed in text (languages modality) with design outcome as images (visual modality). Mansimov was among the first to propose the use of deep neural network for text-to-image synthesis with Neural Language models (Mansimov et al., 2016). Later, Reed et al. (Reed et al. 2016) showed that GANs (Goodfellow et al., 2015) increased picture quality in place of recurrent variational auto-encoders. Vaswani et al. introduced transformers as a novel Natural Language Processing (NLP) architecture (Vaswani, A. et al. ,2017). In 2021, NVIDIA Research developed GauGAN2, which uses a deep learning model that turns a simple written phrase or sentence into a photorealistic masterpiece.

Attension-based large-scale pre-trained models such as Transformer (Vaswani, A. et al. 2017) have played an increasingly important role in image generation from text prompts. Recent neural language models such as DALL·E (Ramesh, A. et al. ,2021) and CLIP (Radford, A. et al.， 2021) showed the potential of image generation with joint embeddings. CLIP-Guided Diffusion Models generate images from text by connecting a class-conditional ImageNet diffusion model (Dhariwal et al., 2021) to a CLIP model. More recently, models like Disco Diffusion (Rafael-Patino et al., 2021) and Stable Diffusion gained increasing attention in the design field. Not only does these model generate objects with recognizable properties, they are capable of generalizing to held-out scopes (Reed et al. 2016).

2.2 TEXT-TO-IMAGE GENERATION IN BUILT ENVIRONMENT

In the domain of Computer-Aided Architectural Design (CAAD), text-to-image models are used extensively to fulfil the task of architectural space generation. Theodore Galanos (Wang, P., 2021) developed Archi-Text to generate architectural floor plans based on text descriptions. Archi-Text was trained on more than 150 thousand paired floor plans and their corresponding text descriptions.

In addition, Hang proposed an approach that compiles the logical semantic structure for the input texts and generates the 3D architectural form expressed by the language descriptions (Hang Z., 2021). Li et al. proposed a method to generate floor plan from semantic program graph description (Li et al., 2022). Lin demonstrates the ability of neural networks to integrate the language of form through written texts and has the potential to interpret the texts into sustainable architecture with a novel approach using Machine Learning-based language models (Lin, Y., 2022). More recently, Ladron proposed a multimodal sequence-to-sequence model that takes design images and their corresponding descriptions as input and outputs a probability distribution over regions of the images in which design attributes are grounded (Ladron d. G., 2022).

These novel research shows the possibility of generating architectural designs at scale from text representations. Traditional Chinese Gardens are a special sub-domain for architecture design; therefore, it is worth experimenting these new approaches for heritage restoration at large.

3. Methodology

Our novel methodology of generation and restoration consists of the following steps (Figure 1):

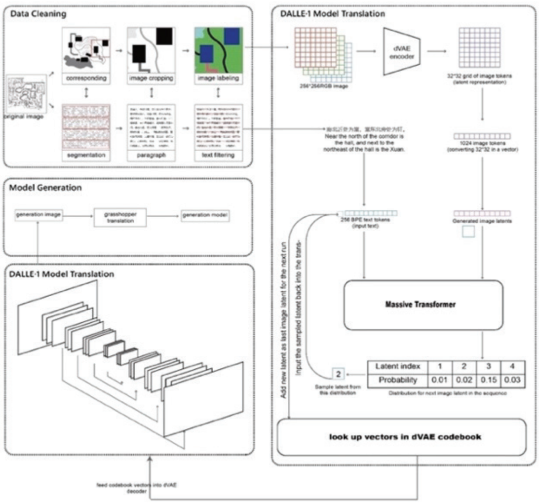


Figure 1. Diagram of Research Methodology

- Dataset Creation. We first create a novel dataset of paired text descriptions and corresponding layout plan of Traditional Chinese Gardens. The text descriptions are excerpted sets of literature about existing gardens and the images are selected from official image archives of such gardens. We canonicalize the text and images into data representation required by DALL·E.
- Text-to-Image Model Training. Then we fit text-to-image model DALL·E on our dataset. We first used the garden's description to train a discrete variational autoencoder (dVAE) to obtain the language-visual codebook, then use the dVAE for training an auto-regressive transformer for text-to-image generation.

- Testing and Evaluation. We use CLIP to evaluate the results by testing the model on test data and compare the results from all six groups to obtain the optimal data representation for the restoration task on un-held data points.
- 3D Representation. A parametric modelling pipeline in modelling software Grasshopper to procedurally generate 3D models from the plan images is developed to generate visual reference for restoration.

3.1. DATASET CREATION AND AUGUMENTATION

Data Collection from Literature. We first collected text descriptions and site plan drawings of a total of 40 gardens from book Collection of Traditional Chinese Garden from Each Dynasty, an important encyclopaedia literature and ethnographic documentation about Traditional Chinese Gardens (Yang G, 2005). The original site plan is represented as CAD line drawings of various garden elements such as water body, path, architectural elements, topography and plantation.

Image Data Scaling. We scale each garden drawing image to the same scale and crop out multiple 256×256 square RGB sub-images of out of each garden. Each sub-image represents a partial site plan of the garden. Since the total area of each garden varies, each garden drawing is disassembled into 41 ~ 96 pieces.

Drawing Element Classification and Color-coding. Chinese gardens have rich connotations and consist of different types of archetypes. Based on the reference, we extracted common architectural archetypes across different gardens and created a mutually exclusive classification systems that represents all architectural elements with a unique label/text token. Our labels are indicated in form 1-1. (Table 1) For each architecture type, also specified a unique encoding for the element, to generate a colour-segmentation variant that semantically map out the programs in the site plan.

Architecture Type(Chinese)	Architecture Type (English)	Icon	Color Code	Architecture Type(Chinese)	Architecture Type(English)	Icon	Color Code
路径	Pavement			廊	Varandal		
祠堂	Ancestor Hall			塔	Deck		
庵、斋	Retreat Space			榭	Terrace		
楼、馆、阁	Communal Space			舫	Boat		
堂、轩、厅	Hall			竹林	Bamboo Forest		
房、屋	Living Space			大片植物	Bushes		
台	Platform			假山	Rockery		
亭	Pavilion			廊	Varandal		

Table 1. Atlas of Architectural Element Classifications

Pairwise Spatial Relationship Description. As a next step, we create a set of rules to process the original text data. The text descriptions are canonicalized with a uniform, pairwise format about the spatial relationships between different garden elements. More specifically, we abbreviated and extracted important spatial information with some standard features from the original garden texts before the training. There are three types of spatial information in original text descriptions.

- **Cardinal Directions.** Cardinal direction is an essential factor in the original spatial relationship. We use both cardinal directions (East, West, South, North) as well the ordinal directions (Northeast, Southeast, Southwest, and Northwest) to represent the pair-wise cardinal direction relationship between architectural elements.

- **Distance.** To make the spatial elements more accurate, we distinguish the distance between Far and Near. The concept of distance and proximity in Chinese gardens is fundamental as it relates to the spatial perception.

We also create varied text and image representations to search for the optimal data engineering. As for the text, we made three versions of text description with different amounts of information. The 6 groups of data are shown in detail in Figure 2.

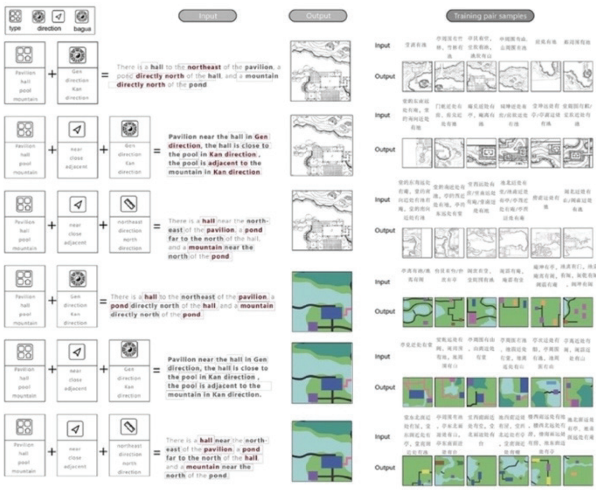


Figure 2. Dataset Creation

Data Augmentation. To augment the dataset, we performed rotation and flip to expand the original dataset eight times to optimize the training results to increase the dataset further. We developed a program that alters the synthesize text when performing the augmentation (the cardinal and ordinal directions may be changed during the data augmentation but can be generated accordingly based on the augmentation rules.) Although our data augmentation method cannot increase the data diversity fundamentally, it provides more samples when our data source is scarce and help reduce overfitting effectively (Shorten and Khoshgoftaar, 2019).

After data processing, we obtained 6 groups of datasets, each with 2284 pairs consisting of garden site plans and corresponding text descriptions.

3.2. TEXT-TO-IMAGE GENERATION WITH DALL·E

DALL·E is a CLIP-assisted generative network trained on a dataset of text-image pairs that create images from text descriptions of a broad range of natural language topics. We retrained the DALL·E model on our dataset from scratch, including both the dVAE and the autoregressive transformer.

3.2.1. dVAE

The dVAE compress the original 256×256 RGB image into 32×32 grid of image tokens, each element can be one of 8192 different values. This can compress the model and reduce the memory requirement of Transformer without sacrificing image quality.

3.2.2. Autoregressive Transformer

Up to 256 BPE-encoded text tokens were concatenate with the 32×32=1024 image tokens to train an autoregressive transformer that models the joint distribution over text and image. The distribution is modelled $\ln p_{\theta,\psi}(x,y,z) = p_\theta(x|y,z)p_\psi(y,z)$, thus yielding the lower bounds:

$$\ln p_{\theta,\psi}(x,y) \geq \mathop{\mathbb{E}}_{z\sim q_\phi(z|x)}(\ln (p_\theta(x|y,z) - \beta D_{KL}(q_\phi(y,z|x), p_\psi(y,z))$$

Where q_ϕ represents dVAE generated distribution over image tokens given the 256×256 RGB image, p_θ denotes dVAE generated distribution over 256×256 RGB image given image tokens and p_ψ is the joint distribution over text and images inferred from the transformer (Ramesh, A. et al. ,2021).

3.3. 3D REPRESENTATION

We developed a system in parametric modelling environment, Grasshopper to transform 2D images into 3D models. We first use raster-to-vector extraction tool Rooster (https://parametrichouse.com/rooster/) to extract key geometric shapes with the colour coding scheme. These shapes are then piped, arrayed or extruded into 3D shapes based on corresponding procedural rules.

4. Experiment

4.1. TRAINING DIESCRETE VAE(DVAE)

The training consists of training for dVAE and training for DALL-E. Our dVAE training was conducted In the first part of training, the pre-training of dVAE, the images are encoded through the encoder to transform the images into tokens. We created a dVAE model with codebook dimension 1024, temperature 0.9 and hidden dimension 64. We train the model for 200 epochs on our dataset in Google Colab environment with a dedicated Nividia GPU. The model converged after 4 hours of training. We fit the model on all dix datasets. As shown in Figure 3, group 4-6 (model D,E,F) with colour segmentation representation converges better then group 1-3 (model A,B,C) dataset with line drawing. (Figure 3)

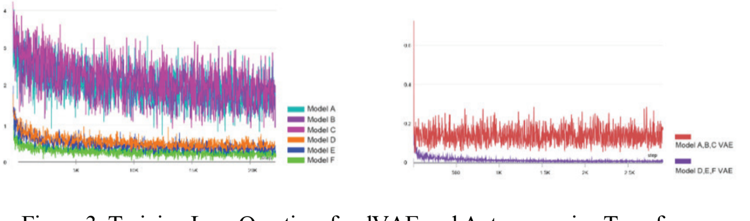


Figure 3: Training Loss Overtime for dVAE and Autoregressive Transformer

4.2. TRAINING AUTOREGRESSIVE TRANSFORMER

We train DALL·E with dVAE checkpoints obtained from Section 4.1. The epoch is set to 200, the step is set to 100, the decay rate is set to 1e-6, and the learning rate is set to 3e-4. During the training process, the primary DALLE-1 model takes our data pairs as input and generates corresponding styled images as output. We fit the model on the six datasets and as the training result shows in Figure 3, group 4-6 (model D,E,F) with colour segmentation representation converges much better then group 1-3 (model A,B,C) with line-drawing representation.

4.3. RESULT ANALYSIS

We tested the model's performance using languages on test set with descriptions that the model has not been fit on during training process (Figure 4).

Figure 4. Comparison Between Model Fitting Results on Various Dataset

We used CLIP model to rank the generated result and select the images that best match the descriptions. Through testing, we found that the images generated by the model numbered images 2-2 performed best. The model performs best when the dataset is a color-encoded image, directly describing orientation and proximity.

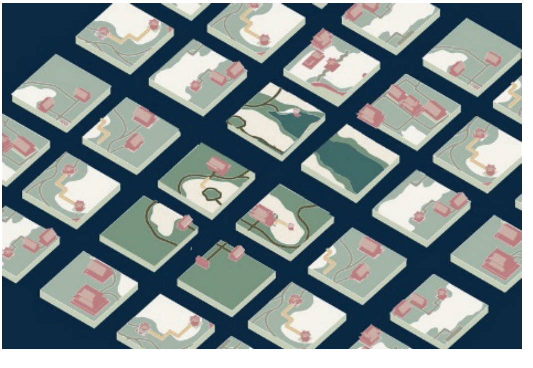


Figure 5. 3D representation

The main spatial elements in the generated images are similar to human designs. The machine learns the relative distribution of critical components based on orientation and distance, consistent with human thinking when describing a garden tour in text. Based on the generated images, we used our Grasshopper-based parametric pipeline described in section 3.3 to generate the 3D Model. (Figure 5).

We also designed a user-interface for our restoration tool through which users can type-in text descriptions for a Traditional Chinese Garden design and obtain 2D drawings as well as 3D models using our developed pipeline (Figure 6).

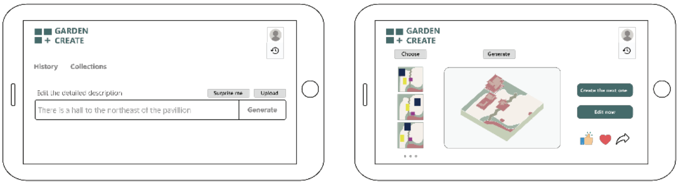


Figure 6. User Interface Design

5. Conclusion and Future work

Based on the DALL·E model, this study implements a method to reconstruct the plan layout of a Traditional Chinese Garden site plan using textual language. Our methods generate a 3D model as a product that can be used for heritage restoration. Our model also shows the potential of multi-modal machine learning in design generation. We believe with the advancement of other generation algorithms such as Text-to-Video and Text-to-3D, our approach could be applied to an even wider range of applications.

Other Societal Impact. Apart from the AEC design domain where architects can use the method to rapid prototype plan drawings from text design intent, our model extends the potential of generative AI into heritage restoration, where our method generates restoration references at scale. The prototypical 3D model from our parametric pipeline provides a solid foundation for speculative design-based restorations and can be used together with other digital restoration technologies to improve the restoration efficiency. Our method could also be generalized to other restoration tasks such as landscape heritage or even city heritage.

Limitations. The data we collected are mostly located in Pan-Yangtze River delta cities. The dataset labeling and feature engineering needs to be polished for higher generation quality. The dataset can be further expanded, with refined labels, more fine-grained definition of building types and more diverse geographical locations.

Human-Factor Consideration. Although our approach automates generation at large, designers play an important role in the entire process. Our dataset is collected, processed and curated by human designers. The bias embedded in their knowledge about Traditional Chinese Garden and architecture may have an impact on the outcome of our methodology. Nevertheless, through our experiment, we see the importance of design decisions from human how and their concurrence with machine intelligence can bring new possibilities in a future where increasing use of AI can be foreseen.

References

Amorim Côrtes, M. (2007) "Digital Photogrammetric Restitution of the Pelourinho's Façades, in Salvador - Bahia, Brazil", Proceedings of the 11th Iberoamerican Congress of Digital Graphics, pp. 375-379.

Brown, A. (2020). "Inhabiting 'prosperous Suzhou' through smart VR - interrogating an ancient artwork and documents to manifest tangible and Intangible Heritage." Available at: https://doi.org/10.26686/wgtn.12922121.

Cai, C., & Li, B. (2021). Training deep convolution network with synthetic data for architectural morphological prototype classification. Frontiers of Architectural Research, 10(2), pp.304-316. https://doi.org/10.1016/j.foar.2020.12.002

City of Suzhou (2022). Yuanlinminglu-Suzhoushiyuanlinju [Garden Directory - Suzhou City Garden Bureau]. Retrieved from: http://ylj.suzhou.gov.cn/szsylj/ylml/nav_list.shtml

Dhariwal, P., & Nichol, A. (2021). Diffusion models beat GANs on image synthesis. Advances in Neural Information Processing Systems, 34, pp.8780-8794.

Goodfellow, I.J. (2014). On distinguishability criteria for estimating generative models. arXiv preprint arXiv:1412.6515. https://doi.org/10.48550/arXiv.1412.6515

Ladron de Guevara, M. et al. (2022) "Design intents disentanglement: A multimodal approach for grounding design attributes in objects," CAADRIA proceedings

Liu, Y. et al. (2021). "Exploration on machine learning layout generation of Chinese private garden in Southern Yangtze," Proceedings of the 2021 DigitalFUTURES, pp. 35–44.

Lin, Y. (2022). "Rhetoric, writing, and ANEXACT Architecture: The experiment of natural language processing (NLP) and Computer Vision (CV) in architectural design," CAADRIA proceedings.

Lu, Y. et al. (2021). "Deep multi-task learning for Cooperative Noma: System Design and principles," IEEE Journal on Selected Areas in Communications, 39(1), pp. 61–78.

Li, A., Tian, R., Wang, and, X. and Lu, Y. (2022). Plan Generation from Program Graph. Machine Learning and the City: Applications in Architecture and Urban Design.

Mansimov, E., Parisotto, E., Ba, J.L. and Salakhutdinov, R., (2015). Generating images from captions with attention. arXiv preprint arXiv:1511.02793. https://doi.org/10.48550/arXiv.1511.02793

Park, T., Liu, M.Y., Wang, T.C. and Zhu, J.Y. (2019). GauGAN: semantic image synthesis with spatially adaptive normalization. In ACM SIGGRAPH 2019 Real-Time Live! (pp. 1-1). https:// doi.org/10.1145/3306305.3332370

Radford, A., Kim, J.W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J. and Krueger, G., (2021). Learning transferable visual models from natural language supervision. In International Conference on Machine Learning (pp. 8748-8763). PMLR.

Ramesh, A., Pavlov, M., Goh, G., Gray, S., Voss, C., Radford, A., Chen, M. and Sutskever, I., (2021). Zero-shot text-to-image generation. In International Conference on Machine Learning (pp. 8821-8831). PMLR.

Rafael-Patino, J., Girard, G., Truffet, R., Pizzolato, M., Caruyer, E. and Thiran, J.P., (2021). The diffusion-simulated connectivity (DiSCo) dataset. Data in Brief, 38, p.107429. https://doi.org/10.1016/j.dib.2021.107429

Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. Journal of big data, 6(1), 1-48. https://doi.org/10.1186/s40537-019-0197-0

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł. and Polosukhin, I. (2017). Attention is all you need. Advances in neural information processing systems, 30.

Wang, P., Page, C. T., & Save, R. (2021). DALL-E in Pytorch.

Yang, G. (2005). Zhongguo Li Dai yuan Lin Tu Wen Jing Xuan. [Collection of Traditional Chinese Garden from Each Dynasty] Shanghai: Tongji University Press.

INTEGRATION OF EEG AND DEEP LEARNING ON DESIGN DECISION-MAKING

A Data-Driven Study of Perception in Immersive Virtual Architectural Environments

MINGMING ZHAO[1], CAO DING[2] and TATJANA CROSSLEY[3]
[1,3]*School of Architecture, The Chinese University of Hong Kong*
[2] *Department of Electronic and Information Engineering, The Hong Kong Polytechnic University*
[1] *mingmingzhao@link.cuhk.edu.hk, 0000-0001-9940-490X*
[2] *caosnow.ding@connect.polyu.hk, 0000-0002-7532-5809*
[3] *tatjana.crossley@gmail.com, 0000-0002-7902-0674*

Abstract. Immersive virtual reality(IVR) as an emerging architectural design tool is utilized by many architecture firms to assist in better design decision-making. It allows users to immersively experience the simulated architectural environment prior to real construction. However, compared to conventional computational design tools, IVR faces more challenges in assessing the perception of designed simulations and visualizations. This paper attempts to examine the possibilities for incorporating human biological data and deep learning technology into the process of immersive visualization in architectural design. It aims to objectively understand human perception in an immersive virtual architectural environment, and ultimately assist in design decision-making and human-centered architectural design. The study proposes a novel and multidisciplinary use of techniques derived from psychology, computer science, and architecture disciplines to explore how biological data might be understood architecturally and vice versa. It also provides an opportunity to explore ways of using IVR-based computational design in the new metaverse era. The experiment results illustrate that there is a significant correlation between environmental experience and brain activation. It indicates the integration of EEG and deep learning is helpful to perform as complementary tools for better understanding human perception in immersive virtual architectural environments.

Keywords. Architectural Design Decision-Making, Eye Tracking, Electroencephalogram(EEG), Convolutional Neural Networks(CNN), Virtual Reality(VR).

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 89-98.

1. Introduction

1.1. RESEARCH BACKGROUND

With the impact of the pandemic and the rise of virtual reality techniques, the whole world is witnessing a social shift to a metaverse era and an unprecedented focus on virtual spaces. After decades of experimentation in gaming, tourism and education, immersive virtual reality (IVR) has shown considerable promise as a professional tool for conducting experimental research and exploring new possibilities in the field of architecture. Immersive design simulation with head-mounted virtual reality (VR) has been utilized by many architecture firms to assist in better design decision-making. Active spatial exploration enabled by IVR can enhance users' spatial perception rather than passive watching on a two-dimensional screen offered by traditional VR (Chalmers, 2022). It allows architects and clients to immersively experience the simulated architectural environment prior to real construction (Noghabaei et al., 2020). Compared to conventional computational design tools, IVR has been validated as an effective visualization tool for design review resulting in a better understanding of the proposed design (Liu et al., 2020). In this sense, VR technology is revolutionizing the conception of “human-centered” design by prioritizing and maximizing realistic simulation of spatial perception, which is essential to phenomenological architecture (Rasmussen, 1964). Nevertheless, computational simulations have been always questioned to be “an accurate representation of the real-world aspects that concern us” (Koutamanis, 2014). While IVR offers a more immersive and “realistic” experience in design simulation and visualization, it is still underexplored with significant challenges in assessing spatial perception. IVR users may experience sensations such as neophobia and motion sickness due to lack of experience, as well as poor design efficiency due to the fact that users need to process more visual information than in the desktop environment (Umair et al., 2022).

The primary purpose of this paper is to examine the possibility of cooperating EEG and CNN deep learning to support architects in better understand the occupants’ perception in architectural design. Through human biometric data generated by EEG with brain activation and the VR headset with eye tracking, it aims to explore a multidisciplinary approach that can help objectively understand human perception in an immersive virtual architectural environment, and ultimately assist in human-centered design decision-making. This study involves two research hypotheses listed as follows: H1. EEG data of arousal and valence value show consistency with subjective questionnaires from participants, regardless of an architectural design background. H2. The integration of EEG and deep learning facilitates a better understanding of the participants’ environmental perception of architecture for architects. It, therefore, contributes to architectural design decision-making, particularly when the decision maker has limited architectural design experience.

1.2. LITERATURE REVIEW

Neuro-architecture is an emerging field based on the wide exploration of neuroscience, environmental psychology, and architecture in history. It explores how the human brain responds when interacting with built environments by utilizing neuroimaging technologies. (Karakas & Yildiz, 2020). According to Gregory (1998), the brain is the highly active engine of understanding, and it constructs perceptions from hardly adequate information from the senses. When the olfactory, auditory, visual, gustatory and tactile nerves of the brain are stimulated, their stimulus-response signals can be expressed through brain waves, therefore revealing the psychological correlation between perception and environmental stimuli. As one of the neuroimaging tools to record the electrical activity in the brain, electroencephalography(EEG) has been utilized to measure central nervous system responses for human emotion study in numerous research fields. Conventional emotion recognition is primarily focused on the study of facial features, body movements, and speech, all of which can be easily disguised and fail to accurately reflect true emotions. The EEG-based affection recognition method provides a quantitative assessment of human emotion through continuous measurement of processing between a stimulus and response, and EEG signals can reflect the neurophysiological activities related to emotion processing in the brain. This makes it an efficient tool for studying environmental perception, which can well compensate for the shortcomings of traditional research methods (Freeman & Quiroga, 2013). In the field of architecture, EEG has been used to gain new insight into how people perceive built environments since the 1980s (Ulrich, 1981). EEG technology has commonly been implemented to quantitatively measure an occupant's physiological reaction to changes in environmental factors, including sleep comfort and thermal temperatures (Pan et al., 2012), stress and green space (Qin et al., 2014; Choi et al., 2015), wayfinding behaviors to urban landmarks and space planning (Erkan, 2018), and cognition and performance in a given architectural environment for specific tasks (Liu et al., 2020). These studies confirmed the feasibility of applying EEG in architecture research for better understanding human perceptions in architectural environments, however, most of the studies were conducted through desktop-based computer simulations with a low level of immersion.

Taking advantage of the IVR's potentialities and the increased accuracy of commercial portable physiological devices, the integration of VR and EEG has been evidenced in many studies to be an effective way to collect objective data for exploring human perceptual feedback based on the change of architectural elements (Kim et al., 2021). However, the environmental simulation in their research was based on the 360° image in VR, which means passive spatial observation with limited ability to elicit real emotions (Baños et al., 2004). Overall, the previous research studies present the following limitations: (1) Few studies apply a psychological data-driven approach to analyze human emotions in immersive virtual environments(IVEs) for architectural design decision-making; (2) The combination of EEG and VR for analyzing spatial perception is scarcely concerned with active stimuli that involve behavioral manipulation; and, (3) There is limited research in architectural design decision-making that tries to recognize the user's emotions in an IVE through EEG signals and convolutional neural networks(CNN) algorithms of attentive object detection.

2. Methods

The study employs a data-driven approach by using a VR headset (HTC Vive Pro Eye) for generating immersive design simulations and eye-tracking coordinates with the SRanipal plugin in Unreal Engine, a portable EEG device (Emotiv Insight) for collecting biological data of brain activities, and CNN deep learning for object detection (Fig. 1). The atrium of the School of Architecture building at the Chinese University of Hong Kong(CUHK) was selected as a case study and designed with three different scenarios (Fig. 2). The virtual environment was developed using Epic Games' Unreal Engine 4 with blueprint scripting.

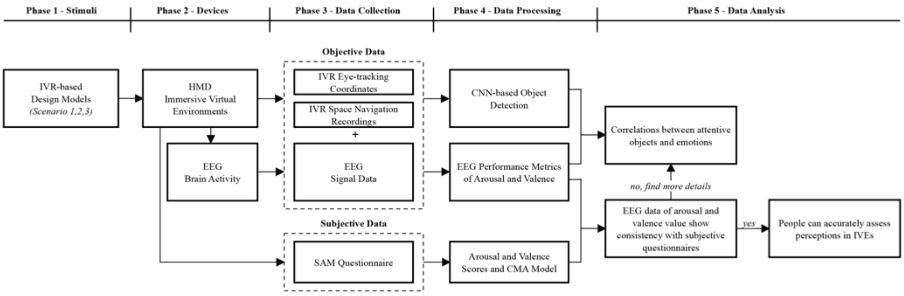


Figure 1. Research framework

Figure 2. Visual Stimuli (VR Scenario 1, 2, 3 from left to right)

2.1. Experiment Procedure

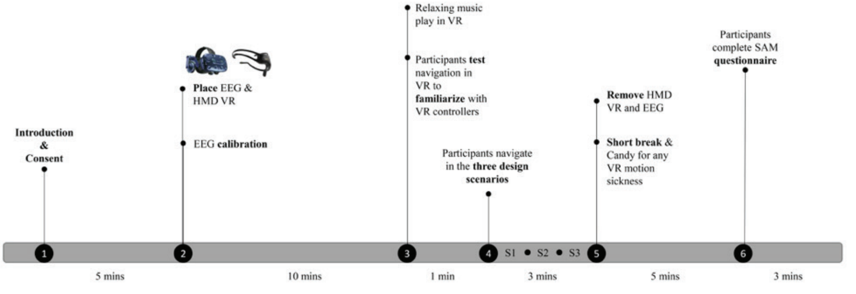
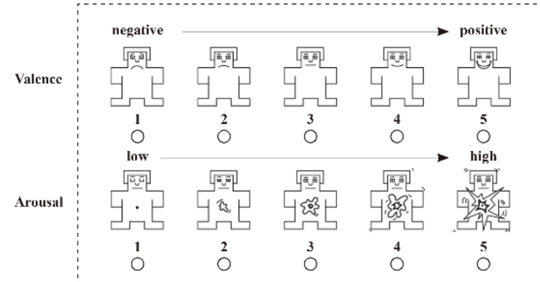


Figure 3. Experiment procedure

As shown in Figure 3, six steps comprised the experiment. Participants were sequentially invited to a quiet laboratory at similar times for the experiments that lasted approximately 30 minutes each. Firstly, they were introduced to how to navigate in the virtual environment using the VR controllers, along with the overall experimental procedure and instructions, including staying seated and refraining from speaking during the entire experiment unless one wished to quit midway. A consent form was signed by a participant after the introduction. In the next step, the participant was then placed on the EEG which was tested to ensure each channel was stably connected.

After placing the VR headset outside the EEG, it was tuned again until each channel had a signal quality of greater than 95%. To restore the participants' mood, a half-minute of relaxing music was played first, following which the interface automatically stepped into the first scene. It was requested that participants follow the cue track on the floor first to ensure a full view of the objects in the space and then explore freely. After the third scenario, the interface automatically displayed the phrase "thanks for your participation" as a reminder of the end of the experiment. Completed recordings of the VR video, eye-tracking and EEG were safely stored. The participants were then assisted in removing the headset and EEG. The sixth step involved completing a questionnaire (Fig. 4) to assess participants' subjective perceptions of each scenario. The questionnaire was developed based on the Bradley and Lang's (1994) Self-



assessment Manikin(SAM) with 5 Likert scales. The SAM scale is a pictorial assessment tool that assesses both valence (the degree to which an emotion is perceived as positive or negative) and arousal (how strongly the emotion is felt) with individuals' affective reactions to stimuli.

Figure 4. Example of the SAM scale questionnaire (Bradley and Lang, 1994)

2.2. PARTICIPANTS

Twelve healthy volunteers (T=12, 6 males and 6 females; mean: 27.82, STD: 3.40) participated in the experiments. To protect personal privacy, they were indicated with a number from 1 to 12. Six students (Group a) from the School of Architecture and six (Group b) from other schools at CUHK participated in the study to gain insight into how people with and without architecture backgrounds react to architectural IVEs.

2.3. DATA PROCESSING

This paper presents three types of data collection and processing, including VR recordings, EEG and self-assessment questionnaires. The EEG data from 12 participants was analyzed using performance metrics algorithms in EmotivPRO. In this study, excitement and interest were the two performance metrics that were primarily analyzed, which can be highly representative of arousal and valence. On the one hand, EEG data (128 Hz sampling rate) were used as objective resources to compare with the subjective SAM questionnaires (Step 5) for exploring their consistency (H1). The dimensional model of the "Circumplex Model of Affects(CMA)" proposed by Russell and Mehrabian was employed to transform the subjective data collected. Using a cartesian system of axes, it classifies emotions according to their valence (the degree to which an emotion is perceived as positive or negative) and arousal (how strongly the emotion is felt) in the two dimensions. It is now one of the most important models for measuring emotion in psychology. On the other hand, EEG signals were integrated with VR recordings and eye-tracking coordinates by using CNN-based object detection technology to evaluate if they could facilitate a better understanding of the participants' environmental perception (H2). We utilized YOLO v5 (Ultralytics, 2020) to discriminate the objects seen by the participants. The backbone structure of YOLO v5 implemented is shown in Figure 5, in which the pre-trained YOLO v5x model was used for real-time object recognition of VR recordings, with its light weight, fast recognition speed, and high accuracy. In this model, nine layers deep and multiple neural network blocks were utilized for object recognition training and testing, such as Focus, Convolution(Conv), BottleneckCSP, and SPP.

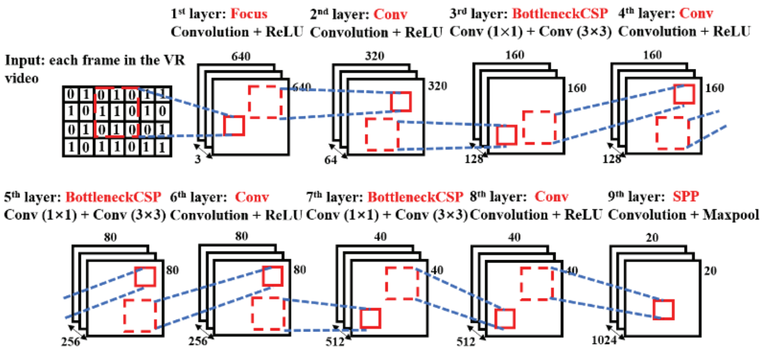


Figure 5. The backbone structure of YOLO v5 model

For each VR video, we compress the frame rate at 20, and input it frame by frame to the CNN network. We set a confidence threshold of 0.5 and an eye fixation level of 0.5 for particular recognition, which indicates that an object will only be shown when the accuracy of the recognition and the eye fixation level are both higher than 50%. From the eye movement coordinates dataset, we obtained the occurrence number of each recognized object position. The higher the eye fixation level is, the higher the number of occurrences in a certain position. The derivation of eye fixation level is as follows:

$$P(x_n) = \frac{m(x_n)}{\sum_{n=1}^{N} m(x_n)}$$

Where x_n is the n-th detected object in the x-th frame, where $n = 1, 2, \ldots, N$, $x = 1, 2, \ldots, X$. N is the total number of detected objects in the x-th frame, $m(x_n)$ denotes the number of detection of the n-th detected object from the eye movement coordinates, $\sum_{n=1}^{N} m(x_n)$ represents the total number of all detected objects in the x-th frame.

3. Data Interpretation and Discussion

3.1. AROUSAL AND VALENCE

In the CMA dimensional model (Fig. 6), we define emotional arousal in the dimension 1 to 5, where 4 to 5 is specified as a high arousal level, 3 as a medium arousal level, and 1 to 2 as a low arousal level. Four regions were divided, including: I - high arousal and positive valence (excited, delighted, happy); II – high arousal and negative valence (tense, stressed, upset); III – low arousal and negative valence (sad, depressed, bored); IV – low arousal and positive valence (relaxed, calm, contented).

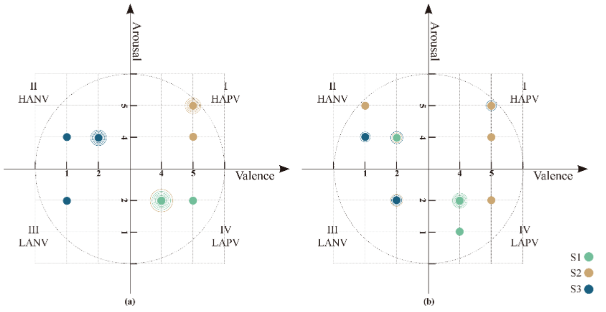


Figure 6. Arousal and valence diagrams for group (a) and group (b)

The self-assessment scores from questionnaires were averaged using the mean and standard deviation (Table 1). It indicates positive valence and high arousal when the mean is above 3. The mean of valence and arousal in the subjective assessment of the three scenarios produced similar region-based results from both architectural and non-architectural students: S1 – LAPV; S2 – HAPV; S3 – HANV. However, it can be observed that there is greater variance in the non-architectural subjects than the architectural subjects from the CMA, as well as the standard deviation in Table 1.

	Valence				**Arousal**			
	Arch		**Non-Arch**		**Arch**		**Non-Arch**	
S1	4.17±0.41	P	3.67±0.82	P	2.00±0	L	2.17±0.98	L
S2	4.83±0.41	P	3.33±1.86	P	4.17±1.33	H	3.67±1.37	H
S3	1.67±0.52	N	2.33±1.51	N	3.67±0.82	H	3.50±1.22	H

Table 1. Arousal and valence resulted from SAM questionnaires with 12 participants.

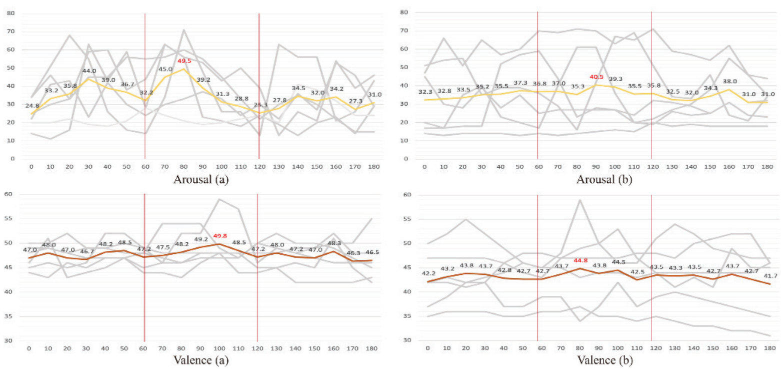


Figure 7. Performance metrics of group (a) and group (b)

The EEG performance metrics data with 12 participants is displayed on a scaled axis y from 0 to 80, and an axis x from 0 to 180 seconds (Fig. 7). Similar to the high standard deviation from the questionnaire, the data of participants in Group b showed more variance. From the mean of the performance metrics of the 12 participants, we can see that the peaks of the average arousal and valence both appear in the second scenario. The first scenario did not demonstrate a significant peak in arousal or enhanced valence, whereas the third scene displayed relatively pronounced arousal as well as a decreasing trend in valence. Generally, the results were consistent with the questionnaire for most participants, with the exception of 3 participants (1 in Group a, 2 in Group b) whose arousal and valence contradicted the questionnaire. It suggests that the introduction of objective data is necessary in the evaluation of perceptions in IVEs.

3.2. CNN-BASED ATTENTIVE OBJECT DETECTION

Six index items (plants, couch, table, ottoman, chair, and desk) were detected with YOLO v5. The number of attentive objects detected was calculated based on 10-second intervals, as well as the arousal and valence values from the EEG (Fig. 8 and 9).

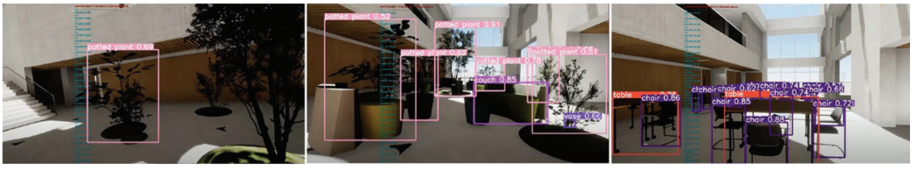


Figure 8. Example of CNN-based attentive object detection in VR Scenario 1, 2, 3 (left to right)

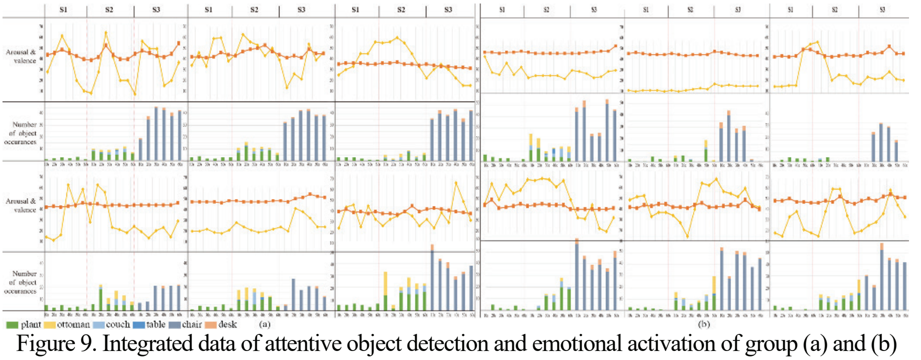


Figure 9. Integrated data of attentive object detection and emotional activation of group (a) and (b)

The results reveal that participants' arousal is highly influenced by the number of attentive objects. Some participants had lower arousal and valence levels than others. The findings of analyzing the VR recordings indicate that this could be the result of some anomalies, such as unfamiliarity with the manipulation system and the act of stepping out of the main building space and into a dark corner in the simulation. To verify the relationship between the attentive objects and participants' emotions, these values were further analyzed with the Pearson correlation algorithm in SPSS (Table 2). The correlation coefficients, which are bolded, vary between +1 and -1, where +1 is a perfect positive correlation and -1 is a perfect negative correlation. It was discovered that plants had a significant positive correlation with arousal, implying that the presence of plants elicited arousal in the participants. Additionally, valence showed a negative correlation with plants, couches, tables, and chairs, whereby the value of valence decreased with the presence of those items. Conversely, a pronounced positive correlation was observed between valence and the ottoman.

		Plant (S1&S2)	Couch (S2)	Table (S2)	Ottoman (S2)	Chair (S3)	Desk (S3)
Arousal	PC	.315**	.026	-.082	-.051	-.002	.172
	Sig. (2-tailed)	0.010	.861	.689	.712	.991	.075
Valence	PC	-.337**	-.459**	-.455*	.427**	-.361**	.212
	Sig. (2-tailed)	0.001	0.001	0.007	0.019	0.001	0.088

**Correlation is significant at the 0.01 level (2-tailed).

*Correlation is significant at the 0.05 level (2-tailed).

Table 2. Pearson correlation analysis in SPSS

4. Conclusion

To conclude, the findings suggest that incorporating biometric data into the design decision-making process can be beneficial. It was found that most of the EEG data and the self-assessments were generally consistent, regardless of whether the participants had an architecture-related background. There is, however, a greater variance in the results of participants from other schools than those from architecture. To further interpret the few inconsistencies between EEG data and self-assessments as well as the fluctuations of EEG signals, a practical workflow was established to investigate the correlation between brain activation and the frequency of attentive objects in virtual architectural environments. While space navigation within IVEs has been demonstrated to be beneficial for studying spatial perception, it was considered too dynamic to be used as an environmental stimulus for general analysis, which is the challenge that this study attempts to address. This research developed a multidisciplinary approach from psychology, computer science, and architecture disciplines to quantify design iterations and facilitate human-centered architectural design decision-making. Furthermore, it provides profound insights for other IVR-based design-related research, and potential applications of IVR for design in the new metaverse era.

The results of this preliminary study have identified some challenges and limitations. The size and diversity of the sample were limited. Due to the lack of familiarity with virtual navigation by using controllers, participants frequently wandered into non-experimental areas, despite the brief training session conducted prior to the experiment. As part of future research, we intend to increase the sample size and diversity of the samples and explore the applicability of the approach in a wider range of situations. A comprehensive training session and navigation test will be conducted before the experiments, and the boundaries of the experimental virtual environments will be strictly restricted. Additionally, it is intended to simulate more realistic spatial experiences by integrating auditory and tactile sensations to enhance sensory perception in IVEs, which will also provide opportunities for exploring diverse ways of understanding biological data architecturally and vice versa.

References

Baños, R. M., Botella, C., Alcañiz, M., Liaño, V., Guerrero, B., & Rey, B. (2004). Immersion and Emotion: Their Impact on the Sense of Presence. *CyberPsychology & Behavior*, *7*(6), 734–741. https://doi.org/10.1089/cpb.2004.7.734

Chalmers, D. J. (2022). *Reality+: Virtual worlds and the problems of philosophy*.

Choi, Y., Kim, M., & Chun, C. (2015). Measurement of occupants' stress based on electroencephalograms (EEG) in twelve combined environments. *Building and Environment*, *88*, 65–72. https://doi.org/10.1016/j.buildenv.2014.10.003

Erkan, İ. (2018). Examining wayfinding behaviours in architectural spaces using brain imaging with electroencephalography (EEG). *Architectural Science Review*, *61*(6), 410–428. https://doi.org/10.1080/00038628.2018.1523129

Freeman, W. J., & Quiroga, R. Q. (2013). Electroencephalography. In W. J. Freeman & R. Q. Quiroga, *Imaging Brain Function With EEG* (pp. 1–19). Springer New York. https://doi.org/10.1007/978-1-4614-4984-3_1

Gregory, R. L. (1998). *Eye and brain: The psychology of seeing* (5th ed). Oxford University Press.

Karakas, T., & Yildiz, D. (2020). Exploring the influence of the built environment on human experience through a neuroscience approach: A systematic review. *Frontiers of Architectural Research*, *9*(1), 236–247. https://doi.org/10.1016/j.foar.2019.10.005

Kim, S., Park, H., & Choo, S. (2021). Effects of Changes to Architectural Elements on Human Relaxation-Arousal Responses: Based on VR and EEG. *International Journal of Environmental Research and Public Health*, *18*(8), 4305. https://doi.org/10.3390/ijerph18084305

Koutamanis, A. (2014). *Computer-Mediated Briefing for Architects: Part of the Advances in Civil and Industrial Engineering Series*. IGI Global. https://doi.org/10.4018/978-1-4666-4647-6

Liu, Y., Castronovo, F., Messner, J., & Leicht, R. (2020). Evaluating the Impact of Virtual Reality on Design Review Meetings. *Journal of Computing in Civil Engineering*, *34*(1), 04019045. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000856

Noghabaei, M., Heydarian, A., Balali, V., & Han, K. (2020). Trend Analysis on Adoption of Virtual and Augmented Reality in the Architecture, Engineering, and Construction Industry. *Data*, *5*(1), 26. https://doi.org/10.3390/data5010026

Pan, L., Lian, Z., & Lan, L. (2012). Investigation of Gender Differences in Sleeping Comfort at Different Environmental Temperatures. *Indoor and Built Environment*, *21*(6), 811–820. https://doi.org/10.1177/1420326X11425967

Qin, J., Sun, C., Zhou, X., Leng, H., & Lian, Z. (2014). The effect of indoor plants on human comfort. *Indoor and Built Environment*, *23*(5), 709–723. https://doi.org/10.1177/1420326X13481372

Rasmussen, S. E. (1964). *Experiencing architecture*. MIT Press.

Rounds, J. D., Cruz-Garza, J. G., & Kalantari, S. (2020). Using Posterior EEG Theta Band to Assess the Effects of Architectural Designs on Landmark Recognition in an Urban Setting. *Frontiers in Human Neuroscience*, *14*, 584385. https://doi.org/10.3389/fnhum.2020.584385

Ulrich, R. S. (1981). Natural Versus Urban Scenes: Some Psychophysiological Effects. *Environment and Behavior*, *13*(5), 523–556. https://doi.org/10.1177/0013916581135001

Ultralytics. (2020). *YOLOv5 2020*. https://github.com/ultralytics/yolov5

Umair, M., Sharafat, A., Lee, D. E., & Seo, J. (2022). Impact of Virtual Reality-Based Design Review System on User's Performance and Cognitive Behavior for Building Design Review Tasks. *Applied Sciences*, *12*(14), 7249. https://doi.org/10.3390/app12147249

INCORPORATING PHYSICAL EXPERIMENTATION INTO CREATIVE DL-DRIVEN DESIGN SPACE EXPLORATION

SHERMEEN YOUSIF[1], YAGMUR AKYUZ[2], LUISA GIFFONI[3] and MATTHEW CRAVEN[4]
[1,2,3,4]Florida Atlantic University.
[1]syousif@fau.edu, ORCID 0000-0003-4214-1023
[2]yakyuz.design@gmail.com, ORCID 0000-0002-8817-0412
[3]luisa@fabricka.design, ORCID 0000-0001-7345-4669
[4]craven.design2016@gmail.com, ORCID 0000-0002-8034-8423

Abstract. In the context of ongoing research on incorporating deep learning (DL) strategies in architecture, this paper proposes a proof of concept, for developing a viable DL-driven design workflow with multiple connected DL models that enable various levels of agency. The approach allows design intentions to manifest systematically throughout the process, through identifying the ways of dataset curation, DL models' selection and connection. Importantly, in parallel to the interconnected DL models, a series of physical experiments were conducted for dataset augmentation and evaluation, and to inform the overall process. The formulated system involved protocols where multiple DL models are employed and interconnected to address specific architectural systems and design tasks. Applying this prototype, a test-case experiment was carried out with a parallel logic of the two processes: (1) a physical experiment (material research) and (2) the DL-driven process (a combination of multiple neural networks), incorporated into the design workflow. The physical experiment was directed at learning from fungal natural systems (mycelium) to understand growth behavior and its physical qualities, which influenced the DL testing and evaluation.

Keywords. Full Paper, Deep Learning, DL-Driven Design Workflow, Connected DL Models, Encoding Design Agency.

1. Introduction

Within artificial intelligence (AI) methods, rule-based expert systems were introduced (developed in the 1970s and proliferated in the 1980s and 1990s) as logic-driven systems, designed to solve complex problems based on reasoning and ‘if-then’ rules. The structure of the expert system is often composed of three parts, a production rule, database and control strategy (Tan, 2017). Characteristics of such systems include solving problems with stored and pre-programmed solutions, within a confined and predetermined space of solutions, built upon a pre-defined knowledge (Gupta and Nagpal, 2020). While expert systems require data and rules to yield answers (inputs +

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 99-108.

rules > results), and rely on hardcoded solutions, self-learning systems, (introduced more recently around 2006), learn solutions from first principles, directly from raw data where in unsupervised, the main input is data (inputs > rules), and in supervised learning systems it becomes the data and answers/results (inputs + results > rules) (Hassabis 2018). Deep learning (DL) networks, as examples of learning systems, are inspired by the structure and function of the human brain and the way in which humans acquire specific types of information. DL does not acquire knowledge by hard-coded answers, but rather through observation (Hassabis, 2018). Despite their promising potential for informing new design processes, issues arise at the intersection of DL models and architectural design processes. One issue this research addresses is how to employ multiple DL models to tackle multiple design aspects and phases within the evolving DL-driven design process? and how to curate the overall DL-driven process for achieving creative exploration? (Bolojan et al., 2022). Another issue addressed here is how certain levels of agency can be identified and made explicit in interacting with DL systems, especially when design intentions change within the process of design exploration. To address the first issue, this work aims at developing a new DL-driven design framework where multiple connected models are employed, tackling multiple architectural systems. This approach was informed by systems thinking in architecture, adopting Christopher Alexander's "systems generating systems", in a process where the parts within the system and their interactions lead to holistic system behavior (Alexander, 1968). We focus here on how a creative process unfolds, depending on the connection of the DL models (nodes) and their interaction. The objective is to formulate a DL-driven design system that allows for an expanded design space and creative exploration. The significance of this "hybrid" process is to leverage human intuition and machine capability in a carefully structured process. In addressing the issue of design agency, the investigation has led to the identification of multiple levels of agency, allowing design intentions to manifest systematically throughout the process. Overall, design agency was targeted in encoding design intentions which require identifying ways of dataset curation, DL models' selection, and connection, as well as post-rationalizing of the generated data. Importantly, in parallel to the interconnected DL models, the workflow was leveraged by a series of physical experiments for dataset augmentation and evaluation and incorporating intentions as curated datasets to inform the process.

To demonstrate the proposed method, a system prototype was formulated to identify methods of connecting multiple DL models and rules to guide the workflow structure. Applying the prototype, a test-case experiment was carried out with a parallel logic of the two experiments: (1) a physical experiment (material research) and (2) the DL-driven process (a combination of multiple neural networks), both integrated into the design workflow. The physical experiment was directed at learning from fungal natural systems (mycelium) to understand the growth behavior and the physical qualities of this living material. Emphasis was placed on the feedback from the physical experiment to guide and inform the DL experiments and enact human agency in supervising the process, as described in the methods section. Mycelium was selected as an intelligent growth behavior, a "neighbor-sensing system," and was employed in two phases: informing DL experimentation for developing the DL-driven system, and as a possible building material for the design proposition in the test-case application.

2. Background

Artificial intelligence (AI)-driven processes have been heavily researched in architecture in the past six years (Leach, 2021). One of the earliest architects to consider (AI) was Makoto Watanabe (2004) who established an "inductive design approach". The goal of his method is to generate unpredicted 'magic-like' aesthetics without a preconceived notion of 'what good is'. It involves an inter-exchange between the designer and the machine, where the designer begins with sketches, the computer infers and offers further ideas, followed by review until the objective is satisfied. In interactive media art, Refik Anadol's process represents a prototype for controlling the data curation to produce a creative artistic expression (Forbes, 2020). Chaillou's (2019) work represents early research of AI and architecture. His method, ArchiGAN, is a Pix2Pix-based generative adversarial network (GAN) model that employs DL models at multiple scales to generate floor plans automatically. Bolojan and Vermisso's work uses CycleGAN to perform a heuristic search, cross-pollinating domains of Sagrada Familia with forests, hinting at ways that the network can learn structures and morphology (2020). In a different approach, Immanuel Koh uses discrete sampling to produce 3-dimensional representations for AI learning, in a hybrid workflow of Computational + AI steps (Koh, 2022). An important precedent related to our proposed design workflow has been established by the Architecture Machine Lab (Negroponte, 1973), which interrogates human-machine communication. According to Negroponte, the design task would be developed and completed in the designer's own way, thus, a constant flow of ideas could be channeled directly from the designer to the machine and back. Importantly, a designer may have a limited grasp of the design problem, necessitating machine tolerance and compatibility with their search for consistency among criteria, as well as between intent and purpose. The DeepHimmelBlau project represents one of the most-related works to this research (2022) that amplifies an "open process" of design where interconnected AI models are employed to address different "nodes" in the design process. On another level, Huang et al. (2021) offer an experimental search for human and artificial intelligence symbiosis where both can inform each other to produce new designs that are culturally and architecturally meaningful. Speculating on the role of AI in architecture, Neil Leach (2021) calls this fusion "architectural intelligence", predicting it to be a dominant approach. Overall, AI-driven design is still new and experimental without leveraging AI's full capabilities in a system-driven approach. This paper fills a gap in research, by examining a possible structure of an interconnected DL-driven workflow, addressing its potential to address multiple design tasks during multiple phases.

3. Research Methods

The scope of our research is to investigate the application of DL strategies through a holistic approach to design processes, where a variety of interconnected DL models and strategies are implemented at multiple architectural design intention levels. We propose a framework made up of a series of complementary DL models to look at the possibility of logical continuity in DL-driven workflows for a new open-ended process. The methods discussed herewith consist of an extensive literature review, experimentation, framework development, prototyping the proposed workflow, and testing and evaluation (test-case application). The test-case workflow (Figure 1) shows how it was composed and connected, as described in the following sections.

3.1. TEST CASE APPLICATION

Developing the design workflow, our investigation led to a framework consisting of: (1) physical experiments: material investigation; (2) artificial experiments (DL models) that decode complex, otherwise undetectable, properties of natural systems. In this application, the brief was to design a speculative bio-centric architecture, that can leverage a new symbiotic relationship with natural ecosystems. The intended outcome was a speculative design prototype that utilizes mycelium as a building material. DL neural networks were used in parallel to the experimentation with this material. Additionally, the project brief involved grounding this experimental approach in a city on the verge of a climate crisis. A site within Ancienna Medina was selected, representing the oldest quarter in Casablanca, Morocco, and classified as "infamous bidonville", a "shantytown" (Beier, 2019). The site selection was rationalized by its challenges as an informal settlement and its potential to employ mycelium as an affordable low-tech material for the proposed structure. The first goal of employing mycelium in this research was to understand its network system and growth behavior for DL training. The second objective was at the design conceptualization phase; the potential use of mycelium as an emergent structural material. Physical experimentation in growing mycelium was aimed at a small scale was based on reviewing material research on growing mycelium as a building material.

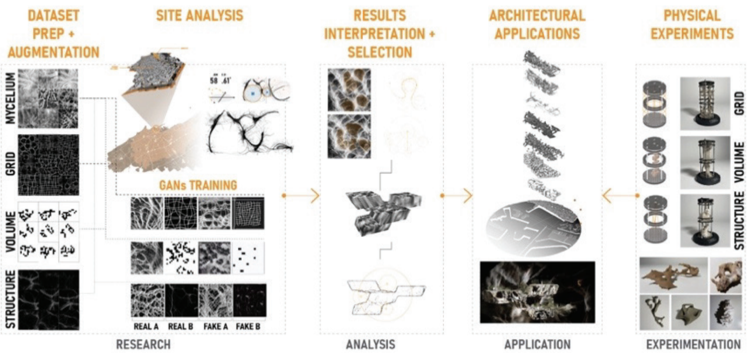


Figure 1. The proposed design workflow applied to a test-case with DL experiments conducted in parallel to physical experiments, informing each other.

3.1.1. Physical Experiments: Material Investigation

Mycelium is formed in the early stages of mushroom growth; from the spores released by mushrooms, when in contact with organic matter, they grow cell by cell to create strands linking with others until it becomes a mass of branched, tubular filaments (hyphae), known as the fungal network. Researching this material was rationalized by its biodegradable, recyclable, low-tech, inclusive, and low-cost properties. To grow mycelium for material use, various types of organic matter can be used as substrate (food), based on desired performance properties. In recent years, mycelium has been heavily studied in the realm of design and has become an alternative building material. Processed mycelium can serve as a building construction material that can repel water, resist fire, and insulate when fully grown. Block Research Group has investigated the geometric properties of a modular design that resulted in a load-bearing mycelium structure (2017). Blast Studio has found methods for converting urban waste into substrate paste, which can be 3D printed into host objects for mycelium to grow on, producing various scaled objects (2020). Most of the work on mycelium so far has been focused on baking (killing) the material after it grows to a satisfactory stage. However, other studies have been conducted to examine mycelium as a forever alive material. For us, to understand the growth behavior and the physical qualities of this living material, it was necessary to interact with it directly. Through a series of experiments (Figure 2) we looked at how mycelium grows in strategically designed systems (spatial grid, structure, and volume) when left in a hospitable environment. By proposing a dynamic structure, as opposed to a static one, the project structure can be resilient and adaptive to the expected changes in the informal settlement and its environment.

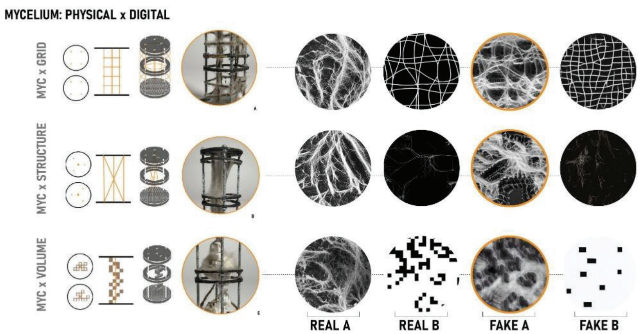


Figure 2. The physical experiments of mycelium growth were carried out in parallel to the DL experiments, addressing grid-based, structural, and volumetric explorations.

3.1.2. Artificial Experiments (DL Models)

The artificial experiments (DL model training and connecting) were formulated based on assumptions guided by the mycelium research. However, as these physical experiments progressed, some assumptions on the material's behavior were eliminated, which led to the curation of new datasets and evaluation of the outcome of the DL training processes. In a systematic procedure, we formulated a sequential connection strategy of training multiple deep neural networks with mycelium datasets to encode 'mycolic' activity into 2D and later 3D architectural explorations. Each neural network had a design task, and each node (DL model) instructed the next.

Phase 1: Natural and Generative System Explorations

To incorporate the physical experiment results into the artificial experiments, we curated datasets that reflect the structure of mycelium at the microscopic scale to decode complex network intelligence (Figure 3). Initially, three main context-specific parameters informed the following architectural systems: an organizational system, a spatial system, and an environmentally responsive envelope system. Utilizing deep neural networks to address those systems, we developed the following generative systems: a grid-based generative pattern, a cellular automata (CA) system, and a stigmergic behavior system, respectively. The DL models used were CycleGAN, StyleGAN, and the datasets were curated to represent the input needed for the networks' training:

1. Volumetric Configuration: Cellular Automata (CA) models in 3D were used, and 2D depiction of their aggregation at different levels was pursued. CA can be interpreted as formal or spatial organization methods with rules to allow "growth" to occur in emergent ways of self-organization (Wolfram, 2018). The generated CA patterns, which reflect the self-organized aggregation inspired by the local "bidonvilles" (Figure 3), were configured into three distinct density programs, divided by site axes. into zones with different densities (dataset: 2000 images of 1024x1024).

2. Organizational grid system: The investigation of the possible aggregation of modular grid-based cells on various scales and how they can relate to an architectural program was pursued. As such, generative grid patterns were parametrically designed at various scales, for dataset generation using a generative Rhino/Grasshopper algorithm, and dataset augmentation (dataset: 2000 images of 1024x1024) (Figure 3).

3. Structural System: Stigmergy is a mechanism of indirect coordination through environments between agents and/or actions (Heylighen, 2015). Like the mycelium network, this agent-based modeling system can be controlled and generate natural emergent responses. The stigmergic simulations were performed by controlled parameters that simulated site conditions and environmental stimulus to provide a susceptible environment for mycelium growth (dataset: 2000 images of 1024x1024), (Figure 3).

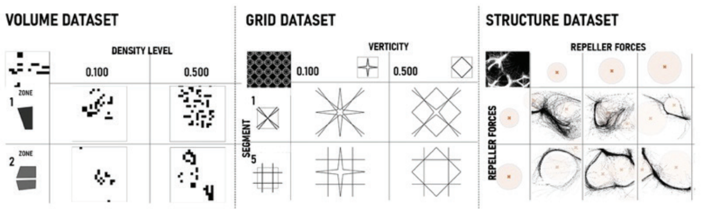


Figure 3. Dataset curation through three systems explored.

Phase 2: Systems for Encoding Design Intentions

Coupled with each of the aforementioned three datasets, microscopic 2D visual references of mycelium which represent the structure of material and its growth were used (dataset: 2000 images of 1024x1024) (Figure 4). The first round of DL trainings involved training three CycleGAN models for breeding the grid-based, volumetric, structural/environmental, each with mycelium datasets in parallel explorations. The grid-based dataset was bred with the mycelium growth patterns to organize the natural mycelium network and lead to a regulated organization. For the volumetric training, the mycelium pattern was intended to adopt the proposed spatial aggregation, while for the stigmergic system, breeding it with the mycelium system was to offer domain translation and let the mycelium growth pattern follow the controlled vertical structure.

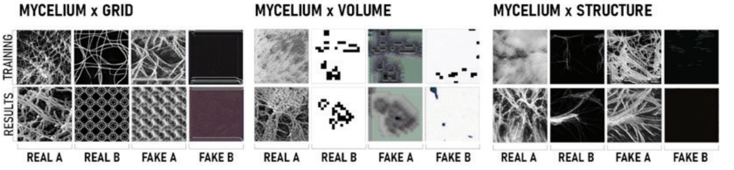


Figure 4. Datasets and results of three DL model training experiments using CycleGAN.

For each of the 3 CycleGAN models, training was stopped after 200 epochs (training iterations) and the models were tested with new datasets (different from the training datasets). For the grid-based case, the training objective was to examine the planar relationships of grid organizations. For each of the 3 CycleGAN models, training was stopped after 200 epochs (training iterations) and the models were tested with new datasets (different from the training datasets). For the grid-based case, the training objective was to examine the planar relationships of grid organizations. The stigmergic model training yielded mycelium-like spatial organizations that followed the agent-based network. This training was aimed at referencing the diagrammatic building section frames and reimagining them in mycelium-like patterns. The growth of the mycelium network offered novel design opportunities. For training the volumetric datasets with mycelium, processing, and encoding design intention, the 3D information needed to be decoded into 2D, to achieve architectural spatial intelligence (Fox, 2010). Therefore, nolli map-like images were used to show spatial positive-negative interactions, and during the structural pattern learning, the network produced architecturally meaningful images. The training aimed to reference diagrammatic building spaces and rethink spaces in terms of mycelium language. Testing was performed using new cellular automata system images in contextualization with the site. As output of this breeding, the CA behaviors were translated into the mycelium structure spatially at various scales. All testing experiments were performed using 200 images of 512x512 pixels.

Phase 3: Dimensionality Reduction and Post-Processing

At the project global scale (site level), the results from the DL models were used to inform five contextualized scenarios of volumetric designs. The selected volumetric CA models became input for a dimensionality reduction method, to create a matrix of 100 models, utilizing a self-organizing map (SOM), for design space expansion and articulation (Harding, 2017). Next, the SOM map was utilized in training a custom 3D GAN to find discrete relations among them. In post-rationalizing, four models were selected for further exploration, representing the most comprehensive preferred conditions for all the intelligence that has been encoded in the process. Moving forward, one model was selected for project development (Figure 5). In further processing, the overall volume was processed using the Cocoon tool, yielding a malleable structure (Figure 5). For the overall structure, defined by stigmergy, the same cloud point model was integrated to simulate the stigmergic behavior of live agents to define optimum collective paths, representing the speculative alive mycelium, as intended. In terms of the project envelope, using the rationalized geometry and the stigmergy simulation, the resulted geometry was identified as the stationary (dead) mycelium. Finally, using the Chromodoris tool, assigning 3D volumetrics was pursued, yielding curvilinear representation of the project at the global scale.

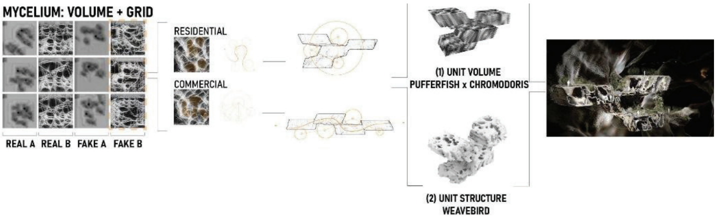


Figure 5. Post-processing phases and using the SOM method for design space articulation.

Phase 4: Project Development

In the final portion of the workflow, at local scale, we develop two cluster cases according to the bottom-up strategies that have guided our process thus far. We interacted with the same training-evaluation processes as before, however, here, the referenced point cloud is made up of the average color values of selected seeds from the grid training. The seeds are selected based upon their spatial qualities, which are most appropriate considering the residential and commercial typologies (Figure 6).

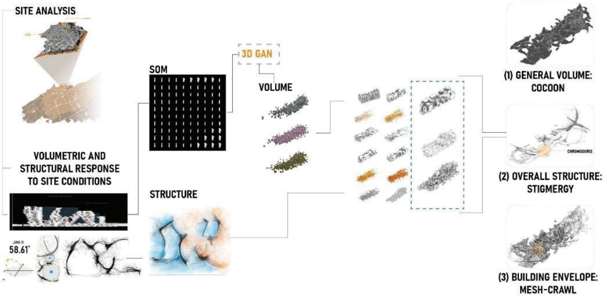


Figure 6. Using grasshopper's Chromodoris, we refined each massing to achieve habitable clusters.

4. Discussion of Results

This adaptive workflow and the node-based approach allowed risk-taking and open-ended exploration. In evaluating the nodes and their connections in the workflow, attention was drawn to three investigated types of connections: sequential, parallel, and branching. The overall workflow followed a sequential logic, with the result of each model's output feeding into the input of the next. Parallel strategies were used when two sets of interconnected DL models were pursued in a linear parallel approach. Branching became useful to break down a task into sub-tasks defined by separate sets of DL that can be later combined into a design solution. In terms of design agency, different levels were experimented with, including dataset curation, identifying the types of DL networks to train, and formulating different types of connections and combinations of those connections. An important issue addressed in this investigation was the evaluation of the DL-assisted design process, through the lens of creativity. According to Demis Hassabis (2018), AI's creativity still lies in its "interpolation" and "extrapolation" capabilities. While interpolation can be defined as the averaging of training examples, that are still contained within the field, extrapolation suggests a higher- level of creativity, in which the boundaries of that field are extended. We can draw the conclusion that CycleGAN has interpolative qualities, and was therefore heavily used for the mapping of two different domains in an unanticipated way, which allowed for a higher degree of exploration. StyleGAN was averaging the input datasets, serving more as an interpolation strategy; augmenting datasets for feeding CycleGAN in some cases. Importantly, AI-creativity combined with human creativity creates a more complex sense of creativity that can be evaluated as process-based rather than product-based (Leach, 2021). We investigated how a creative state can emerge when different levels of design are addressed by different DL models. This way, the design space becomes ever-growing, and decision-making alters that space, widening and narrowing it to unfold design exploration. In terms of employing mycelium as a natural system, its growth behavior informed the work in two phases. First, mycelium's growth has led to informing intelligent system behavior in design exploration, within the DL experimentation process. Second, this intelligent growth was interpreted into design with dynamic growth characteristics (employing mycelium as a building material with the possibility of live mycelium) towards an adaptive structure (Figure 7).

Figure 7. The design outcome in section perspective as the result of the experimental prototype.

5. Conclusions and Future Work

The methodology described in this paper tackles the application of AI in a process-based workflow. Unlike several existing DL applications which target experimental results through applying a discrete model, our work utilizes DL to handle the complexity and multiplicity of the architectural design process. This type of design approach increases the chance for creative design thinking by focusing on iterative collaboration. Our approach favors clarity of architectural intentions; augmenting agency for evaluating the DL-generated outcomes. The method facilitated the breakdown of the architectural design process into systems and tasks with specific layers, allowing more control over local decisions and result evaluation. Essential to this is the investigation of such levels of agency, whether at dataset curation, or selecting the type of DL model that can address a specific design task, in addition to the modes of connections amongst multiple models. Evaluating the success of the DL results was achieved by simulating the natural behavior of the material in the physical experimentation, which helped in understanding the mycelium, and the patterns it generates. Also, integral to this work is the attempt to integrate and insert data from physical experiments to inform the datasets of the neural networks used in the proposed process. As a current limitation, like any other novel method, the DL models' computation was time-consuming, with failing experiments and adjustments of workflow connections. However, in applying this research, establishing familiarity with DL models facilitates conducting such experiments with higher efficiency. The research is part of a boarder project to investigate AI systems and combinations of different connection strategies, report on their success or failure, and push AI research forward in architecture. Future work involves further testing and applications, as well as experimenting with new connection strategies for AI models. Further development also involves identifying methods that support encoding design intentions, like the way the behavior of agent-based systems was curated in this work.

References

Alexander, C. (1968). Systems generating systems. Architectural Design, 38 605-610.

Beier, R. (2019). *From the City to the Desert: Analysing shantytown resettlement in Casablanca, Morocco, from residents' perspectives*, Logos Verlag Berlin GmbH.

Blast-Studio. (2020). *Tree Column, 3d printed mycelium column from used coffee cups* [Online]. Available: https://www.blast-studio.com/tree-column.

Block, P. & Van Mele, T. (2017). *MycoTree* [Online]. Retrieved from: https://block.arch.ethz.ch/brg/project/mycotree-seoul-architecture-biennale-2017.

Bolojan, D., & Vermisso, E. (2020). Deep Learning as heuristic approach for architectural concept generation. ICCC, 2020. 98-105.

Chaillou, S. (2020). ArchiGAN: Artificial Intelligence x Architecture. In: YUAN, P. F., XIE, M., LEACH, N., YAO, J. & WANG, X. (eds.) *Architectural Intelligence: Selected Papers from the 1st International Conference on Computational Design and Robotic Fabrication (CDRF 2019).* Singapore: Springer Nature Singapore.

Forbes, A. (20220). "Creative AI: From Expressive Mimicry to Critical Inquiry". Artnodes, 2020, Num. 26, pp. 1-10, https://doi.org/10.7238/a.v0i26.3370.

Fox, W. L. (2010). Spatial Intelligence: New Futures for Architecture. *Places Journal*. Retrieved from: https://placesjournal.org/article/spatial-intelligence-new-futures-for-architecture/?cn-reloaded=1.

Gupta, I. & Nagpal, G. (2020). *Artificial Intelligence and Expert Systems,* Stylus Publishing, LLC.

Hassabis, D. (2018). Creativity and AI. *The Rothschild Foundation Lecture*. The Royal Academy of Arts.

Huang, J., Johanes, M., Kim, F. C., Doumpioti, C. & Holz, G.-C. (2021). On GANs, NLP and Architecture: Combining Human and Machine Intelligences for the Generation and Evaluation of Meaningful Designs. *Technology| Architecture+ Design*, 5, 207-224. https://doi.org/10.1080/24751448.2021.1967060

Koh, I. (2021). AI Sampling.

Leach, N. (2021). *Architecture in the Age of Artificial Intelligence: An Introduction to AI for Architects*, Bloomsbury Visual Arts.

Negroponte, N. (1973). *Architect-Machine Symbiosis. The Architecture Machine: Toward a More Human Environment*. The MIT Press. https://doi.org/10.7551/mitpress/8269.001.0001

Watanabe, M. (2004). *ALGORITHMIC DESIGN / INDUCTION DESIGN, Three Kinds of Flow / Three Stations* [Online]. Retrieved from: https://www.makoto-architect.com/kashiwanohaCSt.html.

Wolfram, S. (2018). *Cellular automata and complexity: collected papers*, crc Press.

AN IMAGE-BASED MACHINE LEARNING METHOD FOR URBAN FEATURES PREDICTION WITH THREE-DIMENSIONAL BUILDING INFORMATION

BOWEN QIN[1]and HAO ZHENG[2]
[1]*Archi-tectonics NYC, LLC.*
[1]*qinbowen0725@gmail.com, 0000-0001-5307-4933*
[1]*City University of Hong Kong.*
[2]*hazhen@cityu.edu.hk, 0000-0001-5769-6035*

Abstract. Machine learning has been proven to be a very efficient tool in urban analysis, using models trained with big data. We have seen research that applies a generative adversarial network (GAN) to train models, feeding the street map and visualized urban characteristics to predict certain urban features. However, in most cases, the input map is a two-dimensional (2D) map that only stores the land type data (e.g., building, street, green space), hence reducing building information to only the ground-floor area. The identities of buildings with similar floor areas can be hugely different, which may contribute to the prediction errors in previous machine-learning models. In this research, we emphasize the importance of the use of an image-based neural network to analyze the relationship between urban features and the constructed environment. We compare the model that uses traditional street color maps as the input set, against a new input set with more detailed building data. Once trained, the model with the enhanced input set yields output at a higher level of accuracy in certain areas. We apply the new model framework to three selected urban features predictions: rental price, building energy cost, and food sanitary ratio. A broad range of new research could be conduct with our new framework.

Keywords. Artificial Intelligence, Generative Adversarial Network, Urban Features, Building Elevation, Open-source Data, Prediction.

1. Introduction

Since urban planning has become a mainstream discipline, numerous urban data, subjective (by massive questionnaire) or objective (measured or simulated data), have been collected and shared as open-source data, and continue to be updated today. Many cities provide public access to their datasets. Thanks to the accessibility of those datasets, researchers have been able to conduct a wide range of interesting and valuable studies such as interactive urban data visualization (Louna Al Bondakji, 2019), urban data-aided pedagogy (Bjork, 2001), and urban growth simulation (Cemal et al., 2016). In the past decade, artificial intelligence has developed a lot. Machine learning, as the most successful branch of artificial intelligence, has been widely embraced in urban

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 109-118.

analysis. Compared with traditional data-driven method, machine learning offers many advantages. First, it always covers a wide range of data, and sometimes unearths hidden factors. Secondly, it can quantify the influence of different kind of factors, even subjective ones. Thirdly, machine learning provides a convincing testing method to evaluate the accuracy of the model. Applications of machine learning in urban planning includes generating street images (Steinfeld, 2019), street image semantic segmentation to aid urban renewal (Ziyi Tang et al., 2020), and predicting human perceptions of a place (Zhang et al, 2018).

1.1. GAN IMAGE-BASED MACHINE LEARNING.

The generative adversarial network (GAN), first proposed in 2014, is a learning framework that automatically generates deceptive data for comparison with the training sets (Goodfellow et al., 2014). The underlying idea is ingenious, being an adversarial process that simultaneously trains two models: a generative model (G) captures the data distribution, and a discriminative model (D) estimates the probability that a sample came from the training data rather than G (Goodfellow et al., 2014, p.1). The goal of this double model is to improve the performance of both simultaneously. If all goes well, the two adversarial models eventually reach a Nash equilibrium, in which the generator is demonstrably able to generate compelling deceptive data (Figure 1).

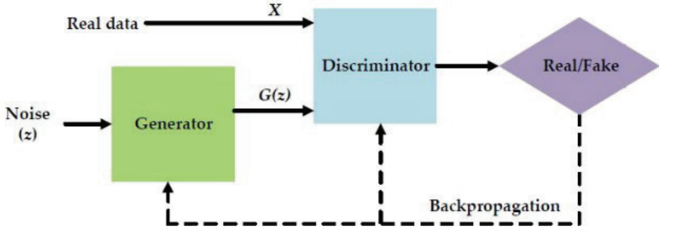


Figure 1. Basic structure of generative adversarial networks (Jie et al., 2020).

This model framework can hybridize with other machine learning models and is widely used for image/voice/video generation. For example, in one type of GAN image generation, input and output data are images of the same size. This kind of model is called an image translator (Figure 2). The eventual goal of its whole training process is to find a credible relationship between the input and output images so that a translated output image will be generated when a new image is fed into the trained model. The existing GAN models for image translation include pix2pix, CycleGAN, IcGAN, MUNIT, DRIT, and pix2pixHD.

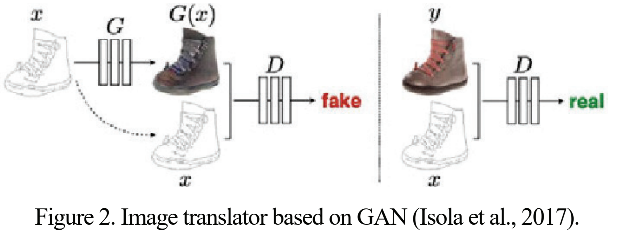


Figure 2. Image translator based on GAN (Isola et al., 2017).

1.2. PROBLEMS IN THE PREVIOUS GAN-AIDED URBAN RESEARCH

Image-based GANs have been widely used in urban studies. Examples include urban filling (Jiaqi et al., 2020), urban vitality prediction (Yunjuan et al., 2020), crime rate prediction (Jingyi and Hao, 2021) and land type classification (Sun and Hu, 2020. That research shares a common ground: the training input data are maps labeled by land type categories.

Many predicted urban indices are closely related to occupants' subjective feelings. However, the input bitmap developed by previous research only contains information about the horizontal surface, i.e., the ground type. Several famous theorists, such as Yoshinobu Ashihara, Manuel Delanda, and Jacobs J. Planner, have proposed fundamental explanations of that point. The hypothesis that information in ambient light specifies affordance is the culmination of ecological optics. Surfaces of the environment, horizontal and vertical, provide affordances for people; they provide the basic unit to perceive the environment (Jacobs J., 1986) (Figure 3). The identity of buildings with similar floor areas can greatly differ. Manhattan, for example, was planned on a grid that makes most street blocks similar in dimension and shape, resulting in proximal buildings on the same block being merged into one big building block in the eyes of AI. A 30-floor tower may be considered the same as six connected two-floor townhomes (Figure 4). Such elisions may contribute to the prediction errors in previous machine learning models. An optimized model, including the missing training information, is needed.

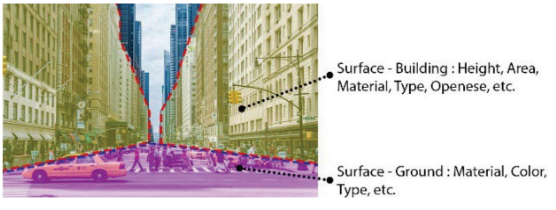


Figure 3. Surfaces in a city provide the basic unit to perceive the environment.

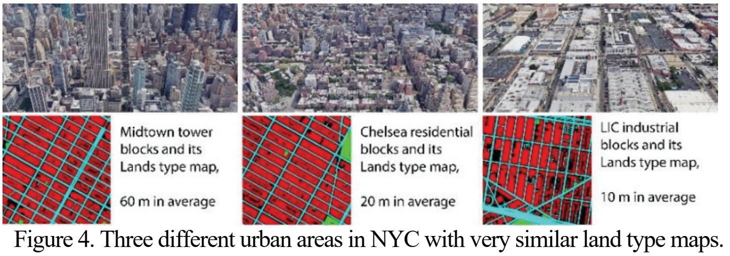


Figure 4. Three different urban areas in NYC with very similar land type maps.

1.3. PROJECT GOAL

This research emphasized the importance of using an image-based neural network to analyze the relationship between urban features and the constructed environment.

Based on insights from previous and existing research methods, we discovered a point that can significantly improve the application of the model in urban research. This research aimed to find an optimized framework for image-based GAN-aided urban research that can increase the accuracy of existing and potential future research, thus promoting the application of GAN models in urban research. The effectiveness of the model was also tested on sample research environments.

2. Methodology

The proposed method can be summarized into three steps: collecting and labeling the input map data, predicting output data collection and visualization, and training and testing the GAN models.

2.1. COLLECTING THE INPUT MAP DATA

The input maps are the most vital part of this research. The general idea behind their selection was to collect two sets of data, one based on labeling rules in previous research and the second with more building facade information and a more detailed labeling rule to represent them.

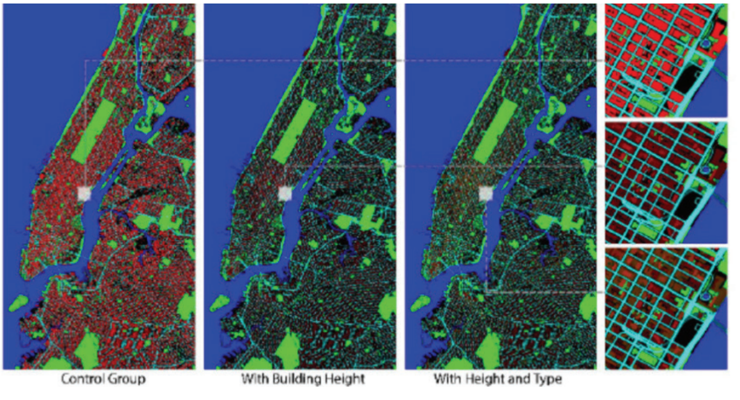


Figure 5. Input group from left to right: control Group 1, Group 2 with height information, and Group 3 with height and building type information.

New York City (NYC) was selected as the exampling city for the convenience of data collection from open sources. The research area includes the area of New York City north as far as Fort Tryon Park in Manhattan, south to the south end of Prospect Park in Brooklyn, east to the east end of St. Michael's Cemetery in Queens, and west to cover Governor's Island. The selected area contains the most representative urban texture in NYC (Figure 5). Mapbox and Google Maps Platforms were used to customize the map style, which contains geographic information, including detailed façade information. In the first training input set, we use a labeling color rule from former research, comprising five colors to represent the five basic land types: buildings, roads, green space, water, and land. This labeled set is the control group (Figure 6).

As mentioned, many levels of detailed building information can be added to the input maps. Among those, height is the most vital since it can directly convert the 2D data into 3D data. The building type information was also collected and applied to the input maps. Other factors, such as building age, are collected via Pluto open-source data, and façade material is collected by semantic segmentation machine learning on street-view data. This more detailed building façade information will be introduced in the next stage of our research. Height was added to the second input data set. Instead of using white (R:255, G:255, B:255) for buildings in the control group, we use the R-value to represent the height of the buildings. Buildings under 10 meters are coded by R = 50, buildings above 100 meters by R = 255, and buildings in-between are distributed linearly from 50 to 255. The third data set incorporates building type information into the G to represent residential and commercial types (Figure 6). These three map sets are cut into small image sizes to fit our model's input interface. These four variations on the input maps will be fed into different models for comparison.

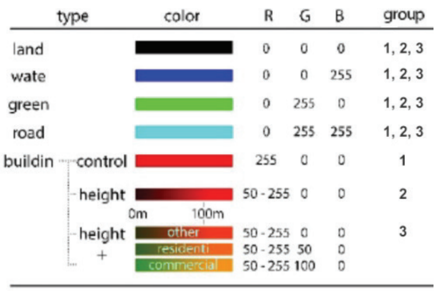


Figure 6. Labeling method for the three input groups

2.2. OUTPUT DATA COLLECTION AND HEATMAP MAKING

The primary purpose of this research was to achieve better performance for existing methods since successful processing of three-dimensionality could be applied to many different topics. The urban map with height and land type information reflected the figure-ground relation of some urban regions and the scale of urban space, which impacts how people perceive and react to the space. To test the efficiency of this augmented model, we selected three urban features related to 3D building information: crime arrest rates, noise complaint rates, and food hygiene rates. Crime arrest rates are highly correlated with development status and urban density, noise complaint rates are correlated with building and population density, and food hygiene rates relate to building development status and type. Data in these categories were collected via the NYC open data platform and the Kaggle open data source. Using the food sanity ratio as an example, the ratio summarises restaurants' NYC food hygiene grades from A to C. Each index is aligned with positional information indicating the location of the restaurant. Further, Rhino, Grasshopper, and Python were used to depict the heat distribution map, showing the distribution of food quality across NYC. Finally, the picture size was refined to align the heat map and the input map in latitude and longitude. The completed map was then partitioned into small squares, like the input data.

The three input and three output sets collected will pair into nine groupings: three for food sanity ratio prediction (Group A), three to predict arrest ratio (Group B), and three for noise complaint ratio (Group C) (Figure 7). These 3 × 3 models will allow performance comparison of the new model along two dimensions. First, by horizontally comparing the errors in each of the three models in the same group, we can see how the new model performed with the former model. Additionally, vertically comparing error analysis between different groups can reveal what urban features are more related to building façade information.

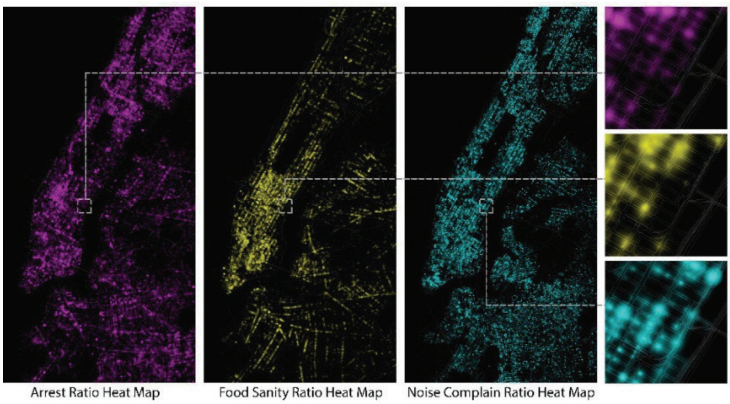


Figure 7. From left to right, three different output heatmap images: crime arrest ratio, food sanity ratio, and noise complaint ratio.

2.3. STRUCTURE AND TRAINING OF THE GAN MODEL

Different GAN models were trained using the prepared data sets. As mentioned, the GAN framework selected for this research is an image translator, specifically pix2pixHD. Technically, there are more parameters to fine-tune the training process. After some debugging, the image size was set to 512 × 512 to match the collected image size, the epochs for constant training to 80, the learning decay rate to 20, the initial learning rate for Adam (the optimizer for training) to 0.0003, and the use of instance maps were turned off. All other hyper-parameters were set to their default values. When pre-processing the training data, we divided the training set (314 maps) into a training group of 276 maps (88%) and a testing group of 48 (12%). (Figure 8).

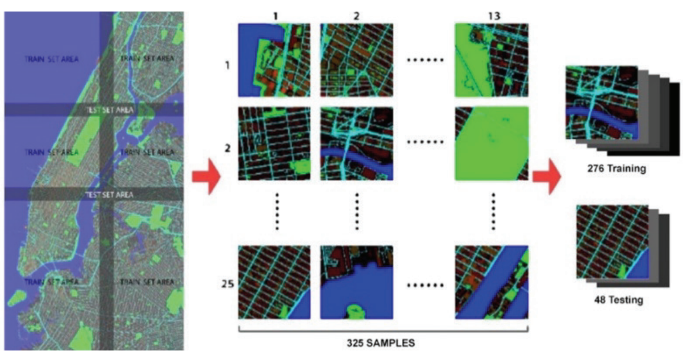


Figure 8. Training data and dividing methods

3. Results and discussion

To compare the performance of models in the same training group, the prediction error between the synthesized map and the ground-truth map needs to be quantified. A method developed by Yunjuan (2020) was used to predict city vitality for better comparability with former research. The error can be calculated by dividing the testing set and the corresponding real image into many 16 × 16-pixel blocks and calculating the average difference of one pixel in each RGB channel (Equation 1) (Yunjuan, 2020, p.163).

$$Error\ R.G.B = \Sigma \mid sum(M\_(R.G.B)\ ) - sum(m\_(R.G.B)\ ) \mid / N$$

$$M\_R\ \ m\_(R\ )\ \epsilon Z\text{^}(16 \times 16)$$

where, N = Number of pixels.

Figure 9 shows the image pairs of the input image, synthesized image, and ground truth in each training epoch of the training process of Group A, model 2. From the first 20 epochs, the accuracy of the synthesized result is low. Then after around 50 epochs, the accuracy is improved to a certain level. The prediction yielded good accuracy after 100 epochs and did not improve significantly after that; thus, the training was stopped and the model stored at 100 epochs.

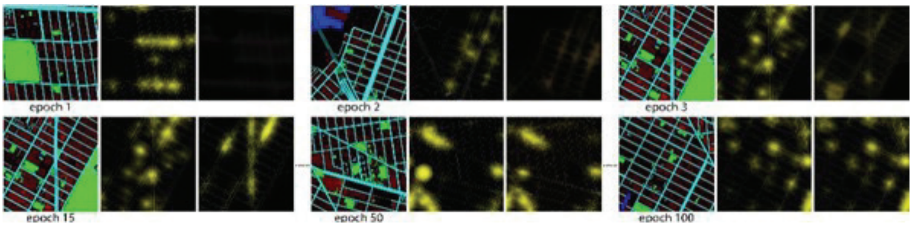


Figure 9. Training images and synthesized images in six training epochs.

3.1. DETAILED RESULT OF TRAINING

After nine models were trained, the predicted result within groups was compared, showing the performance of the original and optimized model. In addition, the result between the three groups was also compared, which showed various improvements by introducing the 3D building information in predicting different urban features.

By analyzing the results among those nine models and the reasons behind these differences, the importance of 3D building information in different contexts can be identified. Moreover, analyzing the characteristics of specific areas where the improvement of new models is relatively low can help identify what other features might be considered missing so that the model can be improved in the next step.

3.1.1. Group A, horizontal comparison

Analyzing the training result with different models of Group A made it possible to make conjectures about factors related to the food sanity ratio in NYC. Figure 10 shows the testing result of all three models in Group A. Synthesized result 1 shows the result of the old model, synthesized result 2 shows the model including building height, and result 3 shows the model including both building height and functional information.

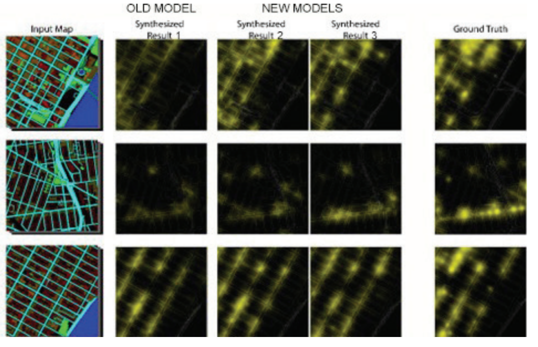


Figure 10. Prediction in different areas in Group A.

The absolute differences between the average errors of the three models are shown in table 1. We used a black image as a reference case for error calculation. Errors B were low in all three models because orange was used as the main color to draw the urban heat map. By analyzing the result, the new model frameworks show better overall accuracy than the former model. At the same time, the model with both height and functional information shows some but not much improvement compared to the model with only height information, which proves that building height information is a more crucial factor than building type in urban food sanity ratio prediction.

Model index	Error R	Error G	Error B	Average Error
1	41.3	37.5	4.9	27.9
2	32.3	28.1	4.3	21.6
3	31.1	27.7	4.4	21.1
Compare Case	61.6	67.4	11.2	46.7

Table 1. Error value of three models in Group A.

The improvement, however, varies by area. The reason for this can be seen by analyzing the characteristics of specific areas. For the areas in figure 10 top, the accuracy of models 2 and 3 show an apparent improvement. However, the building type information does not show as much improvement as building height since the building heights are changing drastically in that area. In figure 10 middle, the accuracy of model 3 shows improvement, but model 2 does not show much improvement because the building type is complex in that area, while the height of buildings is not changing much. In figure 8 bottom, all three models show high and similar prediction accuracy because, in those areas, the building height is low, or type is not changing much. As noted earlier, many more input factors are missing.

3.1.2. Vertical comparison of Groups A, B, and C

Compared to the food sanitary ratio, the prediction of crime arrest ratio shows similar but not identical results (Figure 11). Unfortunately, there was no apparent improvement in complaint ratio prediction when using the new model. Therefore, height and building type are not the primary factors influencing crime in a city. The new model performs better, especially when the input area has drastic height and building type changes. Results also showed that the height factor has more weight on the food sanity ratio than the crime ratio.

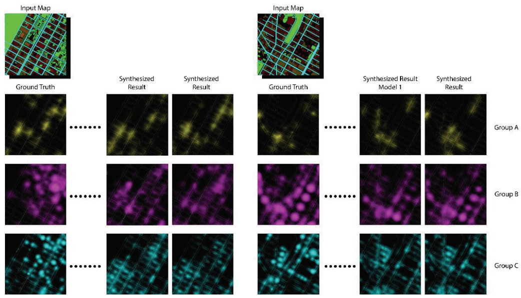


Figure 11 Sampled images show the vertical comparison of three Groups.

4. Conclusion

This paper considers the use of big data in urban research. An optimized GAN-based neural network algorithm was used to replace 2D land-type maps with 3D maps to provide better predictive power. Based on this model, a real-time feedback system could be established to help urban designers adjust the design. The model can also help urban theorists to find the weights of contributing factors to urban characteristics. It solves some problems of previous research while proving and extending the broad applicability of machine learning, specifically image-based models in urban studies. However, this case study is not the primary goal of this research, but it does, nevertheless, show relationships between built environment and food sanity ratio and crime arrest ratio. These results are encouraging for the future use of similar exampling.

Subsequent research will try to incorporate more detailed building information, such as building age and material, and attempt to find ways to increase the feature dimensionality. The current reliance on bitmaps limits representation to three color channels, making it hard to include more than six separate features. While 3-channel bitmaps are suitable for output data visualization. The purpose of using the same-sized bitmaps in the input is only to ensure that the information bears the same positional relationships as in the output. It is likely possible to adjust the machine learning model to accept more channels in each pixel so that more input factors can be added. Multiple bitmaps will also be tested as the input data to increase dimensionality.

A problem in this research is that the map has to be cut into smaller images for training. This leads to some large-scale factors being ignored, such as water bodies just outside the training area. A better model, in our opinion, should be able to address both the neighborhood and its overall environs simultaneously. The development of machine learning is rapid; ever more model frameworks are being proposed in computer science. Given that errors still exist, it is necessary to keep up with the newest methods.

References

Al Bondakji, L., Lammich, A. L., & Werner, L. C. (2019). ViBe (Virtual Berlin) - Immersive Interactive 3D Urban Data Visualization Immersive interactive 3D urban data visualization. In 37th eCAADe and 23rd SIGraDi Conference - Architecture in the Age of the 4th Industrial Revolution, eCAADe (pp. 83-90), https://doi.org/10.52842/conf.ecaade.2019.3.083.

Björk, B.C. (2001). Open source, open science, open course ware. In H. Penttilä (Ed.), Architectural Information Management, In 19th eCAADe Conference, Helsinki University of Technology, Department of Architecture, Espoo, (pp. 13-17), https://doi.org/10.52842/conf.ecaade.2001.013.

Bingöl, C.K. & Çolako?lu, B. (2016). Agent-Based Urban Growth Simulation - A Case Study on Istanbul. In the 34th eCAADe Conference - Volume 2, University of Oulu, Oulu, Finland, 22-26 August 2016, (pp. 41-48), https://doi.org/10.52842/conf.ecaade.2016.2.041.

Gibson, J.J. (1986). The Ecological Approach to Visual Perception. Psychology Press

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. & Bengio, Y. (2014), Generative adversarial nets, Advances in neural information processing systems, 2672-2680.

He, J., & Zheng, H. (2021). Prediction of crime rate in urban neighborhoods based on machine learning. Engineering Applications of Artificial Intelligence. Volume 106, 104460. https://doi.org/10.1016/j.engappai.2021.104460.

Isola, P., Zhu, J. Y., Zhou, T., & Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks, arXiv preprint, (pp. 1125-1134), https://doi.org/10.48550/arXiv.1611.07004.

Shen, J. I. A. Q. I., Liu, C. H. U. A. N., Ren, Y., & Zheng, H. (2020). Machine Learning Assisted Urban Filling. In the 25th CAADRIA Conference Volume 2, Chulalongkorn University, Bangkok, Thailand, (pp. 679-688), https://doi.org/10.52842/conf.caadria.2020.2.679.

Steinfeld, K., Park, K., Menges, A., & Walker, S. (2019). Fresh Eyes: A Framework for the Application of Machine Learning to Generative Architectural Design, and a Report of Activities at Smart geometry. In the 18th International Conference on CAAD Futures, Daejeon, Korea. 2018. https://doi.org/10.1007/978-981-13-8410-3_3.

Sun, J., Jiang, L. & Zheng, H. (2020). A Machine Learning Method of Predicting Behavior Vitality Using Open Source Data. In the 40th Annual Conference of the Association for Computer Aided Design in Architecture, ACADIA, (pp. 160-168),.

Sun, C., & Hu, W. (2020). A Rapid Building Density Survey Method Based on Improved Unet. In the 25th 5th International Conference on Computer-Aided Architectural Design Research in Asia, CAADRIA, (pp. 679-688)

Tang, Z., Ye, Y., Jiang, Z., Fu, C., Huang, R., & Yao, D. (2020). A data-informed analytical approach to human-scale greenway planning: Integrating multi-sourced urban data with machine learning algorithms. Urban Forestry & Urban Greening, 56, 126871, https://doi.org/10.1016/j.ufug.2020.126871.

Zhang, F., Zhou, B., Liu, L., Liu, Y., Fung, H. H., Lin, H., & Ratti, C. (2018). Measuring human perceptions of a large-scale urban region using machine learning. Landscape and Urban Planning (pp. 148-160), https://doi.org/10.1016/j.landurbplan.2018.08.020.

AN INTEGRATED APPLICATION OF BUILDING INFORMATION MODELING, COMPUTER-AIDED MANUFACTURING, MACHINE LEARNING, AND THE INTERNET OF THINGS

A Hybrid Stadium as a Case Study

BOWEN QIN[1], TIANYU WANG[2], WINKA DUBBELDAM[3], JUSTIN KORHAMMER[4] CONG HUANG[5], DENGGUO WU[6] and HONGMING JIANG[7]
[1,3,4]*Archi-Tectonics, NYC, LLC.*
[2,5,6,7]*Huadong Engineering Corporation, Ltd*
[1]*qinbowen0725@gmail.com, 0000-0001-5307-4933*
[2]*david.wty@163.com, 0000-0003-3108-721X*
[3]*wd@archi-tectonics.com*
[3]*jk@archi-tectonics.com*

Abstract Using building information modeling (BIM), design teams can maximize decision-making efficiency by using one data-sharing platform and one integrated digital model. We present a novel built-project that integrates computer-aided manufacturing (CAM), machine learning, and internet of things (IoT) with a BIM process. The goal is to solve design and manufacturing problems and tightly connect virtual design data with the real-world construction process. Introducing CAM in the early phase reduces the cost of reverse engineering and optimizes material and construction costs. Machine learning has been used for curtain wall panel standardization and to solve manufacturing problems. The IoT concept focuses on directing the data stream onto real-world objects and updating the digital model using feedback from real-world monitoring. Using the integrated application in this project, a hybrid Olympic stadium, helped to accelerate the process of design and reduced the time and cost for manufacturing and construction.

Keywords Building Information Modeling (BIM), Computer Aided Manufacturing (CAM), Machine Learning, Internet of Things (IoT), Panel Standardization.

1. Introduction

In recent years, building information modeling (BIM) has become the most powerful tool for increasing efficiency and accuracy within AEC (Architecture, Engineering, and Construction) industries. By integrating different disciplines into one data-sharing platform during the design phase, a much more accurate digital representation, or model, can be created, significantly decreasing data loss during the transitions between project phases (Borrmann et al., 2018). Many data-based advanced technologies can be integrated with the BIM data stream to address specific problems. Recent work in the BIM field has focused on BIM-enabled AR for construction (Chernick et al., 2021), data transmission for cloud-based BIM (Kereshmeh et al., 2016), BIM vocabulary and syntactical study (Al-Assaf & Clayton, 2017), diagram-aided design in BIM (Al-Assaf & Clayton, 2021), and machine learning-aided fire code classification in BIM (Albassel & Waly, 2021). These studies expanded the scope of BIM application.

1.1. MACHINE LEARNING AND DATA CLUSTERING

Over the past decade, machine learning has proven to be a very powerful tool in many fields, including architecture. Clustering is a basic but important machine-learning method. Clustering is the task of dividing data into several groups based on similarity. It has been widely applied in many fields, e.g., in marketing to characterize different customers, in biology for species classification, and in business for fraud detection. In architecture, clustering has been used for architecture type classification (Abdulrahman et al., 2020), urban features recognition (Chaszar & Beirã, 2013), and data visualization in building space (Meng et al., 2020). The potential for applications of clustering related to design construction is substantial, not least because it could help rationalize the application of non-linear parametrical design to buildable segments. In this paper, we refer to the Soumaya Museum project conducted by FR-EE in 2011, which applied a clustering technique to rationalize the hexagon panels on its façade.

1.2. CONSTRUCTION AND USER MANUAL

Among the AEC industries, architecture and engineering are the most digitized and can smoothly embrace BIM, achieving a significant reduction in lost information. However, the use of BIM in construction is lagging. Traditional drawings from the design team continue to be the primary output for the contractor or manufacturer, even in BIM projects. Previous study has proposed an AR-aided method based on real-world alignment (Chernick et al., 2019), but this promise to complement the drawing set rather than replace it. Construction is a dynamic process that combines raw materials based on the contractor's experience and the detailed user manual. Designs focus more on the product than the process. An analogy can be made to assembling an Ikea cabinet after seeing only a model of the final product; unless you are an expert in Ikea products, you may also require a step-by-step manual. In most cases, the modern practice in BIM is not to focus on creating the step-by-step manual but rather on the design result.

Moreover, the processing technology required for complex structures may not fully satisfy the requirements of the engineered structure. In conclusion, there remains a gap between the design goal and the resulting construction that is not being filled by BIM.

1.3. PROJECT BACKGROUND

Designed by Archi-Tectonics, the table tennis stadium project is part of a larger master plan for the 2022 Asian Games. It is a 116-acre sustainable landscape with two Olympic stadiums, a shopping mall, an exhibition center, a fitness center, and two underground parking garages. The table tennis stadium itself was inspired by an ancient Chinese object called a Cong, which is an intersection of two different geometries—a cylinder and a square. The shape of the table tennis stadium is two intersecting ellipses which are completely asymmetrical and non-linear.

This complex project involves many stakeholders from different fields of expertise, which poses many challenges in coordinating the massive amount of information. In addition, as the project is being built for the 2022 Asian Games, which will be held in the autumn of 2023, it is extremely time sensitive.

1.4. PROJECT GOAL

As described above, in the traditional BIM process, information continues to be lost throughout the process, especially when transferring data from design entities to manufacturers and construction workers. This paper aims to discover how to create a workflow that can carry information from the digital design to the construction site with much less information loss.

The goals are as follows: to establish a continuous flow of information on a shared platform for all stakeholders throughout the project, including the manufacturers and construction entities; to build a central digital model that contains data from all the disciplines; to apply advanced, data-driven technologies to solve specific problems during a stage when information usually tends to be lost; and to improve accuracy and efficiency and minimize costs and losses by all teams. This paper consists of two parts: the first describes the overall BIM strategy, and the second uses the curtain wall system as an example to show how machine learning, the internet of things, and other techniques have been used to deliver data from design to build.

2. Methodology

2.1. GENERAL METHODOLOGIES

To manipulate data as smoothly as possible, this project fully adopted the BIM process in the second stage, which is when most stakeholders in other disciplines, we included the main manufacturers and contractors early in this stage.

We based our BIM pipeline on the classic method using one software system for cloud data sharing and one central software model by Revit. Rather than using only the central software, we chose to adopt the "open BIM" approach, which allows all stakeholders to use different data media, with certain restrictions, on a shared platform. This approach will be clarified in the next section. We will then show how we applied machine learning and the IoT to the manufacturing of the curtain wall system.

2.2. . WORKFLOW FOR DATA

We assembled a large multi-disciplinary stakeholder team. First, working with the code consultants, we developed a unique set of BIM standards for this project, based on local codes and regulations as well as the actual situation of each participant. Next, a mutual data-sharing cloud-based platform was established. We chose ProjectWise to serve as the main file exchange platform. We choose to use open BIM, which allows different file formats to be coordinated. There are two reasons that open BIM was the preferred choice. First, there is no single software product that could provide the best solution for all disciplines. Second, every industry already has its own preferred software, so open BIM would serve best for interdisciplinary coordination. We set up the Revit model as the central integrated model to be used by the architect, landscape designer, and the structural and MEP engineers. This central model focuses on the result of the building.

Different software systems were used for different purposes (Figure 1). For example, the structural analysis took place on SAP2000, which is a mature finite element analysis software based on mesh.

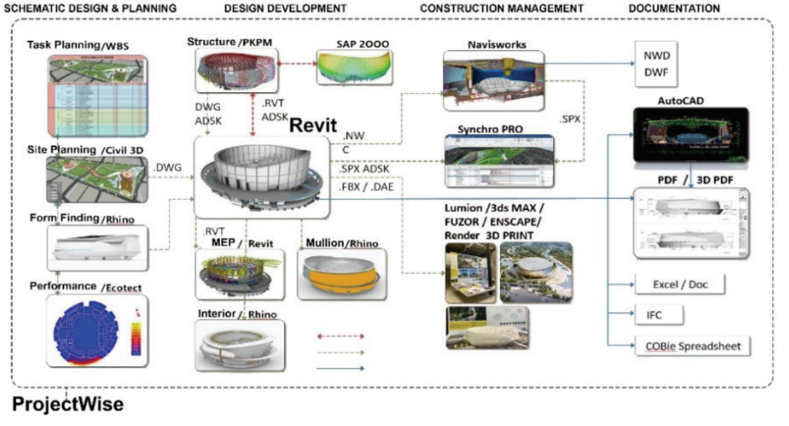


Figure 1. Revit-centered BIM software on ProjectWise platform

At this stage, the goal was to detail the design of every part of the building, which means the output model of this stage would be the same as the built project—a fully integrated digital model. All engineering systems would be simulated and proved to work. The BIM was broken down into eight systems (Figure 2), depending on professions and location.

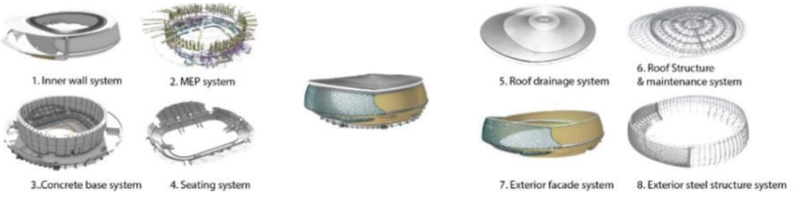


Figure 2. Eight systems for the hybrid building

2.3. EXTERIOR CURTAIN WALL

Among the eight systems, two of them— the façade steel structure and the curtain wall system attached to it—together form the exterior envelope, which was designed as the intersection of several freeform surfaces. Half of the surfaces is covered in brass shingle panels; the other half is covered by tilted rhombus shapes. Figure 3 shows the detailed sub-layers of the glass and metal curtain wall systems. This stage could present challenges in terms of translating the design information into built objects, considering the gap between design and manufacturing in a traditional workflow. However, manufacturing stakeholders—including local suppliers and the procurement team—

were involved in the design development phase of the BIM process, which facilitated effective problem-solving. Our design team understood the requirements of the manufacturing method and thus set up a workflow from design to manufacture.

Figure 3. Exterior curtain wall system of the stadium

2.4. RATIONALIZATION OF STEEL STRUCTURE

The geometrical logic of the twisted diagrid structure is based on the freeform control surface and control grid curves (Figure 4). The challenge was that the tilted glass panels and the structure underneath are integrated, which means the design had to balance structural performance and aesthetic values in ways that were difficult to manufacture.

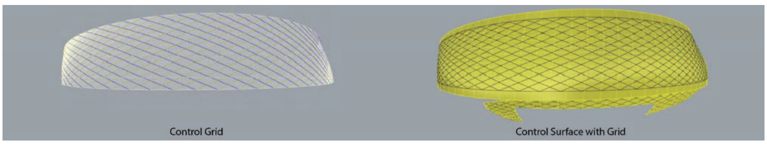


Figure 4. Control surface and control grid

In the original design, we established an equal-thickness shell model by offsetting the control surface in the direction of the surface normal. We then hollowed out the shell according to the control grid. The resulting model has two freeform surfaces: front and bottom. These cannot be flattened to a planar surface. The processing technique could not manufacture the freeform 3D geometry.

The manufacturing method is based on metal bending in which planar steel panels are bent to developable spatial surfaces. Based on this method, we rationalized the model to meet the manufacturing requirements. The mathematical detail has been explained in another paper (Wu et al., 2021). The final model was modified to sweep surfaces with rectangular section frames on the control grid based on arc length so that all surfaces would be developable (Figure 5). During manufacturing, the suppliers can CNC-cut the planar metal sheets, bend them separately to form the four surfaces of the structure, and assemble them onto structural tubes, which will be delivered to site for final assembly.

The rationalization process was conducted using Rhino and Grasshopper. We exported the design model to match the data format used by the CNC-cutting and bending machines, giving the supplier access to the file via ProjectWise.

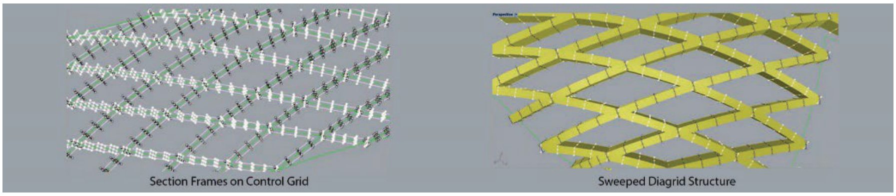


Figure 5. Section frame and diagrid structure by sweep

2.5. MACHINE LEARNING-AIDED PANEL STANDARDIZATION

Unlike the glass façade, the structure and appearance of the brass shingle façade are separate. The challenge here is that there are 6,435 panels, each of which has an identical spatial surface (Figure 6). It was impossible to manufacture each one locally.

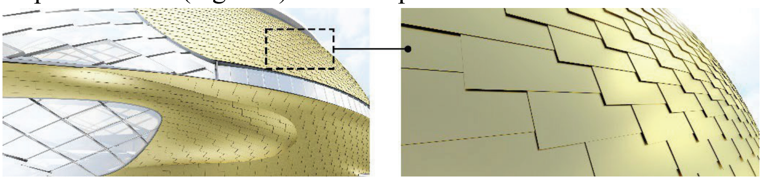


Figure 6. Original parametrical brass shingle panels

The goal was to standardize the panels to an acceptable number of groups and to simplify the spatial panels to planar panels, in response to the manufacturing requirement, the design team applied new techniques to optimize the design, aiming to avoid any reverse engineering and reduce information lost. We conducted a three-step method, aided by the machine-learning clustering technique, to achieve this level of rationalization.

First, we divided the panels into two groups: edge panels and center panels. The edge panels were more irregular and had to be processed separately. Regarding the central panels, the vital challenge was to find a way to divide them into a certain number of groups of similar panels so that they could be standardized to type. An unsupervised machine-learning model was used for this task. Lastly, one standardized panel was generated for each group of panels, based on the average features of all the panels in that group.

2.5.1. Pre-training preparation and features engineering

Features engineering is crucial for machine-learning training. Before applying the machine-learning method, the panels needed to be flattened to planar surfaces to reduce the number of features during training. This is because a planar quadrilateral can be defined only by the length of its four sides and the angles of its four corners, and it is much harder to define an undevelopable quadrilateral. To obtain a sufficiently accurate flattened surface, we extracted the four corner points of each panel, obtained a fitted plane using the PlaneFit battery in Grasshopper, and projected the panel to the fitted plane to generate the flattened surface. During the flattening process, we obtained the first feature for future training—the deviation between the fitted and original surfaces.

After flattening, twelve features were selected: one panel deviation rate (1); the lengths of four sides (4); the angles of three corners (the fourth was not needed mathematically) (3); the area of the panel (1); and the coordinates x, y, and z of the center point (3).

The first eight features are fundamental; the ninth feature, area, can be used as an estimate feature to stand in for lengths and angles features. The location features are not related to the shape of the panels, but the clustering of neighboring panels made the numbering process easier and improved efficiency during site assembly.

Feature normalization was conducted after feature selection. This process is necessary because the above-mentioned features were in varying scales. The scaling technique shifted and rescaled features to the same or a similar range (usually 0 and 1). This helped to speed up the model training and ensured that every feature was training proportionately. The features used in training were determined after multiple iterations of training, more details will be explained in the next paragraph.

2.5.2. Model training and comparison of results

We used two clustering methods for training: the K-mean model and the Gaussian Mixture Model (GMM). After comparing results and feature selection, we determined the best method for clustering.

We used three groups of features for training with both models: Group A contained deviation and area features; Group B contained the side lengths, the corner angles, and the deviation feature; Group C contained all features other than the area feature.

We set the K number for both models and determined the number of groups from 10 to 50. Figure 7 shows the visualized results for different K and feature groups. Each color in one single training result represents one cluster. The training results were very similar for both models in Group A, while the training speed of the K-mean was faster.

The training results in Group C for both models were similar as well; however, the location feature was too dominant and unnecessary. The K-mean algorithm is based on the Euclidean distance of features, while the GMM is based on probability and an assumption that each feature would fit a Gaussian distribution. Without going into detail on the mathematics involved, the conclusion is that when data distribution is more irregular, the GMM is more accurate, and when data distribution is more homogeneous, the K-mean model is accurate enough but faster.

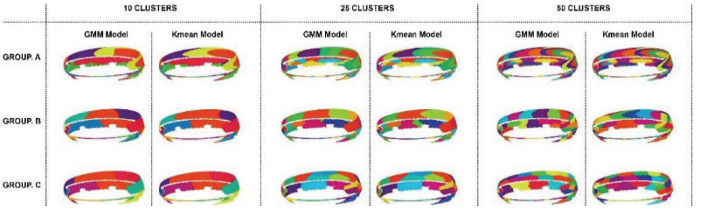
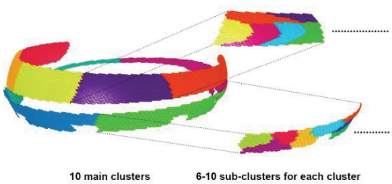


Figure 7. Training results for all three groups

Another way to reduce the training time and increase accuracy is to conduct multi-step clustering. Considering the training result in Figure 7, we conducted a two-step clustering process. First, we clustered the panels into 10 groups by K-mean with area feature; then, we trained each of the ten groups into 85 smaller groups by GMM with Group B features (Figure 8).



B-3	B-rec03	1850	1000	30	1.91	26	49.54	455	907.30
	B-rec04	1900	1000	30	1.96	118	230.93		
	B-rec05	1950	1000	30	2.01	165	331.40		
	B-rec06	2000	1000	30	2.06	124	255.44		
	B-rec17	1850	950	30	1.81	20	36.26		
	B-rec18	1900	950	30	1.86	2	3.72		
B-6	B-rec04	1900	1000	30	1.96	32	62.62	363	716.36
	B-rec05	1950	1000	30	2.01	94	188.80		
	B-rec06	2000	1000	30	2.06	135	278.10		
	B-rec15	1750	950	30	1.72	2	3.43		
	B-rec16	1800	950	30	1.76	26	45.86		
	B-rec17	1850	950	30	1.81	28	50.76		
	B-rec18	1900	950	30	1.86	23	42.83		
	B-rec19	1950	950	30	1.91	23	43.95		

Left: Figure 8. Final clustering method. Right: Figure 9. Sample spreadsheet for each panel

2.5.3. Panel standardization and deliverables

The previous step successfully decreased the number of panels to 85 groups. The last step was to generate a standardized panel in each group and then to number all the panels for manufacturing and later construction.

The method was straightforward; all panels in each group were oriented to overlay each other, then the average corner points of the four corners were calculated. Finally, a planar quadrilateral was created by connecting the four generated points by lines. The optimized model was created to the manufacturer's requirements with a label on each panel (Figure 9). These, combined with Excel spreadsheets providing detailed dimensions of all the panels, were saved on the ProjectWise platform for easy access by the manufacturers.

2.6. IOT-AIDED CONSTRUCTION

Figure 10. QR code "sensor" and user interface of the app

One major challenge was to use the BIM model to create a user manual, to replace the traditional hard-copy drawing output, so that the work of the designers and engineers could be clearly grasped by the on-site workers, thus reducing possible misunderstandings and the loss of data in the transition to construction. We attempted to introduce IoT into the BIM, using mobile applications, to create a basic user manual of the steel structures for each worker. IoT is focused on connecting physical objects in the real world with devices through data. The idea was to track each structural component in the real world, connect each component to its digital representation in BIM, and then create an assembly simulation, allowing each worker on the site to track and understand the process of certain components.

To connect the components, first, we created a unique identifier (UID) for each structural component and connected each UID to the digital model. Then, we created a QR code for each UID, which acted as our "sensor." Those QR codes were sent to the factory via the BIM platform. Immediately after each component was produced, the corresponding QR code was attached to it (Figure 10). These sensors did not continuously send and receive data over a wireless network; instead, they were passive sensors to which location and historical data were updated each time it was scanned.

A mobile application was created that could review a partial BIM model of the exterior steel structure. When the user selected a component in the model, the app would send them the information for that component—the properties of the component, the tracking data from the UID, the physical location of the object, and its estimated time of arrival. Conversely, scanning the QR code in this application gave the user a digital representation of the corresponding physical component, so that the user could identify its relative position in the overall model and would understand when it needed to be assembled. We also created assembly simulations showing how each component should be assembled into the integrated structural system. The simulation was embedded inside the application.

3. Achievement and conclusion

Figure 11. Completion of the table tennis stadium

This paper presents a practical application of BIM to answer the question: how to conduct a workflow that could carry information from the design to the construction with much less information loss? A complete methodology was developed for data sharing by stakeholders throughout the BIM process. Data-based technologies were integrated into this workflow. In this paper, we show that machine learning, IoT, and CAM were practical applications for the design, manufacturing, and construction of the exterior envelope of the project. The final project shows very little deviation from the initial design. Information was shared seamlessly and efficiently through the BIM process from design to construction. Integrating the different data-based technologies led to significant savings in costs, labor, and time. Using the machine-learning application on the brass panels saved approximately 1.2 million dollars by standardizing the panels to 85 types from more than 5,000 center panels. (Figure 11). Using an IoT system along with other technologies for construction site management and human resources management aided the transportation and installation of the steel structures. The final construction period was reduced from ten months to eight months.

The project has limitations. For example, the IoT application sensor could not update the information about objects continuously, and the IoT app eventually served as a supplement to the drawings, even though it had been designed to replace those. There are benefits and drawbacks, but it remains a very innovative project and a bold approach to demonstrate the strengths of the BIM process. This project serves as a good reference for future practitioners and researchers seeking to conduct data-based BIM studies on real-world projects and a valuable case study for BIM education.

References

Al-Assaf, N.S. and Clayton, M. (2017). Representing the Aesthetics of Richard Meier's Houses Using Building Information Modeling. In: The 37th Annual Conference of the Association for Computer Aided Design in Architecture, ACADIA 2017 (pp. 62–71), https://doi.org/10.52842/conf.acadia.2017.062.

Al-Assaf, N.S. and Clayton, M. (2021). The Use of Diagrammatic Reasoning to Aid Conceptual Design in Building Information Modeling (BIM). In: The 39th Education and Research in Computer Aided Architectural Design in Europe, eCAADe 2021 (pp. 39–48), https://doi.org/10.52842/conf.ecaade.2021.2.039.

Albassel, M. and Waly, M. (2021). Applying Machine Learning to Enhance the Implementation of Egyptian Fire & Life Safety Code in Mega Projects. In: The 9th International Conference of the ASCAAD 2021 (pp. 7–22).

Alymani, Abdulrahman & Jabi, Wassim & Corcoran, Padraig (2020). Machine Learning Methods for Clustering Architectural Precedents Classifying the relationship between building and ground. In: The 38th Education and Research in Computer Aided Architectural Design in Europe, eCAADe 2020 (pp. 643-652), https://doi.org/10.52842/conf.ecaade.2020.1.643.

Borrmann, A., König, M., Koch, C., and Beetz, J. (2018). Building Information Modeling: Why? What? How? In: Borrmann, A., König, M., Koch, C., Beetz, J. (eds), Building Information Modeling. Springer, Cham. https://doi.org/10.1007/978-3-319-92862-3_1

Chaszar, A. and Beirã, J.N. (2013). Feature Recognition and Clustering for Urban Modelling, Exploration and Analysis in GIS and CAD. In: 18th International Conference on Computer-Aided Architectural Design Research in Asia: Intelligent and Informed, CAADRIA 2013 (pp. 601–610), https://doi.org/10.52842/conf.caadria.2013.601.

Chernick, A., Morse, C., London, S., Li, T., Ménard, D., Cerone, J., and Pasquarelli, G. (2021). On-Site BIM-Enabled Augmented Reality for Construction. In: The 2020 DigitalFUTURES. CDRF 2020, DOI:10.1007/978-981-33-4400-6_5.

Cui, Q., Zhang, S., and Huang, Y. (2022). Retail Commercial Space Clustering Based on Post-carbon Era Context: A Case Study of Shanghai, In: 27th International Conference on CAADRIA 2022 (pp. 515–524), https://doi.org/10.52842/conf.caadria.2022.1.515.

Kereshmeh A., Eastman, C., and Shelde, D. (2016). Cloud-Based BIM Data Transmission: Current Status and Challenges, In: 33rd International Symposium on Automation and Robotics in Construction, ISARC 2016 (pp. 1073–1080), https://doi.org/10.22260/ISARC2016/0129.

Meng, L.L., Graham, J., and Haeusler, M. (2020). t-SNE: A Dimensionality Reduction Tool for Design Data Visualisation. In: 25th International Conference on Computer-Aided Architectural Design Research in Asia (pp. 629–638), https://doi.org/10.52842/conf.caadria.2020.2.629.

Pottmann, H., Asperl, A., Hofer, M., and Kilian, A. (2007). Architectural Geometry. Bentley Institute.

Wu, Z., Zhang, Q., Wang, T., Fang, B., and Zhou, K. (2021). Geometrical Theory and Methods in the Parametric Design of Bending and Twisting Skew Grid Structures. Journal of Information Technology in Civil Engineering and Architecture 13(3), 138–147, doi: 10.16670/j.cnki.cn11-5823/tu.2021.03.21.

CROSS-DISCIPLINARY SEMANTIC BUILDING FINGERPRINTS

Knowledge Graphs To Store Topological Building Information Derived From Semantic Building Models (Bim) To Apply Methods Of Artificial Intelligence (Ai) Throughout The Life Cycle Of Buildings

CHRISTOPH LANGENHAN[1], JESSICA BIELSKI[2], CHRISTOPH ZIEGLER[3], VIKTOR EISENSTADT[4], KLAUS-DIETER ALTHOFF[5] and ANDREAS DENGEL[6]
[1,2,3]*Technical University Munich.*
[1]*langenhan@tum.de, 0000-0002-6922-2707*
[2]*jessica.bielski@tum.de, 0000-0003-4936-1993*
[3]*c.ziegler@tum.de, 0000-0001-8095-416X*
[4,5,6]*German Research Center for Artificial Intelligence/ University of Hildesheim/ University of Kaiserslautern*
[4]*viktor.eisenstadt@dfki.de, 0000-0001-6567-0943*
[5]*klaus-dieter.althoff@dfki.de, 0000-0002-7330-6540*
[6]*andreas.dengel@dfki.de, 0000-0002-6100-8255*

Abstract. The advancing digitalization in the building sector with the possibility to store and retrieve large amounts of data has the potential to digitally support planners with extensive design and construction information. Large amounts of semi-structured three-dimensional geometric data of buildings are usually available today, but the topological relationships are rarely explicitly described and thus not directly usable with computational methods of AI. To this end, we propose methods for indexing spatial configurations inspired by the similarity analysis of incomplete human fingerprints, since the early design stage of architectural design is characterized by incomplete information. For this, the topology of spatial configurations is extracted from Building Information Modelling (BIM) data and represented as graphs. In this paper, Semantic Building Fingerprints (SBFs) and Semantic Urban Fingerprints (SUFs), as well as use cases for AI methods are described.

Keywords. Conceptual Design, Building Information Modelling, Knowledge Graph, Artificial Intelligence

1. Introduction

The ongoing digitization in the building industry, with the ability to store and retrieve large amounts of data, has the potential to digitally support planners with extensive design and construction information, what we call knowledge supported design. Negroponte's (1970) vision of the architecture machine "that can follow your design methodology" (Negroponte, 1970) is explored against the background of current developments in Artificial Intelligence (AI) in the building industry focusing on the entire life cycle of buildings.

The building industry plays an important role for the further development of our society in environmental, economic, social and cultural terms. People spend more than 80% of their lives in buildings and in the next 30 years, housing and infrastructure will have to be built in urban areas for more than 2 billion people. In Germany alone, about 400,000 new homes are required every year. In addition, demographic change necessitates new building typologies. According to the Federal Statistical Office, at least 21 million people in Germany will be at least 67 years old by 2039.

To address these challenges with digital technologies, our metis projects funded by German Research Foundation (DFG) aim to exploit the full potential of today's technologies for the building industry. Our current DFG-funded project (Fig. 1) aims on combining different technologies to implement auto-completion for building designs, similar to the auto-completion of words and texts smartphone. We combine semantic technologies like BIM and AI approaches like case-based reasoning (CBR) and deep learning (DL).

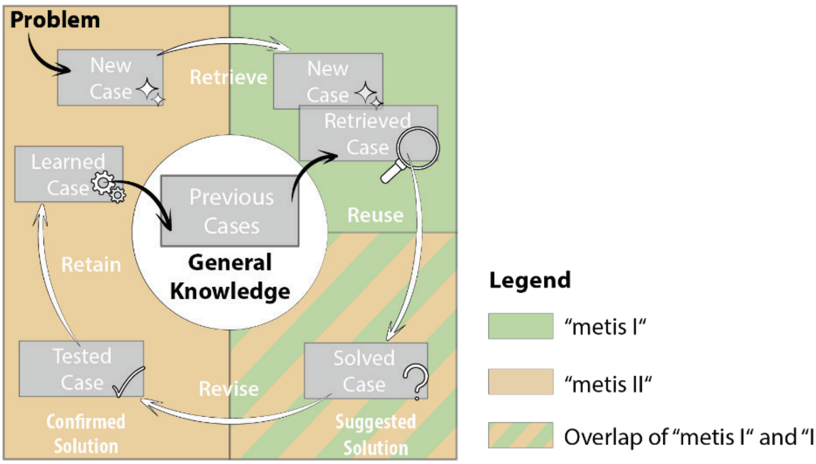


Figure 1. Case-Based Reasoning and the metis projects accordingly.

2. Problem description

From today's perspective, the formalization of complex buildings in digital infrastructures has not been sufficiently solved. There is a data acquisition bottleneck in the building sector. For formalizing building information, object-oriented approaches have been applied to the building sector in the 1990s. However, the semantics of current commercial and open model standards are not strong enough for the application of most contemporary digital methods like machine learning (ML). There is a need for improving the methods to index design knowledge and to formalize this knowledge.

For computer-based support of cross-lifecycle knowledge management in the building industry, we propose Semantic Building Fingerprints (SBFs) (e.g. for early design phases as shown in Fig. 2) as topological data, formalized in knowledge graphs. The paradigm of the SBFs enables us to measure the similarity of buildings and building parts based on their structure, comparable to the identification of human beings by their fingerprints (Langenhan, 2013). The SBFs in Figure 2 are two examples to formalize buildings in the early design phase in terms of accessibility and adjacency. However, already in the consecutive design planning phase material and construction have significant influence on the room arrangement and the whole design.

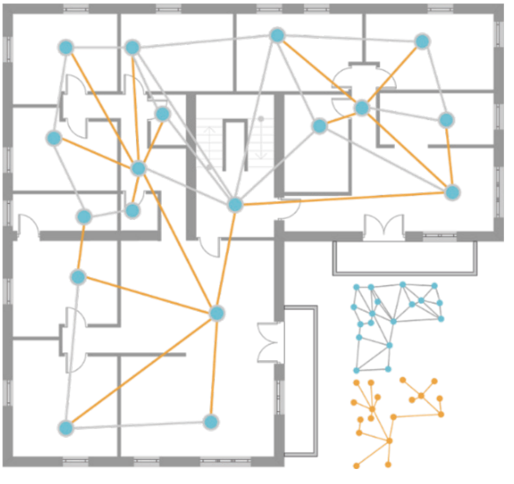


Figure 2. SBFs of accessibility and adjacency in early design phases.

Therefore, a small selection of further Semantic Building Fingerprints (SBFs) is presented in this paper as examples to enable knowledge management with current AI methods for the life cycle of buildings. The SBFs for the different phases of the life cycle can be processed separately and used across phases by linking them. For example, the space at the bottom left or the respective blue graph-node in Figure 2 can be linked to SBFs for materials and their separability or the adjacent buildings and urban spaces.

In the following, an overview of the approaches of Artificial Intelligence (AI) and in particular Case-Based Reasoning (CBR) and Deep Learning (DL) are presented in order to use Semantic Building Fingerprints (SBFs) and Semantic Urban Fingerprints (SUFs) for knowledge management in the construction industry. Further, semantic technologies from the building industry and the associated possibilities of formalizing spatial information are described.

3. Related work

The theory of the first generation of the design methodology movement (Richter, 2010) is based on the assumption that design problems are well-defined and thus exhaustively describable. Building on the systems theory approaches, the first computer-based approaches were developed focusing on design rules (Purcell et al., 1990), but could not exhaustively formalize the complexity of architectural design. Therefore, the second generation of the design methodology movement in the 1970s, represented by Horst Rittel, among others, understood design not procedurally as the fulfilment of requirements, but as an individual process that can only be incompletely described (Richter, 2010). Based on that, mainly Case-Based Reasoning (CBR) approaches of the 1980s influenced architectural design, which led to AI research in architecture, called Case-Based Design (CBD), in the 1990s.

CBD projects such as ARCHIE (Pearce et al., 1992), PRECEDENTS (Oxman, 1994), IDIOM (Smith et al., 1995), CADRE (Hua et al., 1996), FABEL (Voss, 1997) or CaseBook (Inanc, 2000) investigated principles and approaches for case-based support in architectural designing. However, most of these projects focused only on one of the phases of the CBR cycle (e.g. Retrieve or Reuse) and semantic information was only partially considered. In some projects (e.g. ARCHIE and CaseBook), the methods for explaining system processes (e.g. a 'lessons-learned' text or a similarity explanation report) were additionally investigated. Since the 2000s, research began to focus on fully comprehensive and more supportive approaches, e.g. the CBArch project (Cavieres et al., 2011).

The research field of CBR in general is currently increasingly turning to the methods of AI research for sub-processes. Contemporary approaches from the field of AI rely on techniques of DL with an artificial neural network (ANN) for this purpose. An ANN is a dense network of nodes (neurons) that passes the weighted numerical information from the input layer to the output layer through a large number of hidden layers. The goal of the ANN is to find the most likely output (e.g. continuation of a sentence) based on the input (e.g. the first words in a sentence). The learning techniques of ANNs have already been used in the field of architecture (Kazanasmaz et al., 2009; Mozer, 1998). The DANIEL (Deep Architecture for fiNdIng alikE Layouts) approach (Sharma et al., 2017) aims to analyse and search building floor plans, based on DL or ANN approaches. However, in other domains, there are already approaches that link CBR and DL to e.g. provide a generic methodology for CBR, answering in support centers (Amin et al., 2018).

Domain-specific data models for materials and constructs as well as data conversion approaches are described by Werbrouck et al. (2020), while there are data models for timber construction (Châteauvieux-Hellwig, et al., 2021) and approaches for linking data models in construction (Esser & Borrmann, 2021; Senthilvel & Beetz, 2020). Furthermore, monographs on building materials and details, such as by Deplazes (2005), Hegger et. al. (2005), Ballast (2009), Ching (2012) and McMorrough (2018) provide important parameters. For a comprehensive cross-life-cycle formalization of Building Information Modelling (BIM) or Geographic Information System (GIS) approaches are combined and linked as a so-called "Digital Twin".

In digital semantic building models, building information can be formalized for different disciplines. Even though ontology-based approaches can be applied to it, it is not a domain-specific ontology in the classic sense. Examples of standard software tools for ontologies include Open Cyc or DOLCE. Exemplary modelling languages are the Entity Relationship Model and UML (Unified Modelling Language), as well as older languages such as EXPRESS and EXPRESS-G for the ISO standard STEP (ISO 1030) for the exchange of product models (e.g. IFC data model). For IFC, a STEP and XML data model is maintained by buildingSMART (buildingSMART e.V. 2022).

The aim of digital semantic building models is to support various use cases, such as the organisation of construction sites (Borrmann et al. 2009), scheduling (Hartmann et al. 2007), cost planning (Tulke; Ma et al. 2011) or the evaluation of design decisions and energy consumption (Kimpian et al. 2009). In this context, the information exchange between different use cases and their participants is of particular importance.

For digital semantic building models, the project 'IfcOWL' (Beetz et al. 2009) proposes an approach to transform the IFC EXPRESS scheme into the Ontology Web Language (OWL) in order to use standard algorithms for OWL to request building information. As a technology, Beetz et al. (2007) propose the Resource Description Framework (RDF). Therefore, the RDF can be queried with state of the art technologies for the retrieval of specific building information, such as SPARQL, which is impossible in the IFC data model.

4. Methodology

To this end, we have derived information on spatial relations from BIM data as knowledge graphs to build a case base. We have trained deep neural networks with this data converted into e.g. tensor-based data structure, as a so-called "relation map" (Eisenstadt et al., 2021a). The information is derived from BIM data and enriched with further information so that the building data can be processed as a knowledge graph and represented e.g. as one-hot encoding vectors (see Fig. 3), as well as be used as tensor with AI approaches (Eisenstadt et al., 2021b). Our approaches contribute to the formalization and linking of information across the lifecycle of a building and its application in the planning context.

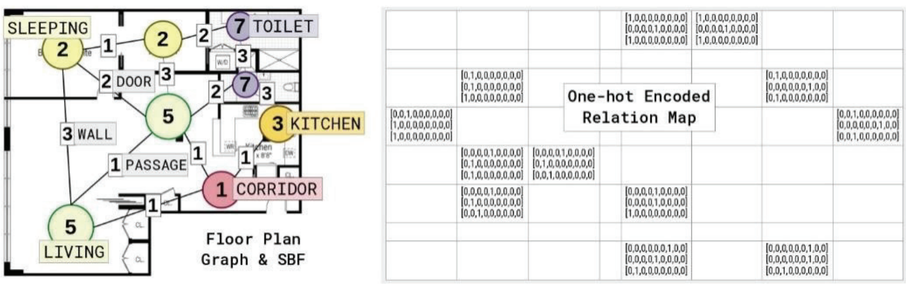


Figure 3. A SBF graph converted into a one-hot-vector encoded map (Eisenstadt et al., 2021a).

In addition to the specification of models and the use of suitable technologies, the transfer of formalizsed information from one data model to another is important for the usability of the information. Depending For example, depending on the focus of the specification, certain data models do not sufficiently support some parts of information and , resulting in information loss during derivation, or necessary automatic or manual post-hoc additions. For data management we propose a federated information system, in which different autonomous information sources (BIM server, graph database and content-management system (CMS) can be queried with our common XML specification, we call AgraphML. We use our AgraphML specification in all frontend and backend prototypes for our Semantic Building Fingerprints (SBFs).

In the specification of AgraphML for early design stages, each space is mapped to a node and the relation of the spaces to each other as edges in a graph (see Fig. 2). The specification is an extension of the XML-based graphml specification and can be processed with numerous standard tools, like Gephi, or our own 'Dolphin' plugin for Grasshopper3D (Langenhan, 2021) used for evaluating our approaches in research related teaching projects. The plugin enables us to transform BIM data into the AgraphML format. Table 1 shows the taxonomies of the node and edge labels for housing.

Table 1: Taxonomies of node and edge labels for the formalization of building information.

Node label	Edge label
ROOM	DOOR
KITCHEN	ENTRANCE
LIVING	PASSAGE
SLEEPING	SLAB
WORKING	STAIRS
CORRIDOR	WALL
TOILET	WINDOW
BATH	
EXTERIOR	
STORAGE	
BUILDINGSERVICES	
CHILDREN	

In addition to the taxonomy for housing in Table 1, we have developed taxonomies for other building typologies, such as school construction. Especially for the methods from graph theory, the labels for the nodes and edges are very important for the reduction of computational costs of the NP-complete problem. The goal is to have as many labels as possible to reduce the computational costs without overlapping or contradictory semantics of the labels.

Our approaches aim to contribute to the formalization and linking of information across the lifecycle of a building and its application in the wider planning context. Therefore, we have developed other Semantic Building Fingerprints (SBFs) and Semantic Urban Fingerprints (SUFs) to support the life cycle of buildings beyond the early design stages, as well as SBFs for material and their constructive connections, and for building elements and their connections (see Fig. 4).

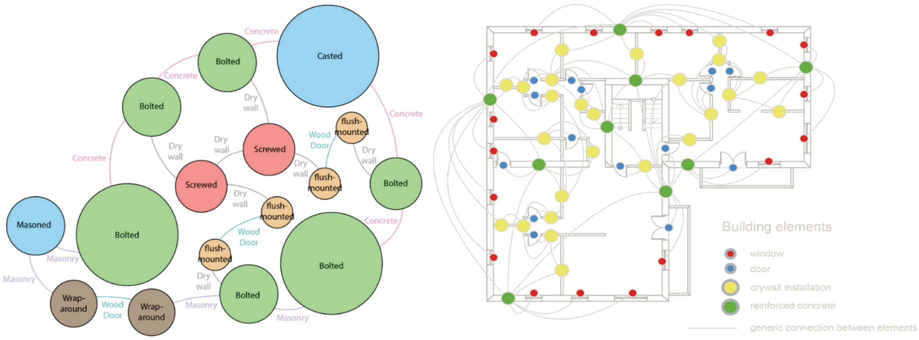


Figure 4. SBFs of the materials and connections (left) and building elements (right).

The left SBF in Figure 4 illustrates how building materials can be represented as edges, while their connection is shown as directed nodes, to provide a representation without displaying the geometry. Therefore, we divide the node types into three main categories: 'wall to wall', 'door to wall' and 'window to wall'. These are also individually sub-divided by material category, e.g. for 'wall to wall': 'concrete to timber', as well as vice versa 'timber to concrete', depending on the direction of the connection. The right SBF in Figure 4 is projected onto a floor plan. In contrast to the left SBF, building elements are represented, forming nodes, but the connecting edge types are undefined.

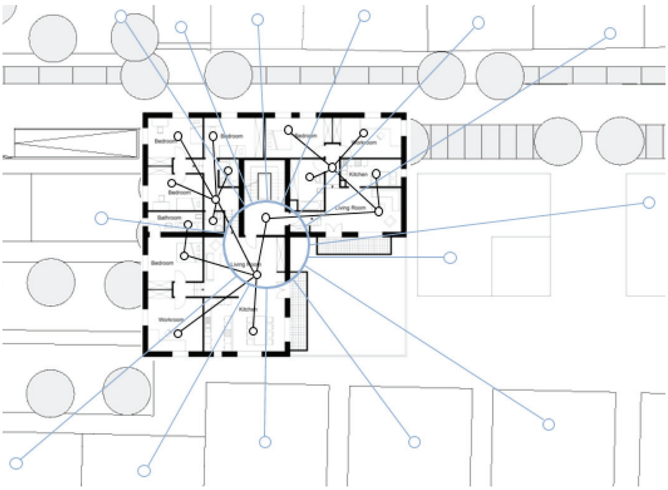


Figure 5. SUF of the urban integration representing buildings and their connection (blue).

In addition to SBFs for buildings, we suggest SUFs for urban planning (Fig. 5). For this, GIS data is analyzed, e.g. with SpatialHadoop, and patterns are derived. These patterns are the basis for the development of SUFs to formalize the relation of objects from GIS into graphs e.g. with attributes like angles, sizes, collinearity, perpendicularity or parallelism.

The SUF shown in Figure 5 is a supplement of the SBF from Figure 2. The knowledge graph of the neighbouring buildings shown in blue represent the relationships of the buildings on an urban scale. The spatial nodes of the SBFs are linked to the building nodes of the SUF, thus linking BIM and GIS data.

5. Conclusion, discussion and future work

Our research so far focuses on supporting the early design stages by formalizing housing construction in knowledge graphs in order to build a case base with these graphs and train deep neural networks. The results of combining Case-Based Reasoning (CBR) and Deep Learning (DL) approaches are very promising and encourage us to explore the transfer of our methods to later design stages, and urban design and planning. The Semantic Building Fingerprints (SBFs) and Semantic Urban Fingerprints (SUFs) in Figure 4 and Figure 5 are initial considerations of how broad planning support can work. The SBFs and SUFs are going to be processed separately in our federated information system and connected by linked open data approaches.

As a result, we derived approaches for the formalization of information in the building industry across the life cycle as knowledge graphs and the conversion for the use with CBR and DL. The four types of knowledge graphs presented are strong digital semantic models that focus topological information that can be used with contemporary methods, e.g. from the field of AI, to support the design method of planners. These four knowledge graphs are selected examples of design information to illustrate the potential for the building industry.

The initial evaluations of the implementation of the AI methods confirm the suitability of the Recurrent neural networks (RNNs) for Analysis-Synthesis-Evaluation (ASE) design phase detection and Graph Neural Networks (GNNs) for link prediction, respectively, both achieving values above 90% correctness (Mete et al., 2022). Explainability methods complementing the AI-based auto-completion approach were evaluated in a case study. The methods were positively received by the participants for facilitating well-informed and confident design decisions through providing additional information and different design solutions (Bielski et al., 2022).

In view of the current climate crisis, digital modelling of the relation of building components in construction is of particular importance, since well separable materials and components can be better disposed of, appropriately recycled and, above all, reused for example. Our work contributes to knowledge management in the construction sector by firstly developing data structures to store design and construction information, secondly processing this information and finally using the information as knowledge in the design and construction process.

Our implementations have shown that neural networks can be trained with our case based building information. For a more comprehensive formalization, much larger amounts of knowledge graphs are required and the corresponding data for training needs to be obtained.

Acknowledgements

We want to thank the German Research Foundation (DFG) for funding the metis projects.

References

Amin, K., Kapetanakis, S., Althoff, KD., Dengel, A., Petridis, M. (2018). Answering with Cases: A CBR Approach to Deep Learning. Cox M., Funk P., Begum S. (eds) Case-Based Reasoning Research and Development. *ICCBR 2018. Lecture Notes in Computer Science, vol 11156.* Springer, Cham. https://doi.org/10.1007/978-3-030-01081-2_2.

Ballast, D.K. (2009). Architects handbook of construction detailing. *John Wiley & Sons, Inc*.

Beetz, J.; de Vries, B.; van Leeuwen, J. (2007). RDF-BASED Distributed Functional Part Specifications for the Facilitation of Service-based Architectures. 22nd CIB W78 Conference on Information Technology in Construction, 183–188.

Beetz, J.; Leeuwen, J. van; Vries, B. de (2009). IfcOWL: A case of transforming EXPRESS schemas into ontologies. Artificial Intelligence for Engineering Design, Analysis and Manufacturing (23), 89–101, https://doi.org/10.1017/S0890060409000122.

Bielski, J., Langenhan, C., Ziegler, C., Eisenstadt, V., Petzold, F., Dengel, A., & Althoff, K.-D. (2022). The What Why, What-If and How-To for Designing Architecture - Explainability for auto-completion of computer-aided architectural design of floor plan layouting during the early design stages. International Conference of the Association for Computer-Aided Architectural Design Research in Asia, 2(27th), 435–444, https://doi.org/10.52842/conf.caadria.2022.2.435.

Borrmann, A.; Ji, Y.; Obergrießer, M.; Rank, E.; Klaubert, C.; Günthner, W. (2009). ForBAU. The virtual construction site project. 26th nCIB-W78 Conference on Managing IT in Construction.

buildingSMART e.V. (2022). buildingSMART. Online: https://www.buildingsmart.de/.

Cavieres, A., Bhatia, U., Joshi, P., Zhao, F., & Ram, A. (2011). CBARCH: A Case-based Reasoning Framework for Conceptual Design of Commercial Buildings. AAAI Spring Symposium Series, 19-25.

Ching, F. (2012). Building Construction Illustrated. John Wiley & Sons, Inc.

Châteauvieux-Hellwig, C., Abualdenien, J., Borrmann, A. (2021). Towards semantic enrichment of early-design timber models for noise and vibration analysis. eWork and eBusiness in Architecture, Engineering and Construction: 13th European Conference on Product and Process Modelling (ECPPM 2020 - 2021). Moscow, Russia.

Deplazes, A. ed. (2005). Constructing Architecture – Materials Processes Structures. Birkhäuser.

Eisenstadt, V., Arora, H., Ziegler, C., Bielski, J., Langenhan, C., Althoff, K.-D., Dengel, A., & Petzold, F. (2021a). Exploring optimal ways to represent topological and spatial features of building designs in deep learning methods and applications for architecture. Proceedings of Conference on Computer Aided Architectural Design Research in Asia (CAADRIA), 1, 191–200.

Eisenstadt, V., Arora, Hardik, Ziegler, C., Bielski, J., Langenhan, C., Althoff, K.-D., & Dengel, A. (2021b). Comparative Evaluation of Tensor-based Data Representations for Deep Learning Methods in Architecture. In V. Stojakovic & B. Tepavcevic (Eds.), Education and research in Computer Aided Architectural Design in Europe Conference. University of Novi Sad, 1, 45-54, https://doi.org/10.52842/conf.ecaade.2021.1.045.

Esser, S., Borrmann, A. (2021). A system architecture ensuring consistency among distributed, heterogeneous information models for civil infrastructure projects. *13th European Conference on Product and Process Modelling*. Moscow, Russia.

Hartmann, T.; Goodrich, W.; Fischer, M.; Eberhard, D. (2007). Fulton Street Transit Center Project. 3D/4D Model Application Report. *TR170*. Stanford University.

Hegger, M., Auch-Schwelk V., Fuchs M., Rosenkranz T. (2005). Baustoff Atlas. *Detail*.

Hua, K., Fairings, B., & Smith, I. (1996). CADRE: Case-based Geometric Design. *Artificial Intelligence in Engineering 10(2)*. 171-183, https://doi.org/10.1016/0954-1810(95)00026-7.

Inanc, B. S. (2000). CASEBOOK. An Information Retrieval System for Housing Floor Plans. *Proceedings of 5th Conference on Computer Aided Architectural Design Research (CAADRIA), 389-398*.

Kazanasmaz, T., Murat, G., and Selcen, B. (2009). Artificial Neural Networks to Predict Daylight Illuminance in Office Buildings. *Building and Environment 44.8 (2009)*, 1751-1757. https://doi.org/10.1016/j.buildenv.2008.11.012.

Kimpian, J.; Mason, J.; Coenders, J.; Jestico, D.; Watts, S. (2009). Sustainably Tall. Investment, Energy, Life Cycle. ACADIA 09. building a better tomorrow: proceedings of the 29th annual conference of the Association for Computer Aided Design in Architecture (ACADIA). United States, 130-143.

Langenhan, C., Weber, M., Liwicki, M., Petzold, F., & Dengel, A. (2013). Graph-based retrieval of building information models for supporting the early design stages. Advanced Engineering Informatics, 27(4), 413–426, https://doi.org/10.1016/j.aei.2013.04.005.

Langenhan, C. (2021, January 29). Dolphin. Grasshopper3D. https://www.grasshopper3d.com/group/dolphin

McMorrough, J. (2018). The Architecture Reference & Specification Book updated & revised. Rockport Publishers.

Mete, B., Bielski, J., Eisenstadt, V., Langenhan, C., Petzold, F., & Althoff, K.-D. (2022). Predicting semantic building information (BIM) with Recurrent Neural Networks. European Conference on Product and Process Modeling, (14th).

Mozer, M. C. (1998). The Neural Network House: An Environment that Adapts to its Inhabitants. AAAI Spring Symp. Intelligent Environments. Vol. 58, 110-114.

Negroponte, N. (1970). The Architecture Machine: Toward a more human environment. Cambridge. Mass: MIT Pr, 1970.

Oxman, R.E. (1994). Precedents in Design: A Computational Model for the Organization of Precedent Knowledge. *Design studies, 15(2)*. 141-157, https://doi.org/10.1016/0142-694X(94)90021-3.

Pearce, M., Goel, A. K., Kolodner, I. L., Zimring, C., Sentosa, L., & Billington, R. (1992). Case-Based Design Support: A Case Study in Architectural Design. *IEEE expert*, 7(5), 14-20, https://doi.org/10.1109/64.163668.

Purcell, P. A., Mitchell, W. J., McCullough, M. (1990). The Electronic Design Studio: Knowledge and Media in the Computer Era. *Cambridge, Mass: MIT Press.*

Sharma, D., Gupta, N., Chattopadhyay, C., & Mehta, S. (2017). DANIEL: A Deep Architecture for Automatic Analysis and Retrieval of Building Floor Plans. *Document Analysis and Recognition (ICDAR), 2017 14th IAPR International Conference on Pattern Recognition,* 1, 420-425, https://doi.org/10.1109/ICDAR.2017.76.

Tulke, J. (2011), Kollaborative Terminplanung auf Basis von Bauwerksinformationsmodellen. *Dissertation*. Weimar.

Richter, K. (2010). Augmenting Designers' Memory: Case Based Reasoning in der Architektur. Dissertation. *Logos-Verlag*.

Senthilvel M., Beetz, J. (2020). A Visual Programming Approach for Validating Linked Building Data. *EG-ICE 2020 Workshop on Intelligent Computing in Engineering. pp. 403-411.* Berlin.

Smith, I., Lottaz, C., & Faltings, B. (1995, October). Spatial Composition Using Cases: IDIOM. *International Conference on Case-Based Reasoning (pp. 88-97)*. Springer, Berlin, Heidelberg.

Voss, A. (1997). CASE DESIGN SPECIALISTS IN FABEL. *Issues and applications of case-based reasoning in design*. pp. 11-38.

Werbrouck, J., Pauwels, P., Bonduel, M., Beetz, J., Bekers, W. (2020). Scan-to-graph: Semantic enrichment of existing building geometry. *Automation in construction 119, Seiten/Artikel-Nr.:103286.* https://doi.org/10.1016/j.autcon.2020.103286.

SYNTHESIZING STYLE-SIMILAR RESIDENTIAL FACADE FROM SEMANTIC LABELING ACCORDING TO THE USER-PROVIDED EXAMPLE

JIAXIN ZHANG[1, 2]; TOMOHIRO FUKUDA[2]; NOBUYOSHI YABUKI[2]; and YUNQIN LI[1]
[1] Architecture and design college, Nanchang University, No. 999, Xuefu Avenue, Honggutan New District, Nanchang 330031, China
[1]{jiaxin.arch|liyunqin}@ncu.edu.cn, 0000-0002-6330-6723, 0000-0002-1886-0477
[2] Division of Sustainable Energy and Environmental Engineering, Graduate School of Engineering, Osaka University, Osaka 5650871, Japan
[2] fukuda.tomohiro.see.eng@osaka-u.ac.jp, 0000-0002-4271-4445, yabuki@see.eng.osaka-u.ac.jp, 0000-0002-2944-4540

Abstract. Example-guided facade synthesis aims to synthesize realistic facade images from semantic labels drawn by architects and example images of user preferences. The automated synthesis approach allows for the efficient generation of facade solutions that will facilitate effective communication between stakeholders and creative inspiration for architects. This study proposes a conditional generation adversarial network with style consistency to solve the problem of example-guided image synthesis. Specifically, the synthesis model is divided into two stages: first, the domain of the semantic label map is transferred to the domain of the realistic image using the pix2pixHD framework to ensure that the synthesized facade in the intermediate stage can be semantically consistent with the designed facade; Second, we use the Deep Photo Style Transfer (DPST) framework to faithfully move the implied features of the realistic facade image synthesized in the previous step to the domain of the provided example to ensure consistency of style. In summary, the proposed method can constrain the synthesis of new residential facades from the semantic labels and example styles. The synthesized residential facades can be consistent with the example styles provided by the client while matching the semantic labels of the facade created by the designer, producing satisfyingly realistic transitions in various cases.

Keywords. Residential Facades, Style Transfer, Image Synthesis, Generative Adversarial Networks, Building Facade Design

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 139-148.

1. Introduction

In architectural design projects, the client often has a preferred building facade style, while the designer must consider the functional and formal requirements of the proposal from a professional perspective. Suppose at the pre-proposal stage, both the user-provided examples and the designer-created sketches with semantic labels can be used as guides to synthesize the building facade images automatically. In that case, it will effectively facilitate the communication efficiency between stakeholders and save the designer's conceptual effort in the draft. Figure 1 illustrates how the proposed methodology can generate a preliminary facade plan to meet the expectations of the architect and the client in their communications.

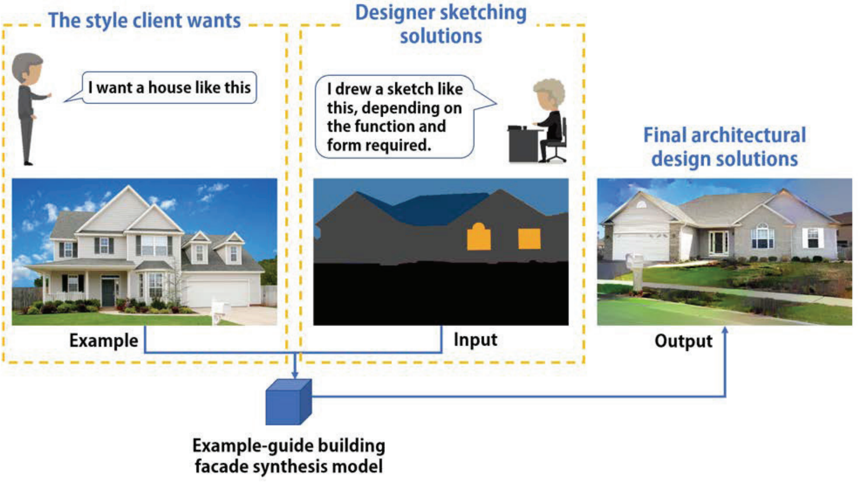


Figure 1. Automatic synthesizing facade images from user-provided examples and designer-created sketches as guides

A priori knowledge-based facade generation method can effectively solve the synthesis problem for limited style samples relying on effective knowledge discovery by designers for specific facade styles (Tang et al., 2019). Recently, data-driven computer vision (CV) methods have proven to be robust in automatically synthesizing building facade images with less reliance on expert know-how guidance (Sun et al., 2022), providing architects with creative solutions. Architects have tried to combine deep learning approaches to the facade synthesis issues. For example, Sun et al. (2022) used pix2pix to generate building facades for historical urban renovation automatically. Zhang et al. (2021) developed a method based on adversarial neural networks to automatically fill missing parts of building facades. Meng (2022) used the StyleGAN2 model to generate images of building facades by analyzing latent spaces without conditional inputs. However, all these methods have challenges in example-guided style-consistent facade image synthesis. The translation of content differences between architectural sketches and reference images is a challenge and may lead to undesirable generation results between unrelated content. The semantic accuracy and transfer faithfulness of each element of the facade is also challenging for synthesizing real-world images. The synthesized facade appearance should match visual experiences. Furthermore, supervised learning-based approaches suffer from a lack of datasets dedicated to diverse building facade styles.

This study attempts to develop a lightweight, example-guided residential architectural facade synthesis tool to address the above issues. The tool can synthesize an image of a residential facade with the tendency of the client's preferences from semantic labeled maps and exemplar images indicating styles. We use the term "style" in this context to refer to the implied characteristics of a facade image, for instance: from an ethnic-geographical perspective, Western, Eastern, From a historical trend, classical architectural style, and modernist style. In these cases, semantic labeling maps denote the segmentation of elements such as windows, doors, roofs, and facades of residences. We propose a two-stage approach and a dataset. We first use an image translation-based framework to fix the semantics of the facade label drawings drawn by the architects. The image translation model generates an intermediate stage facade that visually matches the appearance of real-world buildings. Secondly, we introduce a style transfer model that faithfully transfers the reference style and is used to transfer the real-world domain facade image from the previous stage to the user-provided case style. Furthermore, a paired residential facade dataset containing building facade images (from different styles, periods, and stories) and corresponding facade semantic labels is proposed.

2. Related works

2.1. BUILDING FACADE GENERATION

Typically, facade design requires considering the location of windows and doors in the plan, user-preferred materials, and harmony with the surrounding environment. As a result, building facade design is a laborious task for architects. Several automated facade design studies have recently been proposed to contribute significantly to the conventional design workflow (Cao et al., 2017), freeing the architects' labor. Established studies include both a priori knowledge-based and data-driven approaches. These works show that mechanical architectural design independent of architects' intuition can significantly contribute to the conventional design workflow. Data-driven neural network-based facade generation methods have yielded promising results, including the fidelity of the generated images and the feasibility of style transfer.

2.2. IMAGE SYNTHESIS VIA GAN

A generative adversarial network (GAN) is a neural network structure consisting of a generator and a discriminator. It learns the training set's feature distribution and generates new images. Many GAN variants have been developed to cover different application scenarios. This paper deals with GAN-based image generation tasks, including image translation and style transfer.

Isola et al. (2018) proposed a conditional GAN framework for image-to-image translation tasks with paired images as supervision. We aim to synthesize photographs of real domains using semantic label maps. The pix2pixHD (Wang et al., 2018) is an upgraded version of pix2pix and is a robust framework for image synthesis and interactive manipulation for the generation of large-size facades.

The purpose of style transfer is to transfer the style of the source image to the target image or domain. Established studies include single-example-based and holistic sense-based style transfer. Luan et al. (2017) proposed the Deep Photo Style Transfer (DPST) model, which is locally affine in the color space by constraining the transformation from input to output. The model can relocate the facade style of the original domain to the target domain, where the user provides an example.

3. Methods and datasets

3.1. SYNTHESIS OF STYLE-GUIDED BUILDING FACADE IMAGE

Figure 2 shows the overview of the proposed method. The E provided by the user can guide the synthesis of the real-world image domain y. We aim to synthesize a final facade image y from the semantics labeled map x and the exemplar E: $(x, E) \rightarrow y$. The role of x is to anchor the semantics of the synthetic image. The role of E is to provide stylistic constraints for image synthesis, and the output image y must be consistent with the style of the exemplar E. The two particular requirements we face are: (1) While the semantics of the synthetic facade image is consistent with the semantics of the labeled facade image x, it should satisfy the human visual perception of the realistic facade. We use an image translation model (pix2pixHD) to solve this supervised problem to learn the semantic transformation between paired images (semantic label maps and real-world images). Ensures that the semantics of the generated facade drawings of the real-world domain are consistent with the input label map. (2) Given an input labeled map x, the absence of the ground truth of the bootstrap style exemplar $\{E\}$ should not affect the synthesis result. To solve this weakly supervised problem, we introduce a model of DPST. It constrains the local transformation of the input to the output in color space and expresses this constraint as a fully differentiable energy term.

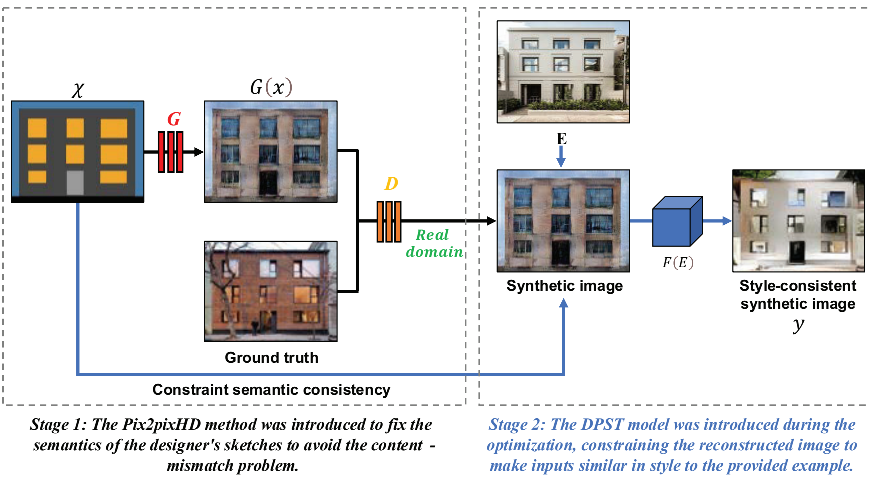


Figure 2. The overview of the proposed method.

The proposed approach builds on the single-scale version of pix2pixHD and DPST, containing (1) a generator G to semantic map x, ground truth image, and a standard discriminator D to distinguish the generated graph as real or fake. (2) A staged synthetic image of the real-world domain, a style-guided example E, and a style-consistent synthetic result. (3) A style-consistent synthesizer DPST, which is used to discover whether the synthetic image and the guide image E are style-compatible. Here, $F(E)$ is an operational procedure that produces a set of synthetic images y that are stylistically consistent with the real-world domain images. Our objective function contains two losses: a standard adversarial loss for synthesizing semantically consistent images of the real domain and an adversarial style consistency loss.

3.2. RESIDENTIAL FACADE DATASET FOR SEMANTICALLY CONSISTENT SYNTHESIS

To synthesize semantically consistent real-world images of residential facades based on image translation, we built a dataset of paired semantically labeled data with real-world residential facade images, which we call Residential Facade for Semantically Consistent Synthesis (RFSCS). RFSCS has 612 residential facades from different countries and eras with corresponding semantic labels. The semantic labels include walls, roofs, windows, the sky, greenery, and road surface in the foreground. Figure 3 shows examples of the paired semantic labels and building facades of RFSCS. Making paired datasets allows designers to draw facade semantic maps that correspond well in abstraction to real-world facade images. The unsupervised approach does not require paired data and is suitable for image transformation between objects where sample correspondence is complex. In contrast to unsupervised learning, supervised learning image translation models trained with RFSCS can accurately and faithfully synthesize the facade elements designed by the architects in line with the semantics.

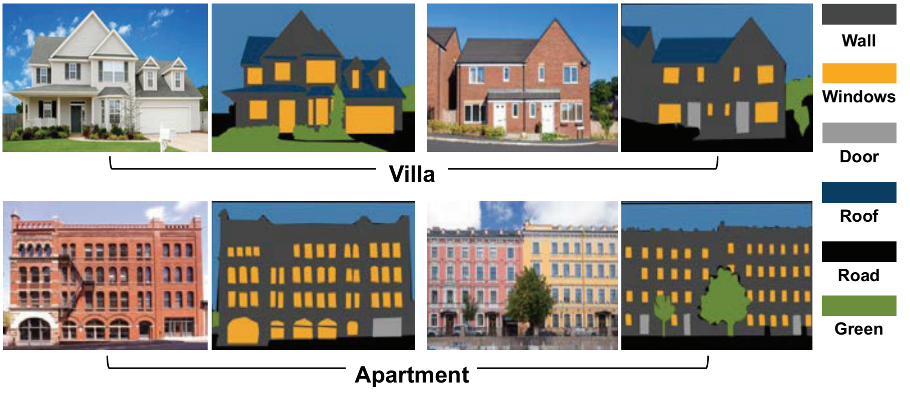


Figure 3. Paired residential facade dataset for semantically consistent synthesis.

4. Experimental results

4.1. SYNTHESIS OF REAL-WORLD DOMAINS OF BUILDING FACADE BASED ON IMAGE TRANSLATION

The training set consists of 612 paired sets of building facades, and we resize all collected images to 1000 × 500. We employ two image processing methods based on the original dataset to extend the number of datasets. As shown in Figure 4, the data augmentation methods are horizontal flip and 6◦ counterclockwise rotation. The final extended dataset has 1836 pairs of images, of which 1560 pairs are set as the training set and the remaining 276 pairs (only 92 pairs for the original data and the remaining 184 pairs as the enhanced data) as the test set. The training was performed on two Nvidia Geforce 1080 Ti Graphics Cards with batch size set to 4 and total epoch to 300. During the training process, we recorded the loss values for the generator and the discriminator. Figure 5 shows the losses for both models. Since the two neural networks compete with each other, the loss of the discriminator increases as the loss of the generator decreases. The performance of the entire training model gradually improves as the generator and discriminator ebb and flow.

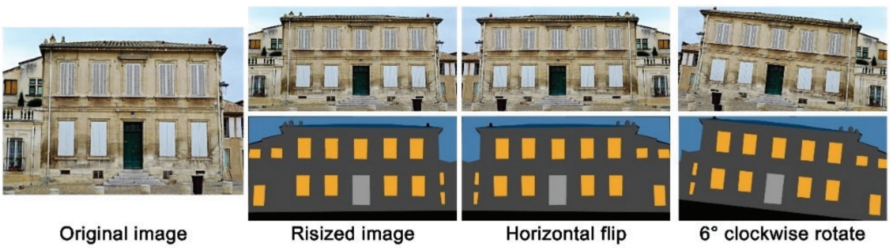


Figure 4. Data processing and augmentation.

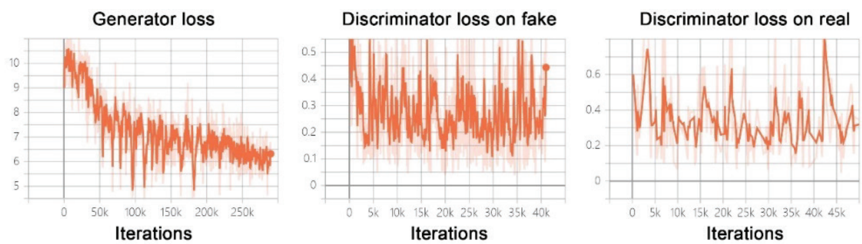


Figure 5. Training Loss. Left is generator loss. The middle is discriminator loss on fake images. Right is discriminator loss on real images.

We used the designer's manual drawings of facade labels, including walls, windows, doors, roads, greenery, and sky, as input for our testing. We experimented with flat roofs, sloped roofs, single-story villas, and multi-storey apartments. The characteristics of residential facades include having similar components and structures, their relatively fixed semantics, and a wide variety of styles. The simple semantic composition of residence facilitates experts to draw semantic labels, while non-experts can also try various preferred styles. In generating real-world domains, we take the residential building as the main object of study.

Figure 6 shows the building facades generated by the pre-trained model pix2pixHD, and the semantics of each facade of the generated results are consistent with the appearance of real-world facades. The process focuses on faithfully transferring the semantic labels of facades to the domain of real-world facades with the help of an image translation model, locating the latent space of residential facades as photorealistic as possible.

Figure 6. Building facades generated by pix2pixHD.

4.2. SYNTHESIS RESULTS OF STYLE-SIMILAR FACADES BASED ON THE PROVIDED EXAMPLES

Figure 7 shows the results of the generated synthetic facades in four different styles. A semantic labeled map drawn by the designer is used as the scheme of the facade. Four examples provided by the client are used as the target style domains. The synthesis process is based on the initial design of a small villa with cases of western and eastern, as well as wood and brick walls. As seen from the results in Figure 7, the color transformation with its same element semantics is localized, and the proposed method can handle context-sensitive color changes. The proposed method applies the exact semantic color mapping to match the color statistics between the input and style-guided images. As a result, the synthesized facade can faithfully respond to the color and material information of the provided example.

Figure 7. Facade synthesis results from semantic labeling and provided examples.

We also try to reveal the image translation process from semantic to real-world domain generation. Figure 8 shows the semantic labeling from the input facade to the target domain facade and then synthesizing a stylistically consistent facade image based on the provided examples. The results show that the first stage of real-world domain facade synthesis affects the final synthesis. When semantic synthesis can be generated with high-quality and seamless borders, the photorealistic facade images can be synthesized according to the style provided in the bootstrap example.

Figure 8. Synthesize stylistically consistent facades using semantic labels and provided examples.

5. Discussion and Conclusion

The proposed method synthesizes facades that meet the designer's requirements for facade semantics, size, and form. Simultaneously, the synthesized facades can also meet the user's requirements for the solution style. The proposed model uses a multi-scale architecture that can synthesise high-resolution images (up to 2048 × 1024 resolution). Our contributions can be summarized in the following three points. (1) The proposed workflow can successfully synthesize facades consistent with the style of the provided examples, with minimal impact on the faithfulness of the transformation. (2) The proposed method solves the problem of the content difference between the input image and the reference image. For example, in the synthesis of the building facade, the result of the synthesis of each element is consistent with people's empirical judgments of the real world. The strategy we adopt is to customize a dataset of facade image translations for supervised learning. The semantic labels of the facades are matched to real-world facades to minimize the chance of inaccurate transfers. (3) When the transfer occurs between semantically equivalent sub-regions, the semantic labels of the input and stylized images are incorporated into the transfer procedure. Besides, the mapping is nearly uniform as the transfer occurs in each sub-region, which helps preserve the richness of synthetic styles.

While the merits are appreciated, several limitations deserve further study.

Synthesizing non-orthogonal projections of building facades can be unsatisfactory because of the limited number of angled datasets. In terms of metrics for evaluating the quality of synthesized facade images, common metrics are not always indicative of the perceived realism of an image, and a human evaluation may provide a more comprehensive understanding of the synthesized image quality. We have demonstrated the feasibility of synthesizing new facades by changing the semantic labels of the old facades. This application will hopefully be used in building facade renewal, where architects redesign the semantic labels of old facades, non-experts provide examples of interest, and the framework will generate new facade solutions that meet the stakeholders' requirements. In addition, the recent excellent performance of the Diffusion Model on image translation (Saharia et al., 2022) will also inspire this work. In future work, we will try to adopt the Diffusion Model framework to realize the example-guided facade style synthesis.

References

Cao, J., Metzmacher, H., O'Donnell, J., Frisch, J., Bazjanac, V., Kobbelt, L., & van Treeck, C. (2017). Facade geometry generation from low-resolution aerial photographs for building energy modeling. *Building and Environment*, 123, 601–624. https://doi.org/10.1016/j.buildenv.2017.07.018

Isola, P., Zhu, J.-Y., Zhou, T., & Efros, A. A. (2018). Image-to-Image Translation with Conditional Adversarial Networks (arXiv:1611.07004). arXiv. https://doi.org/10.48550/arXiv.1611.07004.

Luan, F., Paris, S., Shechtman, E., & Bala, K. (2017). Deep Photo Style Transfer. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 6997–7005. https://doi.org/10.1109/CVPR.2017.740

Meng, S. (2022). Exploring in the Latent Space of Design: A Method of Plausible Building Facades Images Generation, Properties Control and Model Explanation Base on StyleGAN2. In P. F. Yuan, H. Chai, C. Yan, & N. Leach (Eds.), *Proceedings of the 2021 DigitalFUTURES (pp. 55–68). Springer Singapore.* https://doi.org/10.1007/978-981-16-5983-6_6

Saharia, C., Chan, W., Chang, H., Lee, C., Ho, J., Salimans, T., Fleet, D., & Norouzi, M. (2022). Palette: Image-to-Image Diffusion Models. *Special Interest Group on Computer Graphics and Interactive Techniques Conference Proceedings*, 1–10. https://doi.org/10.1145/3528233.3530757

Sun, C., Zhou, Y., & Han, Y. (2022). Automatic generation of architecture facade for historical urban renovation using generative adversarial network. *Building and Environment*, 212, 108781. https://doi.org/10.1016/j.buildenv.2022.108781

Tang, P., Wang, X., & Shi, X. (2019). Generative design method of the facade of traditional architecture and settlement based on knowledge discovery and digital generation: A case study of Gunanjie Street in China. *International Journal of Architectural Heritage*, 13(5), 679–690. https://doi.org/10.1080/15583058.2018.1463415

Wang, T.-C., Liu, M.-Y., Zhu, J.-Y., Tao, A., Kautz, J., & Catanzaro, B. (2018). High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 8798–8807. https://doi.org/10.1109/CVPR.2018.00917

Zhang, J., Fukuda, T., & Yabuki, N. (2021). Automatic Object Removal With Obstructed Facades Completion Using Semantic Segmentation and Generative Adversarial Inpainting. *IEEE Access*, 9, 117486–117495. https://doi.org/10.1109/ACCESS.2021.3106124

VARIABILITY IN MACHINE LEARNING FOR MULTI-CRITERIA PERFORMANCE ANALYSIS

VICTOR OKHOYA[1] AND MARCELO BERNAL[2]
[1,2]*Perkins&Will Architects.*
[1]*victor.okhoya@perkinswill.com*
[2]*marcelo.bernal@perkinswill.com*

Abstract. Parametric analysis is emerging as an important approach to building performance evaluation in architectural practice. Since architectural performance has many competing metrics multi-criteria analysis is required to deal effectively with the complexity. However, multi-criteria parametric analysis involves large design spaces that are expensive to compute. Machine learning is emerging as an important design space reduction method for multi-criteria analysis. However, there are many types of machine learning algorithms and architects can benefit from understanding which algorithms perform well on which tasks. Using a mid-rise commercial residential tower project this paper investigates three common machine learning algorithms for performance against three common performance metrics. The algorithms are multi-layer perceptrons, support vector machines, and random forests, while the metrics are site energy, illuminance, and a value function that combines them both. In addition, we seek to understand what factors are most impactful in improving algorithm performance. We investigate four impact factors namely sample size, sensitivity analysis, feature selection, and hyperparameters. We find that multi-layer perceptrons perform best for all three performance metrics. We also find that hyperparameter tuning is the most impactful factor affecting multi-layer perceptron performance.

Keywords. Parametric Analysis, Machine Learning, Design Space

1. Introduction

Parametric analysis is a broad and generic term to describe a computational approach to problem-solving where a range of input parameters are related and varied. Evins (2013) presented a review of applying parametric analysis to sustainable building design problems. Machairas et al. (2014) reviewed methods for parametric building design optimization. Bernal et al. (2020) compare them against design heuristics. Nguyen et al. (2014) reviewed simulation-based parametric analysis methods focusing on simulation programs, optimization tools, the efficiency of optimization methods, and industry trends.

In attempting to implement parametric analysis in architectural practice several data

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 149-158.

analytical challenges are encountered from early on (Glover & Greenberg, 1989). Some of these can be characterized as having to do with the size of design spaces, the speed of dynamic simulation, and the accuracy of design space reduction (Keogh & Mueen, 2017). Given the relentless pace of current architectural design processes, it is desirable for all design-related and design-supporting activities to be executed as rapidly as possible. Unfortunately, large design spaces take a long time to simulate (Liu et al., 2018). If simulation time is improved by reducing the size of the design space, then questions regarding accuracy arise.

Machine learning algorithms are emerging as a viable approach for solving the parametric analysis simulation bottleneck (Robinson et al., 2017). Not only are they speeding up the process by predicting rather than simulating results (Kazanasmaz et al., 2009; Wei et al., 2018), but they are also proving versatile in their application to different analysis criteria and metrics. There are, however, many types of machine learning algorithms on offer. To effectively use these algorithms in practice, architects need to know which algorithms perform well on which types of problems. In addition, they need to understand what factors impact algorithm performance so they can adjust these factors to extract optimal outcomes.

In this paper, we will investigate the question: what is the best machine learning algorithm for architectural parametric performance analysis? We will perform a comparative analysis between three commonly used algorithms: multi-layer perceptron (MLP), support vector machine (SVM), and random forest (RF). We will investigate the performance of these algorithms against two common metrics: daylight and energy (Magoules & Zhao, 2016) using a real-world commercial residential project. We also investigate the performance of a value function that combines daylight and energy. In addition, we will investigate the impact of four factors on the performance of the algorithms namely sample size, hyperparameter tuning, feature selection, and sensitivity analysis.

2. Methodology

2.1. CASE STUDY

The case study consisted of three towers from 14 to 18 levels in height in a growing, transit-oriented, affordable housing community. It had two separate buildings on two adjacent sites with a total site area of 63,000 sqft. Two of the towers sat on a connected podium while the third sat on the podium of an adjacent sire (Figure 1). The designers proposed a range of accommodations including retail units, market housing units, co-op housing units, townhouses, underground parking, rooftop and outdoor amenity spaces, and an upper-level courtyard.

Figure 1. Mixed use Case Study

2.2. SIMULATION PROCESS

The metric for energy analysis w as Site Energy, while the metric for daylight analysis was Illuminance. Due to the complexity of the project, the designers proposed ten slices to capture all the different adjacency conditions of the design. These were the roof, which had an exposed upper surface; the typical floor, which was sandwiched between two typical floors and the podium adjacent floor, which shared a surface with the podium itself (Figure 2). In practice, due to different program requirements at the podium level (such as retail and commercial spaces requiring substantial mechanical systems and equipment), the process focused only on the eight non-podium slices.

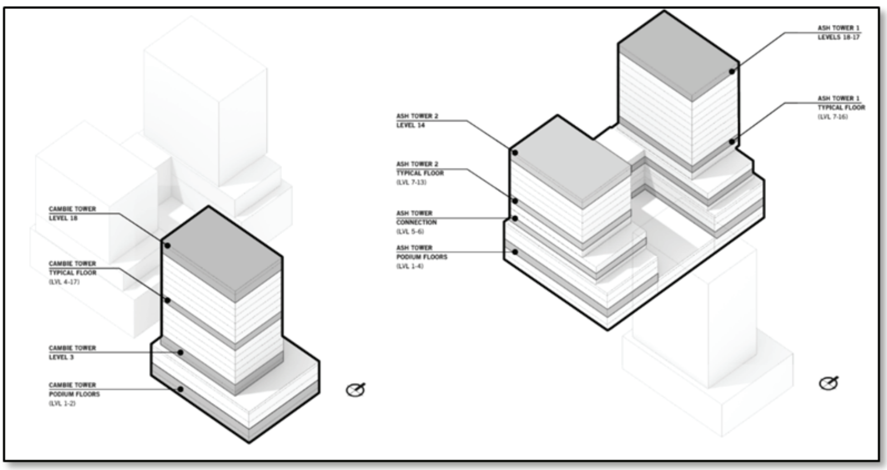


Figure 2. Slices of the three towers.

The design space of the analysis is shown in Table 1, and for each slice there was a design space size of 1296 instances for a total of 10,368 simulations. The zoning was related to the project slicing and suite layouts (Figure 3).

Parameter	Description	Range	Size
WWR North	Window to wall ratio	25,35	2
WWR West	Window to wall ratio	45,55,65	3
WWR East	Window to wall ratio	25,35	2
WWR South	Window to wall ratio	45,55,65	3
Wall_U_Value	Wall U Value	0.379,0.568,0.811,1.136	4
Chi_Value	Point heat transmittance coefficient	0.48	1
Window_U_Value	Window U Value	1.42,1.7,1.99	3
SHGC	Solar Heat Gain Coefficient	0.3,0.35,0.4	3
Slice_ID	Project slice identifier	1-8	8
			10,368

Table 1. Design Space

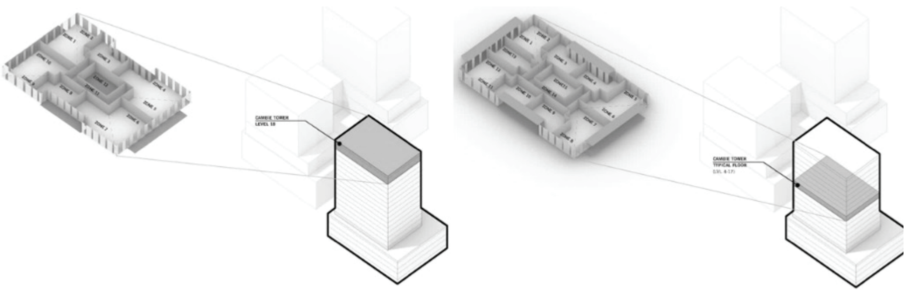


Figure 3. Zoning approach for two slices of the project.

In addition to zoning, other parametric analysis inputs included glazing information such as SHGC values; envelope information such as Wall U-Value; lighting information such as lighting controls; equipment information such as equipment loads; design conditions such as thermostat set-back temperatures, and air system information such as heat recovery system. Zone inputs included program type, lighting power density, occupancy, metabolic rate, ventilation, and equipment.

2.3. MACHINE LEARNING PROCESS

The Weka Experimenter (Waikato University, 2022) was the machine learning interface. It allowed for the upload of several training sets, based on actual simulations, and then the selection of several algorithms to run on the training sets. Three popular algorithms were used for the investigation: Multilayer Perceptron, an artificial neural network; SMOReg, a regression-based implementation of support vector machines; and Random Forest, an implementation of decision tree type of algorithms.

The Weka Experimenter also allowed the user to select suitable experimental and algorithm settings. Sample size, sensitivity analysis, and hyperparameter tuning experiments used the Weka Experimenter. Feature Selection used the Weka Explorer environment with the *WrapperSubsetEval* algorithm. Experimental settings were 10-fold cross-validation with 10 repetitions. The repetitions used a random seed to ensure the 10-fold cross-validation used different folds for each repetition. This means that each training set iterated 100 times.

2.4. EVALUATION METHODS

We considered four algorithm performance factors: sample size, sensitivity analysis, feature selection, and hyperparameter tuning. These performance factors will also be referred to as impact factors. The algorithm performance metric we used was the Relative Ab solute Error (Cichosz, 2014) or RAE.

For the sample size, we ran all three algorithms on all three metrics for sample sizes ranging from 100 to 1000 instances. The goal was to discover which algorithms had a small RAE when using a small sample size. For sensitivity analysis, we used the Access Variable Importance to compute the sensitivity of each input parameter for each metric. For all three metrics, the bu ilding location where performance was analyzed, Slice_ID was the most sensitive parameter (Figure 4) We ranked all the input parameters in order of sensitivity and then ran machine learning experiments for each algorithm on all metrics varying the number of sensitive parameters from 1 to 8. The goal was to discover if a few sensitive parameters could perform as well or better than all parameters taken together.

For feature selection, we used a wrapper method from the Weka Attribute Selection interface. Wrapper methods involve two aspects: evaluation and search. The evaluation was done using the WrapperSubsetEval algorithm. Search was done using the BestFirst algorithm with the backward method. The goal was to establish if any algorithm performed significantly better for any of the metrics after feature selection.

For hyperparameter tuning each algorithm's hyperparameters were varied for each metric in the Weka Experimenter. The goal was to establish if variations in any hyperparameter significantly improved algorithm performance for any metric. The hyperparameters identified for each algorithm together with their range of values are shown in Figure 5.

Sensitivity Analysis, Energy

Summary Report

Column	Main Effect	Total Effect
Slice	0.544	0.563
Wall_U_Val	0.28	0.299
Window_U_Val	0.049	0.068
SHGC	0.014	0.024
WWR_W	0.014	0.023
WWR_E	0.008	0.016
WWR_S	0.006	0.01
WWR_N	0.001	0.003

Sensitivity Analysis, Illuminance

Summary Report

Column	Main Effect	Total Effect
Slice	0.863	0.879
WWR_N	0.079	0.095
WWR_S	0.013	0.021
WWR_E	0.001	0.003
SHGC	4e-4	0.001
Wall_U_Val	2e-4	3e-4
WWR_W	9e-5	2e-4
Window_U_Val	4e-5	1e-4

Sensitivity Analysis, Value Function

Summary Report

Column	Main Effect	Total Effect
Slice	0.742	0.775
Wall_U_Val	0.123	0.156
Window_U_Val	0.015	0.034
WWR_N	0.011	0.026
SHGC	0.008	0.019
WWR_W	0.005	0.012
WWR_E	0.001	0.004
WWR_S	2e-4	0.001

Figure 4. Sensitivity analysis for three output metrics.

Since the investigation was multi-criteria, we needed a method for comparing how the algorithms perform on both energy and daylight metrics combined. We did this by using a weighted value function. In this investigation, both the daylight and energy metrics are given a weight of 50%. We calculated the value function by normalizing the daylight and energy metrics, inverting the energy metric by subtracting its normalized value from one, then multiplying the two metrics by their weight and summing them into a single value.

3. Results

The impact of sample size on all three metrics was that larger sample sizes performed better than small sample sizes. On average all algorithms performed better with a sample size of 1000 than with a sample size of 100 by 10.84% for the energy metric, by 6.51% for the illuminance metric, and by 9.67% for the value function metric. On average MLP with a large sample size performed 8.72% better than RF and 5.26% better than SMOReg. SMOReg performed 3.46% better than RF.

The impact of sensitivity analysis was that all algorithms performed better with all parameters than they did with just some sensitive parameters. For the energy metric the best-performing algorithm was MLP with 8 parameters (RAE 4.96%) while the worst was MLP with 1 parameter (RAE 77.95%). For the illuminance metric, the best-performing algorithm was RF with 8 parameters (RAE 13.1%) while the worst-performing algorithm was MLP with 1 parameter (RAE 47.27%). For the value function metric, the best-performing algorithm was MLP with 8 parameters (RAE 14.76%) while the worst-performing was MLP with 1 parameter (RAE 63.27%).

The best features for each algorithm are shown in Table 2. The change in RAE due to feature selection compared to default with no modifications for each algorithm is shown in Figure 6. This suggests that feature selection did not have a significant impact on algorithm performance.

MLP hyperparameters

Hyperparameter	Description
hiddenLayers	(10,10), (20,20), (30,30), (40,40), (50,50)
learnRate	0.001, 0.01, 0.1, 0.2, 0.3
momentum	0.2, 0.4, 0.6, 0.8

SMOReg hyperparameters

Hyperparameter	Description
complexity	0.001, 0.01, 0.1, 1,10
kernel	Polykernel, NormalizedPolyKernel, RBFkernel
epsilonParameter	0.001, 0.01, 0.1,1

Random Forest hyperparameters

Hyperparameter	Description
numIterations	100, 300, 500, 700
numFeatures	2, 4, 6, 8
maxDepth	50, 100, 150, 200

Figure 5. Hyperparameters for the three algorithms.

For the MLP algorithm the most impactful hyperparameter for energy was *hiddenLayers* while the most impactful hyperparameter for illuminance and the Value Function was *learningRate*. On average hyperparameter tuning for MLP improved performance by 1.9%. For the SMOReg algorithm, the most impactful hyperparameter for energy, illuminance, and the value function was the *kernel*. On average hyperparameter tuning for SMOReg improved performance by 5.54%. For the RF algorithm, the most impactful hyperparameter for energy and illuminance was *numFeatures*. The most impactful hyperparameter for the value function was *iterations*. On average hyperparameter tuning for RF improved performance by 0.66%.

ALGORITHM	OBJECTIVE	BEST FEATURES	RAE

MLP	**Energy**	Slice_ID, WWR_N, WWR_W, WWR_S, WWR_E, Wall_U_Value, Window_U_Value, SHGC	5.08
	Illuminance	Slice_ID, WWR_N, WWR_W, WWR_S, WWR_E, SHGC	17.5
	Value Function	Slice_ID, WWR_N, WWR_W, WWR_S, WWR_E, Wall_U_Value, Window_U_Value, SHGC	14.99
SMOReg	**Energy**	Slice_ID, WWR_N, WWR_W, WWR_S, WWR_E, Wall_U_Value, Window_U_Value, SHGC	11.36
	Illuminance	Slice_ID, WWR_N, WWR_S, WWR_E, Wall_U_Value	23.72
	Value Function	Slice, WWR_N, WWR_W, WWR_S, WWR_E, Wall_U_Value, Window_U_Value, SHGC	19.13
RF	**Energy**	Slice_ID, WWR_N, WWR_W, WWR_S, WWR_E, Wall_U_Value, Window_U_Value, SHGC	18.22
	Illuminance	Slice_ID, WWR_N, WWR_W, WWR_S, WWR_E, Window_U_Value	12.55
	Value Function	Slice_ID, WWR_N, WWR_W, WWR_S, WWR_E, Wall_U_Value, Window_U_Value, SHGC	18.05

Table 2. Feature selection

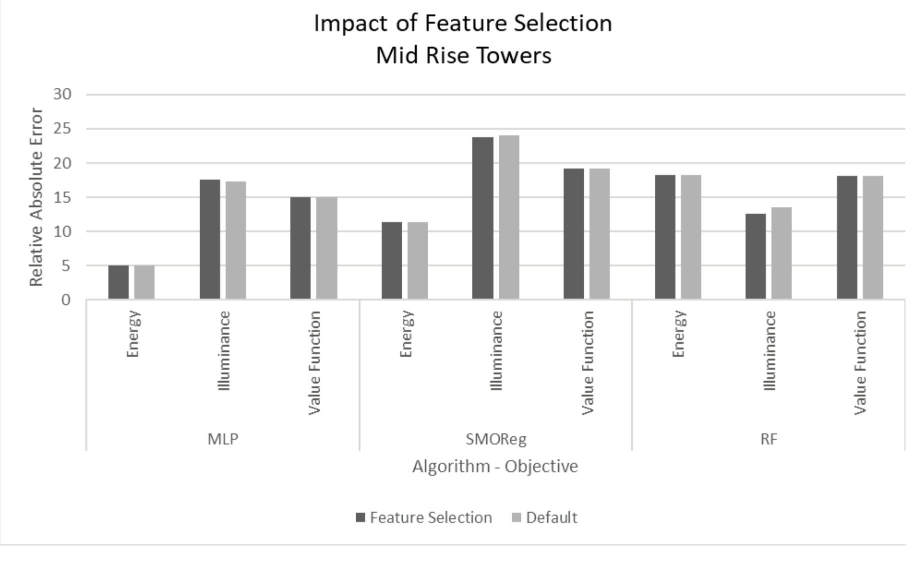


Figure 6. Feature selection compared to the default.

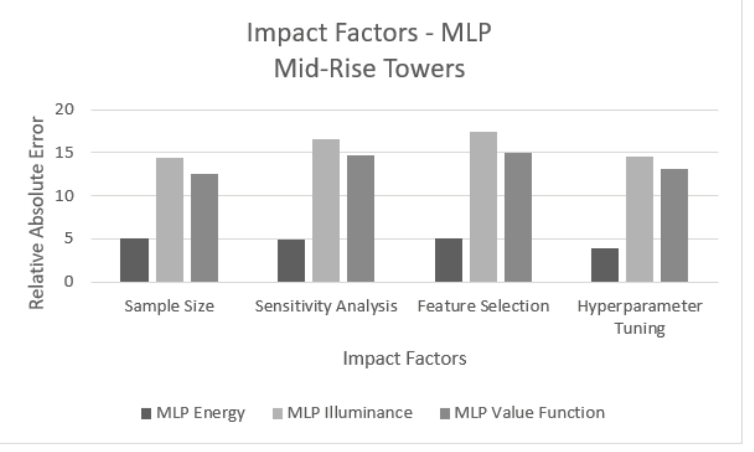


Figure 7. Summary of impact factors for MLP algorithm.

4. Discussion

The best performing algorithm for all metrics was MLP. MLP had an RAE of 3.91% for the energy metric, 14.62% for the illuminance metric, and 12.59% for the value function metric. MLP performed 4.84% better than SMOReg and 13.34% better than RF for energy; it performed 8.52% better than SMOReg and 1.92% worse than RF for illuminance; and it performed 5.88% better than SMOReg and 5.35% better than RF for the value function. The best-performing metric for MLP was energy. On average, for MLP, energy performed 6.66% better than illuminance and 9.48% better than the value function. Illuminance was 2.83% better than the value function. The most important impact factor for MLP was hyperparameter tuning which had an average performance of 10.57% across all metrics compared to 10.72% for increasing the sample size, 12.52% for feature selection, and 14.33% for sensitivity analysis (Figure 7) This investigation involved a single project type and only examined three popular algorithms. It however provides some useful initial information for architects seeking to take advantage of machine learning algorithms for parametric analysis. In addition, it provides a possible framework for approaching the testing and tuning of machine learning algorithms for different performance metrics. Future work will involve algorithm performance tests on a broader range of project types. In addition, a broader range of machine learning algorithms can also be tested.

References

Ahmad M, Hippolyte , J., Mourshed, M. & Rezgui, Y., (2017). Random Forests and Artificial Neural Network for Predicting Daylight Illuminance and Energy Consumption. s.l., s.n., pp. 7-9.

Bernal, M. et al., (2020). Integrating expertise and parametric analysis for a data-driven decision-making practice. International Journal of Architectural Computing, 4(18), pp. 424-440, https://doi.org/10.1177/1478077120940975.

Cichosz, P., (2014). Data Mining Algorithms explained using R.. s.l.:John Wiley & Sons.

Evins, R., (2013). A review of computational optimization methods applied to sustainable building design.. Renewable and Sustainable Energy Reviews, pp. 230-245, https://doi.org/10.1016/j.rser.2013.02.004.

Glover, F. & Greenberg, H., (1989). New approaches for heuristic search: A bilateral linkage with artificial intelligence. European Journal of Operational Research, pp. 119-130, https://doi.org/10.1016/0377-2217(89)90185-9.

Kazanasmaz, T., Gunaydin, M. & Binol, S., (2009). Artificial neural networks to predict daylight illuminance in office buildings. Building and Environment, pp. 1751-1757, https://doi.org/10.1016/j.buildenv.2008.11.012.

Keogh, E. & Mueen, A., (2017). Curse of Dimensionality. In: Encyclopedia of Machine Learning and Data Mining. Boston: Springer, pp. 314-315.

Liu, Y., Colburn, A. & Mehlika, I., (2018). Computing Long-Term Daylighting Simulations from High Dynamic Range Imagery using Deep Neural Networks. s.l., s.n.

Machairas, V., Tsangrassoulis, A. & Axarli, K., (2014). Algorithms for optimization of building design: A review. Renewable and Sustainable Energy Reviews, pp. 101-112, https://doi.org/10.1016/j.rser.2013.11.036.

Magoules, F. & Zhao, H.-X., (2016). Data Mining and Machine Learning in Building Energy Analysis. s.l.:Wiley, DOI:10.1002/9781118577691.

Nguyen, A., Reiter, S. & Rigo, P., (2014). A review on simulation-based optimization methods applied to building performance analysis. Applied Energy, pp. 1043-1058, https://doi.org/10.1016/j.apenergy.2013.08.061.

Perkin&Will, (2019). Image Source.. s.l.:s.n.

Robinson, C. B Dilkina, J Hubbs & W Zhang, 2017. Machine learning approaches for estimating commercial building energy consumption. Applied Energy, pp. 889-904, https://doi.org/10.1016/j.apenergy.2017.09.060.

Waikato University, (2022). Weka (3.8). [Software]. Available at: https://waikato.github.io/weka-wiki/downloading_weka/ [Accessed Nov. 2022]

Wei, Y., Zhang, X., Shi, Y., Xia, L., Pan, S., Wu, J., Han, M. and Zhao, X., (2018). A review of data-driven approaches for prediction and classification of building energy consumption. Renewable and Sustainable Energy Reviews, pp. 1027-1047, https://doi.org/10.1016/j.rser.2017.09.108.

SKYWAYS VERSUS SIDEWALKS

Evaluating the Perceptual Qualities and Environmental Features of Elevated Pedestrian Systems in Hong Kong

SIFAN CHENG[1] and PU JIANG[2] and FRANCESCO ROSSINI[3] and JEROEN VAN AMEEIJDE[4]
[1,2,3,4]*School of Architecture, The Chinese University of Hong Kong*
[1]*1155136989@link.cuhk.edu.hk*
[2]*jiangpu33@gmail.com*
[3]*rossini@cuhk.edu.hk* [4]*jeroen.vanameijde@cuhk.edu.hk*

Abstract. Elevated pedestrian systems can increase urban density and reduce congestion in urban spaces, which contributes to the sustainability of cities. In high-density cities such as Hong Kong, the complex system of underground passages and skywalks at multiple levels is crucial in facilitating efficient pedestrian circulation. While the relationship between street environments and pedestrian experience has been widely discussed, research on how the perceptual qualities of skywalk environments affect pedestrians' walking preferences has not yet been addressed. This study demonstrates how the perceptual environmental features of skyway systems can be characterised in a quantitative and human-centric manner. Taking Central, Hong Kong, as a case study, it has investigated the walking environment of the skyway system and the environmental differences between skyways and sidewalks and has explored the relationships between subjective /objective street qualities and pedestrian's perception. An improved understanding of the environment-behaviour relationships of these spaces can provide suggestions for designing more walkable and comfortable skyway systems that are better integrated in the fabric of the city.

Keywords. Skyways, Walking Environment, Google Street View, Volunteered Geographic Information, Machine Learning.

1. Introduction

In high-density cities such as Hong Kong, the complex system of underground passages and skywalks at multiple levels is crucial in facilitating efficient pedestrian circulation (Rossini et al., 2018). Skywalk systems profoundly impact the public's daily activity structure and well-being (Zacharias, 2013). However, due to the ad-hoc planning principles and design practices for skywalk spaces in Hong Kong (Tan & Xue, 2014), their spatial qualities are perceived as labyrinthine and repetitive (Frampton, 2012). Many studies have emphasised the influence of perceived

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 159-168. © 2023 and published by the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA), Hong Kong.

environmental qualities on pedestrians' experience (Ewing & Cervero, 2010; Ewing & Handy, 2009; Handy et al., 2006) and predicting walking behaviour (Humpel et al., 2004).

Previous research on the environment of skywalks in Hong Kong focuses on its formation, accessibility and economic impact at local and district scales (Woo & Malone-Lee, 2013; Villani & Talamini, 2020). Few studies have discussed the perceived qualities (Frampton, 2012) and walking experience (Kalandides, 2012; Villani & Talamini, 2020) of elevated pedestrian spaces. However, no comprehensively objective exploration of elevated pedestrian systems' design elements and the related perceptual attributes has yet been conducted.

The relationship between street environments and pedestrian experience has been widely discussed (Adkins et al., 2012; Ewing & Handy, 2009; Johansson et al., 2016). Recently, street-level image data, combined with deep learning algorithms, has become a new analytical approach widely used in streetscape perception studies (Ma et al., 2021; Tang & Long, 2019; Zhang et al., 2018). Virtual audits of environmental features using street view images has been broadly applied at city or regional scale. However, prior studies argued that compared with on-site audits, virtual audits may overlook some valuable environmental features which may impact residents' perception. The relationship between subjective/objective measures of environmental features and people's perception is still under-studied, especially on skyways in the context of the compact city.

In this research, a deep learning-based approach using Google Street View (GSV) imagery was used to analyse the environmental features of selected sections of the sky bridges and main sidewalks in the Central district of Hong Kong. The approach audited the five perceptual indices, Enclosure, Imageability, Complexity, Greenness, and Tidiness (Ewing & Handy, 2009). Second, we employed a team of volunteers to walk through the case study area, during which they filled in the Perceived Walkability and Perception Scale. Last, this study explored the correlation between visual qualities and people's perception using Pearson Correlation, aiming to determine the critical perceptual features of skywalks and sidewalks. Various visual elements were identified to be positively or negatively correlated with perceptual indicators through statistical analyses. The analytical framework developed in this study can help evaluate the visual qualities of newly planned elevated walkways in high-density cities worldwide.

2. Methods

2.1. ANALYTICAL FRAMEWORK

The project used a mixed method, qualitative and quantitative approach to document the correlation between the visual quality of skyways/sidewalks and pedestrians' walking perception (Figure 1). A structured study was conducted to collect quantitative data by obtaining GSV images of the skyway system in Central, and we employed SegNet, a machine learning algorithm, to recognize different environmental features. The presence, visual proportion, and diversity of environmental features were used to quantify the visual quality indices. To compare the subjective and objective measurements of environmental features, we recruited several volunteers as observers and collected their walking perceptions through field audits. The qualitative evaluation includes several indices: Physical barriers to walking, Aesthetics, Presence of people, Crowdedness, Traffic and road hazards, social disorder/littering, Crime, and Perceptions. The study combined a short version of an environmental perception evaluation scale (NEWS=CS) and the collection of five perception indices. The results of the visual quality and walking perception analyses were plotted into QGIS with geographic coordinates. Last, this study conducted statistical analysis on 1) the visual difference between skyways and sidewalks; 2) the correlation between subjective/objective environmental features and people's perceptions.

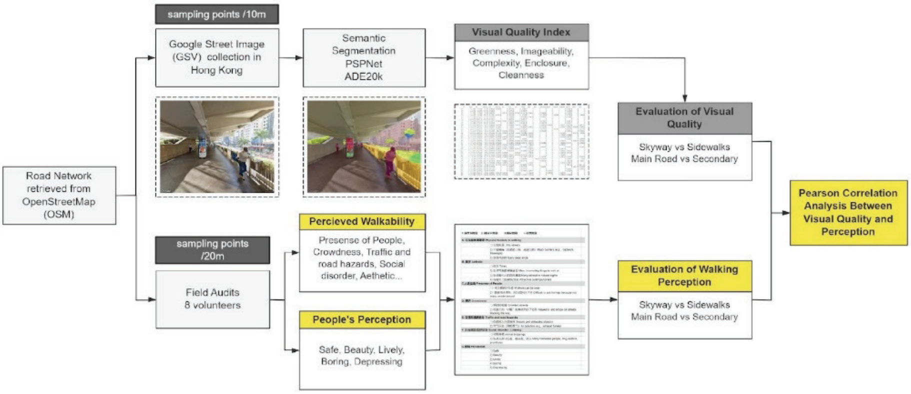


Figure 1. Analytical framework of the study

2.2. STUDY AREA

Our research was conducted within the Central district in Hong Kong, which houses a large network of grade-separated pedestrian routes. The system was initially started by private developers and has gradually expanded through public and private urban development (Tan and Xue 2014), connecting transit stations, shopping malls, office buildings and parks. This study chose several typical skyway and sidewalk segments as pilot areas to test our conceptual evaluation framework (Fig. 2).

2.3. DATA GATHERING

We first extracted road network geospatial information from OpenStreetMap. We sampled points every 10 meters in our pilot study area for virtual and on-site audits, and obtained Google Street View (GSV) images as a virtual audit dataset. GSV images were requested in an HTTP URL form using the GSV Image API (Google, 2014). For our study, a Python script was created to read the coordinates of each collecting point and evaluate the typological features of the surrounding street networks to determine the heading view angles for each collecting point (Fig. 2). This study used a Semantic Segmentation technique, Pyramid Scene Parsing Network (PSPNet), to extract and compute the view index of different elements. PSPNet has been widely used to evaluate streetscape scenarios due to its high pixel-level precision. We used a pre-trained dataset – ADE20K to process the GSV in our research areas.

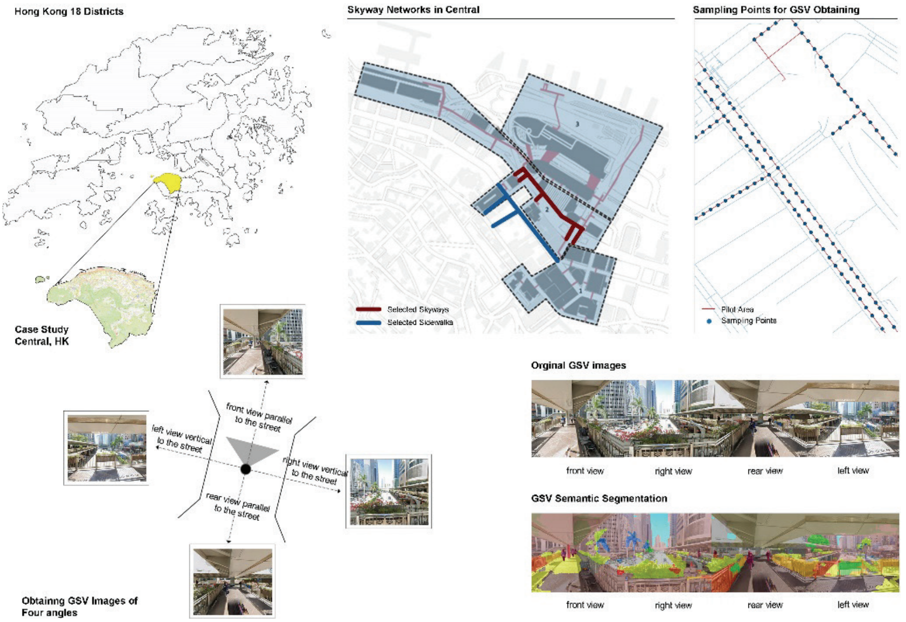


Figure 2. Case study area, sampling methods, and GSV imagery gathering

2.4. OBJECTIVE MEASUREMENT OF PERCEPTUAL QUALITY

Subjective perceptions and physical features may have complicated or subtle relations (Ewing et al., 2006; Ewing & Handy, 2009). Initially, Ewing et al. (2006) objectively measured five subjective urban design qualities from the street environment: imageability, enclosure, human size, transparency, and complexity. It has been argued that the traditional field observations were limited to a single city, involved a small number of participants, and used low-throughput surveys. Recent studies have employed crowdsourced map data and geotagged images to derive fine-scale perceptual quality information of large urban areas. Ma et al. (2021) and Ye et al. (2019) transformed the perceptual quality indices into the subjective evaluation of a combination of different elements. To compare the visual quality of skyways and sidewalks, we selected five visual quality indices: Greenness, Enclosure, Imageability, Complexity, and Cleanness (Figure 3).

This project followed prior studies and defined the proxy of visual greenness as the proportion of greenspace vegetation intermixed with building façades. Previous research has revealed a significant correlation between Openness and Enclosure. We, hence, only employed the Enclosure index in this study. Enclosure is the degree to which streets and other public spaces are visually defined by buildings, walls, trees, and other vertical elements (Ewing & Handy, 2009). The analyse the skywalks' covered outdoor spaces, we included ceiling elements for measuring Enclosure, in addition to buildings and fences. Imageability is the quality or trait of a place that makes it distinct, recognizable, and memorable (Lynch, 1960). To characterize visual imageability, the proportion of buildings, signs, and boards is calculated for each GSV.

Perceptual Quality	Significant Physical Features	Equations
Greenness	Tree, Grass, Palm, Plant	$Green_i = VI_{tree} + VI_{grass} + VI_{palm} + VI_{plant}$
Enclosure	Building, Railing, Fences, Ceilings	$Encls_i = \frac{VI_{bldg} + VI_{tree} + VI_{ceil} + VI_{rail}}{VI_{road} + VI_{sidewalk} + VI_{floor}}$
Imageability	Building, Signs (Signboard, Banister, Poster)	$Imgbl_i = VI_{bldg} + VI_{sign}$
Complexity	Person, Building, Benches, Signs, Windowpane, Stage, Poles, Lights	$Cmplx_i = \frac{VI_{bldg} + VI_{tree} + VI_{persn} + VI_{benches} + VI_{sign} + VI_{pole} + VI_{win}}{VI_{road} + VI_{sidewalk} + VI_{floor}}$
Cleanness	Ashcan, Trash, Garbage, Dusk	$Clean_i = 1 - VI_{trash} + VI_{dusk} + VI_{garbage} + VI_{ashcan}$

Table 1. Subjective Measurement of 6 Visual Quality Indices

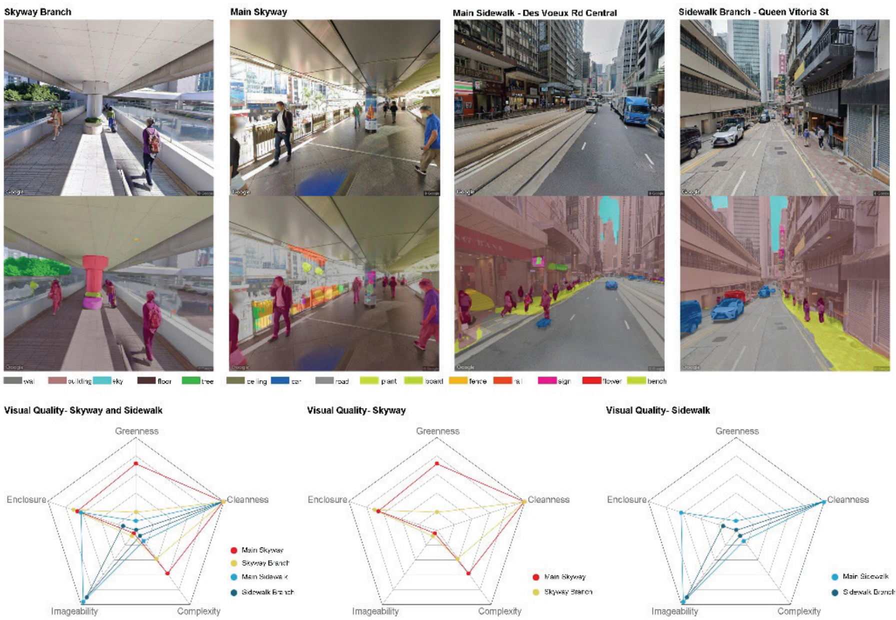


Figure 3. GSV Images and Semantic Segmentation Samples on Skyways and Sidewalks

Complexity, the visual richness of a place, is commonly evaluated by the number of types of buildings, landscapes, and street elements. The proportion of persons, buildings, and varied kinds of street furniture is extracted for evaluation. Cleanness/ Tidiness is a crucial visual quality factor that affect people's perceived safety especially on skywalks (Clifton et al., 2007; Ferrer et al., 2015). The proportion of littering, trash, and trash bins is extracted to present the level of maintenance of the environment. The definition and formula of each perceptual quality indices are listed in Table 1.

2.5. ON-SITE AUDITS-PERCEPTUAL ENVIRONMENTAL FEATURES AND PEOPLE'S PERCEPTION

Two instruments were developed to conduct the on-site audits, based on two established audit instruments developed by Cerin et al. (2010) and Zhang et al. (2018). The NEWS-CS possesses sufficient levels of reliability to be used for measuring the perceived neighbourhood environment in East Asia. We selected fifteen sub-indicators specifically related to the street environment, which are categorized into six parts: Physical barriers to walking, Aesthetic, Presence of People, Crowdedness, Traffic and road hazards, and social disorder / Littering (Fig. 4). To explore the relationship between environmental features and people's perception, we designed a Likert scale questionnaire to collect the perceptual rating of five dimensions: safe, beautiful, depressing, lively, and boring, which is developed by MIT's Place Pulse 2.0' project (Zhang et al., 2018).

1.非常不同意　2. 部分不同意　3.部分同意　4.非常同意 **1.strongly disagree　2.disagree　3.agree　4.strongly agree**	
A. 行走的物理障碍 Physical barriers to walking	
	1) 丘陵街道 Hilly streets
	2) 主要障碍（如道路工程、高速公路）Major barriers (e.g., roadwork, freeways)
	3) 许多死胡同 Many dead ends
B. 美学 Aethetic	
	1) 树木 Trees
	2) 许多有趣的事情要看 Many interesting things to look at
	3) 许多吸引人的自然景观 Many attractive natural sights
	4) 有吸引力的建筑/房屋 Attractive buildings/homes
C.人员在场 Presense of People	
	1）可以看到步行者 Walkers can be seen
	2）很难寻求帮助，因为周围人不多 Difficult to ask for help because not many people around
D. 拥挤 Crowdness	
	1) 拥挤的街道 Crowded streets
	2) 街道上的"小贩"和商店挡住了去路 'Hawkers' and shops on streets blocking the way
E. 交通和道路危险 Traffic and road hazards	
	1) 街道和人行道很滑 Streets and sidewalks slippery
	2) 空气污染（例如废气）Air pollution (e.g., exhaust fumes)
F. 社会混乱/乱扔垃圾 Social disorder / Littering	
	1) 动物粪便 Animal droppings
	2) 许多无家可归者、吸毒者、妓女 Many homeless people, drug addicts, prostitutes
G.感知 Perception	
	1) Safe
	2) Beauty
	3) Lively
	4) Boring
	5) Depressing

Figure 4. The NEW-CS Scale for Subjective Evaluation of Environmental Features

All procedures involving human subjects in this study were approved by the ethics committees of the participating institutions. Eight participants, with no diagnosed cognitive impairment and able to walk unassisted, were recruited to conduct on-site audits in the research area. We sampled 76 points in total along the selected skyways and sidewalks. The participants were asked to fill in the two Likert-scales at each sampling point. The field experiments were conducted on several weekdays in November 2022, under conditions of clear, windless afternoons to minimize the effect of extreme weather on the results.

2.6. STATISTICAL ANALYSIS

First, we calculate the average value for each sample point with four GSV views. A standardized Z-score of five indices was then calculated so that the values would be comparable among different indices. The study area was divided into four types: main skyway, skyway branches, main sidewalk, and sidewalk branches. We then conducted a descriptive analysis of the Visual Quality indices of 158 sampling points, to compare the range, mean, and standard deviation of the values. Second, for each item of CS-NEWS, we calculate the average value of the eight volunteers. The sum values of six indices were calculated separately for each type of walking area. The descriptive analysis was conducted in SPSS. Similarly, the range, mean and standard deviation were processed for evaluating volunteers' perceptions at 76 sampling points. Last, we conducted a Pearson correlation analysis separated for skyway and sidewalk, to compare the difference of significant indicators.

3. Results

3.1. THE SPATIAL DIFFERENCE OF VISUAL QUALITY ON SKYWAYS AND SIDEWALKS

The descriptive results (Table 2) show a significant visual quality difference between skyways and sidewalks.

Visual greenness on skyway segments is much higher than on sidewalks. The semantic segmentation result reveals that trees, flowers, palms, and plants were extracted on the skyway, while only small proportions of trees were found on sidewalk segments. Skyways have a higher value of enclosure, which is consistent with our field observations. Compared to sidewalks, skyway segments have a large proportion of ceilings. In contrast, the value of the imageability of skyways is significantly lower than sidewalks. In general, the distribution of visual elements on the sky bridges is spatially homogeneous, and the different visual elements are relatively evenly distributed. In terms of complexity, skyways achieved a higher value. Richer street furniture, for instance, benches, flower ponds, signboards, and stairs were extracted from their GSVs. Last, litter/trash was not detected on most of the sampling points. This results in a large standard deviation of the dataset. In general, the skyway is much cleaner, likely owing to good maintenance associated with private ownership.

	Skyway Overall (N=56)				Main Skyway (N=34)				Skyway Branches (N=22)			
	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD
Z_greenness	-0.552	5.805	**0.663**	1.319	-0.552	4.881	**0.626**	1.039	-0.552	5.805	**0.720**	1.689
Z_enclosure	-1.178	2.363	**0.233**	0.775	-1.178	1.376	**0.006**	0.554	-0.854	2.363	**0.584**	0.937
Z_imageability	-1.528	0.489	**-0.957**	0.579	-1.525	-0.380	**-1.122**	0.353	-1.528	0.489	**-0.702**	0.754
Z_complexity	-0.825	3.205	**0.538**	0.950	-0.825	2.755	**0.568**	0.863	-0.759	3.205	**0.491**	1.090
Z_cleanness	-1.609	0.217	**0.107**	0.378	-1.005	0.217	**0.145**	0.238	-1.609	0.217	**0.049**	0.529

	Sidewalk Overall (N=89)				Main Sidewalk (N=65)				Sidwalk Branches (N=23)			
	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD
Z_greenness	-0.552	1.260	**-0.437**	0.279	-0.552	0.332	**-0.499**	0.127	-0.552	1.260	**-0.258**	0.468
Z_enclosure	-1.267	5.052	**-0.150**	1.081	-1.267	5.052	**-0.261**	1.079	-0.739	4.353	**0.168**	1.044
Z_imageability	-1.525	1.484	**0.686**	0.585	-1.129	1.387	**0.619**	0.576	-1.525	1.484	**0.877**	0.582
Z_complexity	-1.038	3.738	**-0.327**	0.820	-1.038	3.738	**-0.267**	0.910	-0.893	1.123	**-0.499**	0.443
Z_cleanness	-9.284	0.217	**-0.035**	1.253	-9.284	0.217	**-0.025**	1.233	-6.190	0.217	**-0.061**	1.336

Table 2. Descriptive Analysis of Objective Visual Quality (GSV)

3.2. PERCIEVED FEATURES AND PEDESTRIAN PERCEPTIONS

According to the results (Table 3), participants perceive the physical barriers on the skyway less than the sidewalks. Especially the branches of the sidewalk on the slope are thought to have the most barriers. The aesthetics variable mean values of the skyway are much higher than those of the sidewalk, while both skyway and sidewalk are perceived as not aesthetic. Based on the NEWS-CS items, aesthetics is composite measures of trees, many interesting things to look at, many attractive natural sights, and attractive buildings/homes. As mentioned in the last paragraph, the objective measures of greenness are much higher on the skyway than on the sidewalk. This result shows the concordance between objective and perceived measures. Both footbridges and sidewalks are perceived to have many people around, and the sidewalks are thought to be more crowded. The mean values of traffic and road hazards of the skyway are lower than those of sidewalks. Both skyways and sidewalks are perceived as very clean.

	Skyway Perceived Walkability Overall (N=217)				Main Skyway Perceived Walkability (N=112)				Skyway Branches Perceived Walkability (N=112)			
	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD
A. Physical barriers to walking	1.000	1.667	**1.032**	0.132	1.000	1.000	**1.000**	0.000	1.000	1.667	**1.067**	0.187
B. Aesthetics	1.750	3.750	**2.476**	0.579	1.750	3.750	**2.625**	0.563	1.750	3.500	**2.317**	0.571
C. Presence of people	3.500	4.000	**3.871**	0.222	3.500	4.000	**3.781**	0.256	3.500	4.000	**3.967**	0.129
D. Crowdedness	1.000	2.000	**1.387**	0.308	1.000	2.000	**1.219**	0.315	1.500	2.000	**1.567**	0.176
E. Traffic and road hazards	1.000	1.500	**1.371**	0.222	1.000	1.500	**1.281**	0.256	1.000	1.500	**1.467**	0.129
F. Social disorder/littering	1.000	1.000	**1.000**	0.000	1.000	1.000	**1.000**	0.000	1.000	1.000	**1.000**	0.000

	Sidewalk Perceived Walkability Overall (N=315)				Main Sidewalk Perceived Walkability (N=245)				Sidewalk Branches Perceived Walkability (N=70)			
	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD
A. Physical barriers to walking	1.000	3.000	**1.341**	0.505	1.000	2.000	**1.162**	0.296	1.000	3.000	**1.967**	0.597
B. Aesthetics	1.500	3.250	**1.839**	0.392	1.500	2.750	**1.757**	0.281	1.500	3.250	**2.125**	0.580
C. Presence of people	3.500	4.000	**3.878**	0.217	3.500	4.000	**3.914**	0.191	3.500	4.000	**3.750**	0.264
D. Crowdedness	1.000	3.500	**2.200**	0.537	1.500	3.500	**2.314**	0.501	1.000	2.500	**1.800**	0.483
E. Traffic and road hazards	1.500	2.500	**2.178**	0.285	2.000	2.500	**2.257**	0.254	1.500	2.000	**1.900**	0.211
F. Social disorder/littering	1.000	1.500	**1.011**	0.075	1.000	1.000	**1.000**	0.000	1.000	2.000	**1.100**	0.316

Table 3. Descriptive Analysis of Subjective Environmental Features (NEWS-CS)

Table 4 shows the perception variables. The safe values are the highest among all the positive perception values on the sky bridges and sidewalks. This may be because there are many people on the skyways and sidewalks, and the sky bridges have security guards around the entrances of several buildings. In addition, the sidewalks are perceived as slightly more lively and less boring than the footbridges, while the main skyway is perceived as the most lively among all the selected areas. This may be due to the plazas connecting with the main skyway, and there are many activities besides walking; for example, people sit, chat and wander.

	Skyway Perception Overall (N=217)				Main Skyway Perception (N=112)				Skyway Branches Perception (N=112)			
	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD
1) Safe	2.000	4.000	**3.673**	0.608	2.000	4.000	**3.839**	0.393	2.000	4.000	**3.495**	0.735
2) Beauty	1.000	4.000	**2.641**	0.913	1.000	4.000	**2.795**	0.807	1.000	4.000	**2.476**	0.991
3) Lively	1.000	4.000	**2.622**	1.002	1.000	4.000	**2.732**	0.890	1.000	4.000	**2.505**	1.102
4) Boring	1.000	4.000	**2.111**	0.994	1.000	4.000	**1.848**	0.872	1.000	4.000	**2.390**	1.042
5) Depressing	1.000	4.000	**1.691**	0.973	1.000	4.000	**1.429**	0.846	1.000	4.000	**1.971**	1.023

	Sidewalk Perception Overall (N=315)				Main Sidewalk Perception (N=245)				Sidewalk Branches Perception (N=70)			
	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD
1) Safe	1.000	4.000	**3.505**	0.679	1.000	4.000	**3.486**	0.675	2.000	4.000	**3.571**	0.693
2) Beauty	1.000	4.000	**2.448**	0.874	1.000	4.000	**2.416**	0.853	1.000	4.000	**2.557**	0.942
3) Lively	1.000	4.000	**2.743**	0.903	1.000	4.000	**2.690**	0.906	1.000	4.000	**2.929**	0.873
4) Boring	1.000	4.000	**1.876**	0.948	1.000	4.000	**1.861**	0.931	1.000	4.000	**1.929**	1.012
5) Depressing	1.000	4.000	**1.762**	0.901	1.000	4.000	**1.800**	0.917	1.000	4.000	**1.629**	0.837

Table 4. Descriptive Analysis of Pedestrian's Perceptions

Fig. 5 (left) shows the Pearson Correlation results of overall study area, including values retrieved from skyway and sidewalk segments. Fig. 5 (right) shows the results specifically on skyway segments. The correlation between subjective measurement of visual quality/ objective environmental features (NEWS-CS) and people's perception were analysed separately.

For objective visual quality indices, in general, enclosure is negatively correlated with beauty and lively, and positively (sig.= -.233*, -.288*) correlated with boring and depressing (sig.= .437**, .012**). Imageability shows negative correlation with the perception of boring (sig.= -.308*). Surprisingly, the values of complexity are positively correlated with the perception boring (sig.= 362**). When zoomed into skyways' condition, the correlations between varied visual quality indices and people's perception are more significant. The results indicate that cleanness may be a crucial indicator on skyways, while the correlations between cleanliness and people's perception are not statistically significant after adding the values from sidewalks.

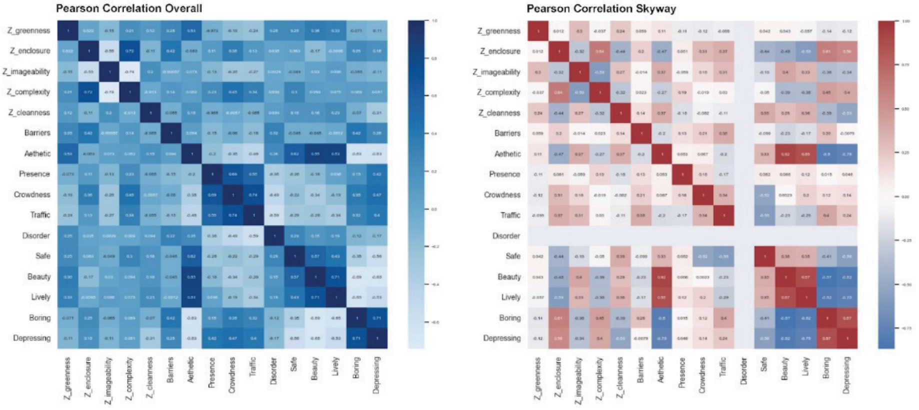


Figure 5. Correlation Analysis Between Subjective/ Objective Street Quality and Perception

Several subjective indicators are correlated with perception indices. Aesthetic, crowdedness, and the traffic and road hazards may have positive or negative impact on people's walking perception. It's noted that the crowdedness is negatively correlated with perception of safety/beauty (sig. =443**, -279*), and it also reveals the positive correlation with the sense of depressing. This also accords with our earlier observations, which showed that overcrowding can be unpleasant for pedestrians (Ferrer et al., 2015). Moreover, traffic and road hazards may affect the perception of safety, which is hard to obtain via virtual audits by GSV images.

4. Conclusions and Future Works

As an innovative empirical study that combines virtual and on-field audits to analyse the spatial difference on skyways and sidewalks in the compact city context, this study has presented a pilot project for analysing the relationships between subjective/objective street quality and pedestrians' perceptions. In short, it can be concluded that (1) there are significant visual quality difference between skyways and sidewalks, especially in greenness, enclosure, and complexity; (2) There is considerable variability in environmental perceptions between skyways and sidewalks, especially in the Aesthetics, Crowdedness and Traffic and road hazards. (3) For the between-skyway and sidewalks variance in perceptions seemed limited. Feeling secure is a common emotion shared by the two groups; (4) The correlated variables differed on skyways and sidewalks. While most of the objective visual quality indices are significantly correlated with perception indices on skyways, subjective indicators, for instance, Aesthetic, Crowdedness, Traffic and Hazards, showed statistical correlations with pedestrian's perceptions, which should be included into the mixed methods criteria.

In addition, several issues warrant discussion in future work. Recent studies have indicated that simply summing up or recombining visual elements could not comprehensively capture or represent more comprehensively defined perceptual quality. While this research combined the subjective and objective measurements, ssignificant street-level adaptations were applied to the Hong Kong version of the NEWS-CS questionnaire, so only a limited number of comparable built environment items could be included in the analysis. The validity of the mixed-method evaluation should be tested in a broader context.

References

Adkins, A., Dill, J., Luhr, G., & Neal, M. (2012). Unpacking Walkability: Testing the Influence of Urban Design Features on Perceptions of Walking Environment Attractiveness. Journal of Urban Design, 17(4), 499–510. https://doi.org/10.1080/13574809.2012.706365

Cerin, E., Sit, C. H. P., Cheung, M. chin, Ho, S. yin, Lee, L. chun J., & Chan, W. man. (2010). Reliable and valid NEWS for Chinese seniors: Measuring perceived neighborhood attributes related to walking. International Journal of Behavioral Nutrition and Physical Activity, 7. https://doi.org/10.1186/1479-5868-7-84

Clifton, K. J., Livi Smith, A. D., & Rodriguez, D. (2007). The development and testing of an audit for the pedestrian environment. Landscape and Urban Planning, 80(1–2), 95–110. https://doi.org/10.1016/j.landurbplan.2006.06.008

Ewing, R., & Handy, S. (2009). Measuring the unmeasurable: Urban design qualities related to walkability. Journal of Urban Design, 14(1), 65–84. https://doi.org/10.1080/13574800802451155

Ferrer, S., Ruiz, T., & Mars, L. (2015). A qualitative study on the role of the built environment for short walking trips. Transportation Research Part F: Traffic Psychology and Behaviour, 33, 141–160. https://doi.org/10.1016/J.TRF.2015.07.014

Johansson, M., Sternudd, C., & Kärrholm, M. (2016). Perceived urban design qualities and affective experiences of walking. Journal of Urban Design, 21(2), 256–275. https://doi.org/10.1080/13574809.2015.1133225

Lu, Y. (2019). Using Google Street View to investigate the association between street greenery and physical activity. Landscape and Urban Planning, 191. https://doi.org/10.1016/j.landurbplan.2018.08.029

Ma, X., Ma, C., Wu, C., Xi, Y., Yang, R., Peng, N., Zhang, C., & Ren, F. (2021). Measuring human perceptions of streetscapes to better inform urban re-newal: A perspective of scene semantic parsing. Cities, 110. https://doi.org/10.1016/j.cities.2020.103086

Rossini, F., Roca, E., & Harris, S. (2018). The notion of ground: A definition of urban permeability in Hong Kong and Barcelona. Architecture, 13, 211–234. https://doi.org/10.5821/ace.13.38.5432

Tang, J., & Long, Y. (2019). Measuring visual quality of street space and its temporal variation: Methodology and its application in the Hutong area in Beijing. Landscape and Urban Planning, 191, 103436. https://doi.org/10.1016/j.landurbplan.2018.09.015

Villani, C., & Talamini, G. (2020). Socialising on a Skywalk: How Hong Kong's elevated walkways become public open space. *Asian Journal of Environment-Behaviour Studies,* 5(15), 57–72. https://doi.org/10.21834/aje-bs.v5i15.361

Ye, Y., Zeng, W., Shen, Q., Zhang, X., & Lu, Y. (2019). The visual quality of streets: A human-centred continuous measurement based on machine learning algorithms and street view images. *Environment and Planning B: Urban Analytics and City Science,* 46(8), 1439–1457. https://doi.org/10.1177/2399808319828734

Zhang, F., Zhou, B., Liu, L., Liu, Y., Fung, H. H., Lin, H., & Ratti, C. (2018). Measuring human perceptions of a large-scale urban region using machine learning. *Landscape and Urban Planning,* 180, 148–160. https://doi.org/10.1016/j.landurbplan.2018.08.020

BESPOKE 3D PRINTED CHAIR

Research On The Digital Design And Fabrication Method Of Multi-body Pose Fusion

KEKE LI[1], HAO WU[2], XIANGWEN DING[3], HAOWEI LI[4] and PHILIP F. YUAN[5*]
[1,2,3,4,5]*Tongji University, College of Architecture and Urban Planning, Tongji University*
1cora@tongji.edu.cn, 0000-0002-1400-4980
22032186@tongji.edu.cn,0000-0002-0503-1696
[3]*DingS0818@outlook.com, 0000-0003-2416-2613*
[4]*15218851337@163.com, 0000-0002-0342-7355*
[5]*philipyuan007@tongji.edu.cn, 0000-0002-2871-377X*

Abstract. Customization has become a vital part of the post-industrial production model. Traditional production methods struggle with the challenge of collecting large amounts of data and the high costs of customization. With advancements in big data and deep learning algorithms, it is now possible to reduce the difficulty and cost of data collection, resulting in more accurate individual customization. This paper presents a proposed workflow for customizing multi-position seating for individuals. Using deep learning algorithms such as OpenPose and parametric design platforms such as Grasshopper, the workflow transforms user-generated photos of the body into a seating model that fits corresponding positions. This process combines deep learning algorithms, simplifies the data collection and processing process, and provides an interface for user interaction on the Grasshopper platform. The workflow provides a comprehensive example of data-driven customization in the context of big data. It explores the potential of a new paradigm in digital design where data is the primary driving force.

Keywords. Customization, Posture Recognition, Multi-platform Collaboration, Deep Learning, Workflow

1. Introduction

The term "customization" originates in the tradition of "bespoke" suits from Savile Row. Its full name, "to be spoken", signifies "tailor-made suits for individuals." This highlights the essence of customization: the precise gathering of individual data and unique differences. As customization evolves into a mass production industry, it encounters the challenge of collecting and processing vast amounts of user data. Without advanced data collection and analysis capabilities, customization is limited to batch production for specific groups and cannot truly cater to individual differences.

1.1. THE CONTRADICTION AND EXPLORATION OF CUSTOMIZATION

In his book Consumer Society, French sociologist Jean Baudrillard referred to customization as a "monopoly concentration of the production of differences" (Baudrillard, 2016). This observation highlights one of the contradictions between customization and industrial production: the inability to fully capture the nuanced differences between users, leading to further classification and modeling of users. Innovation in related technologies is required to overcome this contradiction.

Since mass customization was first introduced in the 1980s, various efforts have been made to reconcile the tensions between customization and batch production. Furniture companies have introduced modular customization options. Y.H. Chen et al. proposed using a data adjustment process for specific product information (Chen et al., 2001). Laura Trautmann proposed a workflow that involves inputting individual information using computer simulation software (Trautmann, 2021). Z.J. Zhang used mechanical sensors to collect real-time body data (Zhang, 2013).

Over the years, the customization process has evolved from a broad, modular process to a more precise one, accompanied by increasing data collection and processing complexity. The challenge of collecting user data conveniently and accurately has become the focal point of the new era of large-batch customization. The advent of big data and the development of deep learning algorithms have provided new solutions to this issue. Today, we can collect part of the user's body information and use deep learning models to predict and simulate the rest of the body information, bridging the gap between customization and industrial production.

This paper proposes a novel workflow for multi-pose fusion that uses photo input to collect user data and makes it easier for users to participate, representing a paradigm shift from parametric customization to data customization.

1.2. INTRODUCTION TO POSTURE RECOGNITION RESEARCH

The study of human posture recognition and acquisition has evolved significantly from single postures to compound postures and finally to multi-posture recognition. The main content of this paper explores human posture as the core design methodology for multi-posture fusion.

Since human body posture is determined by the relative positioning of the skeletal landmarks, and all individuals have similar skeletal structures, posture recording, and encoding is possible. In 1955, Benesh used the Benesh notation to record dance movements(Abbie, 1974). Later, American landscape architect Lawrence Harplin developed the Notation System method to encode human body behavior in the design of landscapes (Merriman, 2010).

Oscar Schlemmer explored the relationship between the human body and stage space during the Bauhaus era through custom dance costumes. By designing costumes that hindered human posture, he effectively designed the "virtual space" represented by human posture trajectory (Preston, 2014).

Based on previous research, this study aims to incorporate human pose into the data collection of the customization process. We propose a workflow that involves deep learning algorithms to convert photos into digital body models, which are then used to generate a chair surface. The chair's structure is optimized and finally generated through the use of software such as Kangaroo, Karamba3D (Clemens, 2013), Biomorpher (Harding & Brandt-Olsen, 2018), and others. The entire chair design process is broken down into several work packages, each with user interaction content allowing personal participation and adjustments to fit individual preferences.

2. Method

The components used in constructing chairs allow for adjustable angles, resulting in the potential for multiple postures, such as regular sitting or lying. Whether it can support multiple postures in a continuous print is a potential proposal this project wants to answer. Accurate user body data is necessary to ensure the chair accommodates specific postures. Moreover, a suitable surface generation method must be found to create a comfortable wrap for the body. Structural verification and stability analysis must be conducted to guarantee that the chair can also support the body's weight. The chair must also be designed to be as thin and lightweight as possible while still meeting structural requirements. Throughout the process, ample customer engagement opportunities should be provided to offer personalized choices.

The study has developed a picture-to-model multi-platform workflow (Figure 1) to achieve these goals. The OpenPose pose recognition algorithm and the SMPL-X (Pavlakos et al., 2019) agent generation algorithm are utilized for inputting user data. First, collecting and processing primary human data, transforming the input human body photos into agent models. Then, the project uses the Kangaroo plug-in to generate a surface that wraps around the body, providing a quantitative measurement standard for the surface prototype. Finite element analysis was performed on the surface using karamba3D, and a toppling test was conducted under external forces. Finally, the multi-objective optimization with Biomorpher ensures the chair's lightweight and slim design while maintaining stability. The following experiment was conducted to demonstrate the entire workflow process.

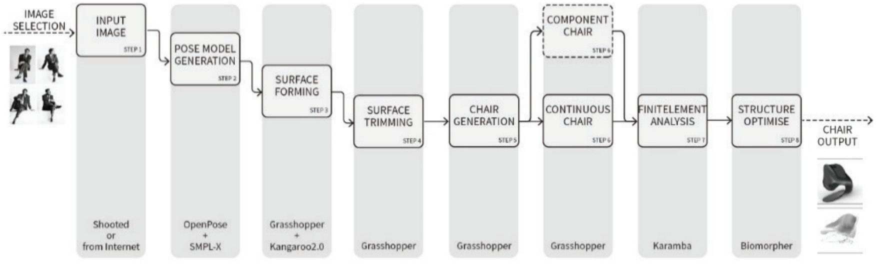


Figure 1. The Process of Workflow.

In the experiment, we recreated the entire workflow process. We obtained input data by retrieving pictures of a seated human body from Pinterest. Using OpenPose and SMPL-X, we calculated several pose-consistent proxy models. We then utilized Grasshopper, Kangaroo, karamba3D, and other plug-ins to generate a 3D model of the seat. Finally, we printed a seat prototype using plastic lamination and tested its bearing capacity and comfort through a human-body sitting test and posture reproduction.

We integrated deep learning algorithms into our custom workflow and streamlined the data collection process using posture recognition on body photos. This significantly reduced the workload in collecting different information. Using the collected human body data, we updated the original ergonomic classification and proposed new, personalized ergonomics. We also made this workflow available to clients, providing them with more customization options.

3. Workflow

3.1. FROM IMAGE TO MODEL FILE

The advancements in digital technology have led to the development of tools for human posture recognition. Utilizing deep learning, computer models can capture human joint checkpoints and record relative positions. This process, known as posture recognition, is commonly used in animation and game production to record human movements. Currently, the most popular posture recognition frameworks include OpenPose, developed by CMU (Cao et al., 2017); DensePose, developed by Facebook (Güler et al., 2018); and AlphaPose, developed by Shanghai Jiao Tong University (Fang et al., 2017). These algorithms focus on solving the problem of a 2D picture to checkpoint reconstruction. 3D human reconstruction involves creating a digital model of the human body and storing a wide range of ergonomic and personalized data such as height, weight, and other parameters. Models like SMPL and SMPL-X are commonly used as 3D parameter representations.

Therefore, we selected OpenPose to work with the SMPL-X algorithm to reconstruct the 2D image file, which was achieved by locally scaling, partially rotating, and fitting the existing proxy model to the image's edges.

Once the user inputs a JPG image, the OpenPose algorithm detects 18 human joint points, including the eyes, nose, neck, shoulders, elbows, hands, hips, knees, and feet. The trained OpenPose model extracts the relative values of the joint point coordinates and scales them to standard coordinates, which are then stored in JSON files. By reading the JSON file, we can access the point coordinates and basic body information processed by the algorithm.

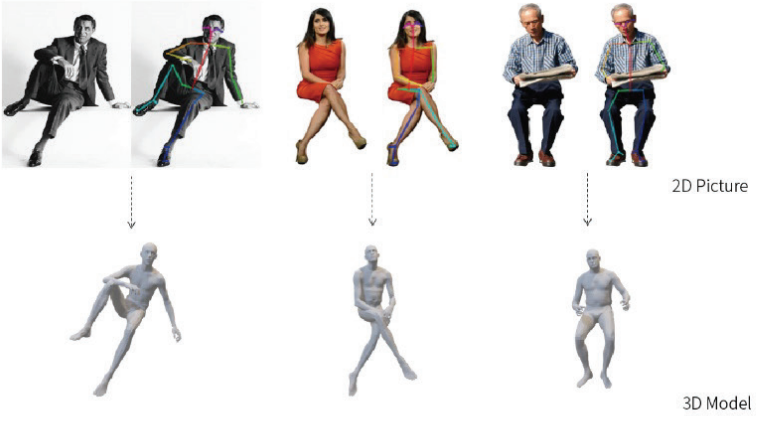


Figure 2. OpenPose and SMPL-X Algorithm Conversion

Next, we send the JSON file and the original input image to the SMPL-X model to generate the proxy model. The pixel information stored in the JPG image is utilized to confirm the human body's edge curve, and the JSON file is used to determine the proxy model's joint points. SMPL-X generates a surrogate model by adjusting the betas matrix, which controls the human body shape, aligning the body joints' relative positions with the JSON file's data. The ten parameters in the matrix correspond to different muscle states of the human body. After the simulation is completed, the SMPL file will produce an obj mesh model of the human body for output (Figure 2). This process enables us to obtain a highly accurate representation of the human body in the photo, which can then be used for surface fitting and form finding.

3.2. SURFACE GENERATION AND TRIMMING

Grasshopper is a powerful parametric design platform integrated into Rhino 3D modeling software, allowing designers to use C#, VB, or Python scripting languages to create models that respond to changes in parameters. The platform is further enhanced by the addition of Kangaroo, a force-based plug-in developed by Daniel Parker. Kangaroo leverages a computer particle spring model to simulate physical phenomena and is widely used in parametric design as a form-finding tool. With Kangaroo, designers can easily simulate the morphological changes of objects under stress, such as tensile films, elastic cloth surfaces, and inverted catenaries.

In the surface generation process, we establish a relationship between the human body and the seat surface, then use Kangaroo to generate the surface. The goal is to create a surface that fits comfortably to the body while providing adequate forward and lateral support. To achieve this, we use Kangaroo to construct a physical model (Figure 3) that generates the seat surface based on the human model.

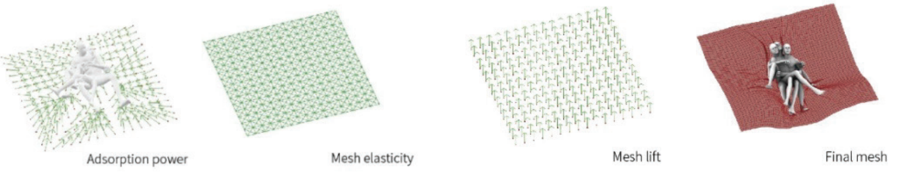


Figure 3. Kangaroo Model Diagram.

In this physical system, four main forces determine the shape and behavior of the fabric: an elastic force that maintains the primary form of the fabric, lifting and flattening forces that help the fabric conform to the human body and reduce wrinkles, and an adsorption force that enables the fabric to follow the body's shape closely. Essentially, it's just like the body is being wrapped in a flexible and magnetic mesh. To refine the surface, we then utilized the Quadmesh subdivision tool in Grasshopper to subdivide and reconstruct the mesh, eliminating folds and achieving the final surface. The parameters of these forces and the accuracy of the reconstruction were determined through experimentation. The offset distance between the surface and the body surface was measured under various parameters (Figure 4a), cases with low redundancy were observed and eliminated, and finally, optimal combination of parameters was selected

(Figure 4b).

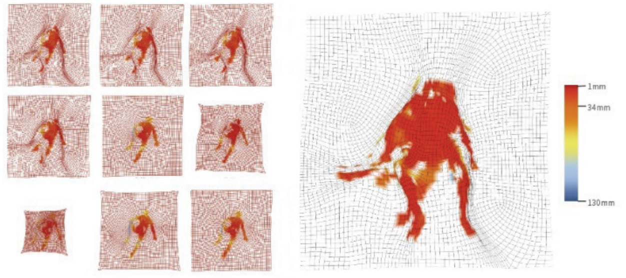


Figure 4a. Comparison of Deviations with Different Parameters. 4b. Deviation of selected parameters

After evaluating the suitability through the parametric mesh, we proceed to trim the surface. The method for trimming seat surfaces (Figure 5) is based on the human body form and construction. Our aim was to create a smooth edge covering most of the human body contour while allowing for personalized adjustments. To achieve this, we utilized the envelope algorithm to generate a grid that conforms to the human body model. The grid's outer edge approximates the human body's plane contour, and we extracted the curve's equipartition points as the human body contour points. Next, we identified the head and knee points from the SMPL model's human body junction points and replaced the corresponding human body contour points with them. We then constructed NURBS curves using the remaining human body contour points as control points. Finally, the curve's projection was used to trim the previously generated surface, yielding the seat surface. In this process, the user can adjust the seat's curved edge by changing the position of the head and knee points.

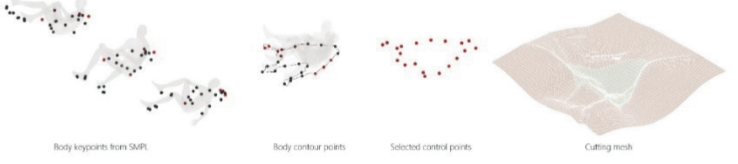


Figure 5. Cutting Methods of Seat Surfaces.

3.3. BASE GENERATION AND FORM MAKING

After customizing the seat surface, this workflow offers two options for generating the seat base - an integral bottom seat or a separate bottom seat.

Integral Base (Figure 6): After trimming the seat surface, we convert it into a mesh, extract its vertices, and reorder them in XYZ coordinates. We identify the vertex with the minimum Z value from the entire chair and use it as the root node for seat base generation. We then offset the vertices in the specified direction to reconstruct the mesh. By default, the seat base is generated at the height of 350mm, depth of 400mm, and width of 700mm. The user can adjust these parameters to accommodate changes in the sitting position. Upon completing the data adjustment, we will have an entire seat surface without thickness. To determine the thickness, the use of the Karamba3D plug-in in conjunction with the Biomorpher multi-objective solution plug-in is required.

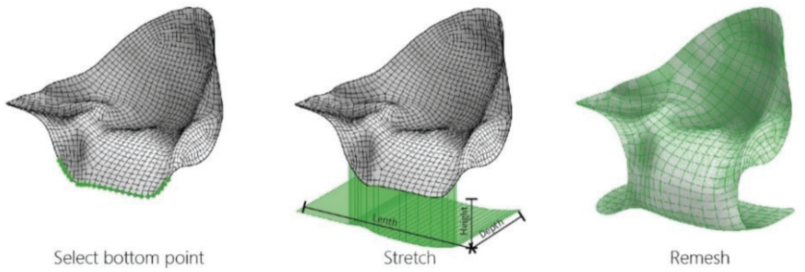


Figure 6. Integral Base Generation.

Separate Base: To start, we calculate the appropriate support points for the seat surface. To minimize surface stress, it is crucial to position the support points strategically:1) Use Karamba3D to analyze surface stress. 2) Set the closest point between the surface and the human buttocks as the initial support. 3) Calculate the surface stress. 4) Select the maximum stress point on the surface, add it to the set of support points, and then repeat the calculation with the updated set. Repeating steps three and four, we obtain the final set of support points. The results of our tests indicate that this iterative method yields a smaller maximum offset of the support points compared to a uniform distribution of the same number of points (Figure 7). Next, we project the support points onto the ground and connect the projected points using an insertion point curve. The dislocation connection between the projection and support points results in a rod system base (Figure 8).

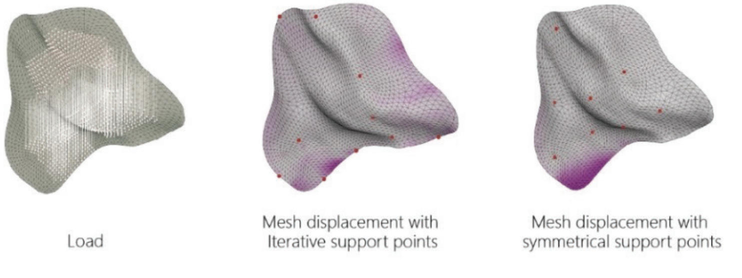


Figure 7. Diagram of Support Points.

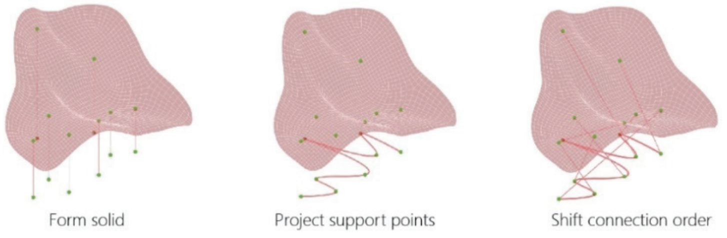


Figure 8. Separated Base Generation.

The stress of the seat surface appears unevenly distributed under human load. We aim to use the material effectively by showing the stress condition on the seat surface through its thickness. Thus, the solid thickness of the surface will be set based on the stress distribution at different positions. We use Karamba3D to simulate the deformation of the model under load and display the stress state in color on the object's surface. The darker the color, the higher the stress force at that position. We convert the force data from RGB to gray data to simplify it to a single value, which we then map to the actual thickness range of the chair surface. Our initial thickness interval is set as 2cm-10cm, which we adjust through Biomorpher and verify the structural strength in Karamba3D by sorting results based on deformation. The relationship between the thickness of the interval and structural strength is positive, but an excessively thick structure goes against the design purpose. Hence, we manually select a lighter yet strong enough chair surface as the final choice.

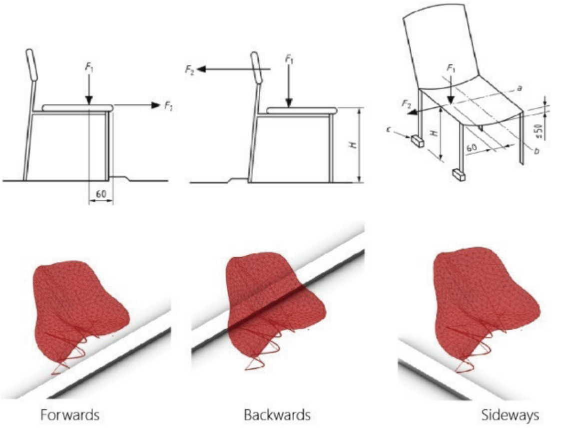


Figure 9. Stability Test.

3.4. STABILITY CHECK

The stability check is a crucial step before printing the seat. In this step, we use Kangaroo to determine if the seat will tip. Adhering to the EN 1022-2018 standard for furniture seat stability determination, we model the seat as a rigid body in Kangaroo and apply loads according to the experimental method specified in the standard. We test its stability in the forward, backward, and roll directions. The results of our simulation show that the seat remains stable under the specified loads (Figure 9).

3.5. 3D PRINTING

We selected the integral seat for exploration in 3D printing manufacturing. Using the FURobot plug-in on the Grasshopper platform, we divided and reconfigured the electronic model based on a 3mm layer spacing to create the robot's printing path. The path was divided into multiple points as the machine's printing head coordinates, allowing the robot to carry out the printing process. Due to the continuity requirements of laminated printing, we could only print with the chair's left or right side as the base plane. The outcome of this process was a prototype chair (Figure 10). Upon testing, we found the prototype seat sturdy and capable of accommodating a wide range of pre-set seating positions.

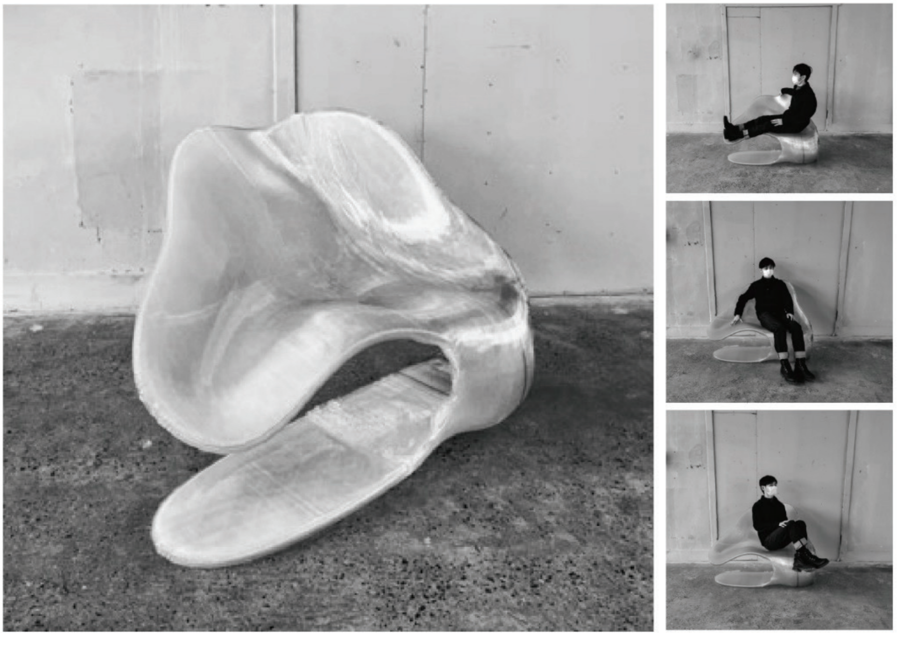


Figure 10. Close-up of the demonstrator

4. Summary and Outlook

In this study, we present a novel approach to customizing chairs. The concept allows users to create a chair that fits their body shape and reflects their personality by inputting photos and preferences. This process is made cost-effective using deep learning algorithms, reducing the need for extensive user data collection. Using parametric design platforms and 3D printing technology reduces the workload of customizing design and manufacturing.

Traditionally, chairs have been regarded as standardized products in industrial production. Our workflow re-establishes the connection between the seat and the body, transforming the chair into a form of body expression. This exploration opens up new possibilities for future industrial customization. The design approach of multi-pose fusion has significant implications in architecture as well. Considering body size and posture, this method results in both minimalist and comfortable spaces. Given the current pandemic, exploring such space ontology is a topic worth further discussion and investigation.

Acknowledgements

This article is funded by Ministry of Housing and Urban-Rural Development (Grant No.2021-R-085)and Shanghai Science and Technology Committee (Grant No.21DZ1204500).

References

Abbie, M. 1974. Movement notation. Australian Journal of Physiotherapy, 20(2), 61-69, https://doi.org/10.1016/S0004-9514(14)61177-6.

Baudrillard, J. 2016. The consumer society: Myths and structures. Sage. (pp.81-82)

Cao, Z., Simon, T., Wei, S. E. and Sheikh, Y. 2017. Realtime multi-person 2d pose estimation using part affinity fields. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 7291-7299), https://doi.org/10.48550/arXiv.1611.08050.

Chen, Y. H., Wang, Y. Z. and Wong, M. H. 2001. A web-based fuzzy mass customization system. Journal of manufacturing systems, 20(4), 280-287, https://doi.org/10.1016/S0278-6125(01)80047-9

Fang, H. S., Xie, S., Tai, Y. W., & Lu, C. 2017. Rmpe: Regional multi-person pose estimation. In Proceedings of the IEEE international conference on computer vision (pp. 2334-2343), https://doi.org/10.48550/arXiv.1612.00137.

Güler, R. A., Neverova, N., & Kokkinos, I. (2018). Densepose: Dense human pose estimation in the wild. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 7297-7306), doi: 10.1109/CVPR.2018.00762.

Harding, J., & Brandt-Olsen, C. 2018. Biomorpher: Interactive evolution for parametric design. International Journal of Architectural Computing, 16(2), 144-163, https://doi.org/10.1177/1478077118778579.

Laura, T. 2021. Product customization and generative design. Multidiszciplináris Tudományok, 11(4), 87-95, DOI:10.35925/j.multi.2021.4.10.

Merriman, P. (2010). Architecture/dance: choreographing and inhabiting spaces with Anna and Lawrence Halprin. Cultural Geographies, 17(4), 427–449. http://www.jstor.org/stable/44251370

Pavlakos, G., Choutas, V., Ghorbani, N., Bolkart, T., Osman, A. A., Tzionas, D., & Black, M. J. 2019. Expressive body capture: 3d hands, face, and body from a single image. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 10975-10985), https://doi.org/10.48550/arXiv.1904.05866.

Preisinger, C. 2013. Linking structure and parametric geometry. Architectural Design, 83(2), 110-113, https://doi.org/10.1002/ad.1564.

Preston, C. J. 2014. Modernism's Dancing Marionettes: Oskar Schlemmer, Michel Fokine, and Ito Michio. Modernist Cultures, 9(1), 115-133.

Z.J.Zhang. 2013 Triangulation Chairs [J]. Bingfen, (101): 160

TRADITIONAL CHINESE VILLAGE MORPHOLOGICAL FEATURE EXTRACTION AND CLUSTER ANALYSIS BASED ON MULTI-SOURCE DATA AND MACHINE LEARNING

XIAO WANG[1], PENG TANG[2] and CHENYI CAI[3]
[1,2,3]School of Architecture, Southeast University, Nanjing 210096, China.
[2]Key Laboratory of Urban and Architectural Heritage Conservation (Southeast University), Ministry of Education, Nanjing, 210096, China.
[1]wangxiao_seu@seu.edu.cn, 0000-0001-9471-5370
[2]tangpeng@seu.edu.cn, 0000-0003-1658-6774
[3]caichenyi@seu.edu.cn, 0000-0002-6249-1816

Abstract. This study of traditional village morphology provides a possible entry point for understanding the growth patterns of settlements for sustainable development. This study proposes a hybrid data-driven approach to support quantitative morphological descriptions and to further morphology-related studies using open-source map data and deep learning approaches. We construct a dataset of 6819 traditional villages on the Chinese official list with geometrical, geographic and related no-material information. The images containing village buildings combined with roads or other environments are represented in binary to explore the integrated influence of these elements. The neural network is implemented to quantify the morphological features into feature vectors. After dimension reduction, cluster analysis is conducted by calculating the distance between the feature vectors to reveal five main types of Chinese traditional village patterns. The proposed method considers their overall spatial form and other factors such as size, transportation, graphical structure, and density. At the same time, it explores a framework using machine learning in the conservation and renewal work. And it also shows the possibility of data-driven methods for design and decision making.

Keywords. Cluster Analysis, Traditional Village, Morphology, Multi-source Data, Machine Learning, Rural Development.

1. Introduction

As China's urbanization process continues to grow, the conflict between urban and rural areas is becoming more pronounced. The study of traditional village morphology provides a possible entry point for better understanding the growth patterns of settlements for sustainable development. On the one hand, the formation of traditional villages is bottom-up self-organization and adaptation to environmental conditions,

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 179-188.

such as topography and climate. On the other hand, the development of traditional villages involves multiple social factors such as population, infrastructure equity, cultural identity and the continuation of local traditions (Yang and Pu, 2020). The morphological characteristics of the traditional village are the result of the interaction between human and nature. Therefore, the analysis of the morphological features of traditional villages cannot only be carried out from a tangible aspects perspective but also incorporates the intangible aspects mentioned above (UNESCO, 2009). From the data point of view, multiple data related to multiple factors need to be integrated (Liu et al., 2020). From a methodological point of view, it is important not only to identify phenomena but also to explore the underlying logic and inner connection of the space, form and other intangible aspects.

The study of traditional villages has always been an important part of urban design or settlement research. Concerning the research object, a remarkable study (Yang, 2014) has conducted a typological analysis of the spatial integrity characteristics of traditional settlements in China. Recent research has mainly focused on two aspects. From a geographical perspective, scholars (Li et al., 2020) treat villages as points and studying their spatial distribution characteristics in the entire region of China, focusing on location distribution rather than form. From a settlement form perspective, researches (Jiang and Kang, 2019) focus on the form of a village in a specific region and background, usually taking the internal and external space characteristics of a single village as a case. Concerning the research content, many studies also explore the relationship between material space and human society. For example, analysis of the coupling relationship between family composition and internal space of the village is conducted to reveal the correlation between traditional village spatial structure and family hierarchy structure (Yang and Cai, 2020) and conclude the intrinsic laws and integration methods of village space changes (He and Deng, 2011). At a larger scale, some studies discussed the relationship between regional village distribution and diseases (Xia et al., 2017), poverty (Wang et al., 2019) and development potential (Nusrang et al., 2022), revealing the factors affecting village distribution that lead to these social impacts.

In addition, quantitative analysis, as an intuitive path to explain spatial characteristics, is increasingly being used to interpret rural spaces. Some scholars analyse the physical space of a village through relational graphs and a series of parameter values (Yang, 2020). Based on the spatial syntax (Hillier et al., 1976), researchers (Cao et al., 2016) analyse the accessibility of the village through the calculation of global integration and spatial depth, and compare it with the current spatial characteristics and actual frequency of use. Alternatively, a new quantitative index system is established to describe the spatial characteristics of the countryside by quantitative indices of spatial structure, planar form, and architectural order to calculate the clustering strength of the building node network (Pu et al., 2020). Furthermore, parameterized analysis and reconstruction (Tang et al., 2019), machine learning (Alvarez and Ochoa, 2020), especially deep learning (Cai et al., 2021) more used in urban problem analysis, can also be referenced in research.

This study proposes a hybrid data-driven approach to support quantitative morphological descriptions and to further morphology-related studies using multi-source data and deep learning approaches. We construct a dataset of traditional villages on the Chinese official list with both geographic information and related intangible information. The images containing village buildings combined with roads or other environments are represented in binary to explore the integrated influence of these elements. The neural network is implemented to quantify the morphological features into feature vectors (Moosavi, 2017). After dimension reduction, cluster analysis is conducted by calculating the distance between the feature vectors (Cai et al., 2021). This study also involved three different algorithms (Krizhevsky et al., 2017) combinations in vector extraction and clustering for comparison and verification.

This study reveals five main clusters of Chinese traditional village patterns, which show their differences in location, terrain, boundary form and development level. And we tried to summarize the underlying reasons and drivers for these differences. At the same time, it explores a framework using data-driven method like cluster analysis in the recognition and conservation of human settlement. And it also shows the possibility of machine learning tools for design and decision-making.

2. Methods

2.1. DATA SOURCE

We construct a dataset of 6819 traditional villages (Figure 1) on the Chinese official list with both geographic information such as satellite images and elevation information, and related intangible information such as population and production type (Table 1). The satellite images and geographical maps of sample villages are obtained in batches through APIs. Sources of data are Google Map, DigitalGlobe, OpenStreetMap, Tianditu, etc. Then these images can be saved locally and imported as a database for subsequent experiments. The social data such as population, income and ethnicity are gathered from the official website of Traditional Chinese Village Digital Museum (Traditional Chinese Village Digital Museum, 2021).

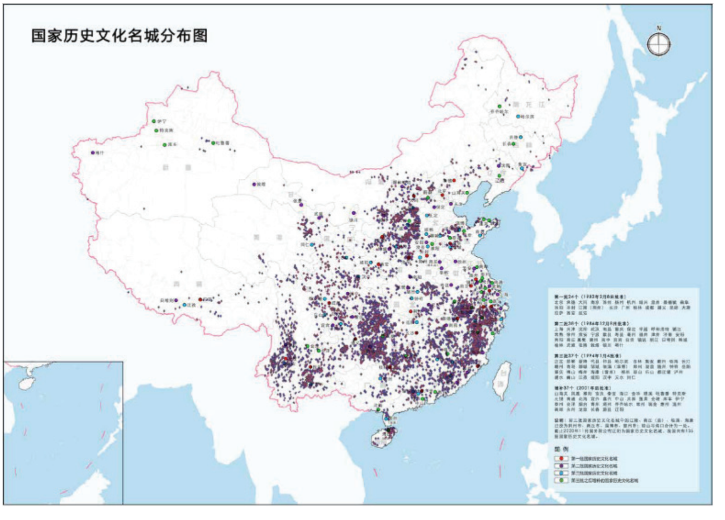


Figure 1. Spatial distribution of 6819 Chinese traditional villages in the official list.

Table 1. Information of villages from Traditional Chinese Village Digital Museum

Content	Attributes					
No.	1	2	3	4	…	6819
Name	Xiye	Zecheng	Taoyang	Shifo		Zhangtong
Longitude	112.4571	112.2964	112.8869	123.3486		117.5935
Latitude	35.32812	35.51886	37.02527	42.1333		31.60173
Location	Dongye Yangcheng Jincheng Shanxi	Gulong Yangcheng Jincheng Shanxi	Yuntu Yushe Jinzhong Shanxi	Shifosi Shenbei Shenyang Liaoning		Huanglu Chaohu Hefei Anhui
Attributes	administrative village	administrative village	administrative village	administrative village		natural village
Age of formation	Yuan	Yuan	Yuan	Yuan		Ming
Altitude (m)	1125	706	720	1240		20
Lang area (km²)	12	2.91	3.57	5		2.74
Building area (km²)	0.506667	0.146667	1.5	0.033333		0.32
Topography	mountain	hill	hill	hill & plain		hill
Household population	2098	1083	1700	186		2142
Resident population	1200	1300	1430	130		2100
Total annual income (CNY)	120000	2000000	22000	0		550000
Average annual income (CNY)	8000	8900	4200	500		15600
Main ethnic groups	Han	Han	Han	Han		Han
Main industries	plantation, aquaculture	agriculture, labor	plantation	agriculture		tourism, education, plantation

2.2. SPATIAL OVERLAY

Image analysis at the same scale provides not only the qualitative insight into the morphology of the settlement itself (e.g. clumped, linear, dispersed, etc.) but also a synthetic perspective of the size, the degree of development (density of the road network), and corroborated by the corresponding population, economy and location. Considering the integrated influence of these elements, the footprints of village buildings combined with roads or other environments are overlaid in one image. All the images are represented in binary to next features extraction (Figure 2).

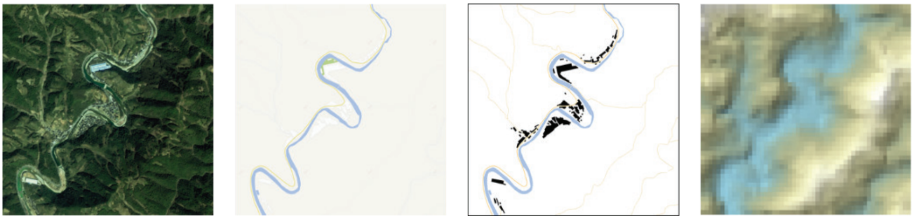


Figure 2. Multisource data of traditional villages.

2.3. CLUSTERING BASED ON BINARY IMAGES

The image-based clustering analysis contains feature vector extraction and clustering using neural network in Mathematica platform. The neural network is implemented to quantify the morphological features into feature vectors. After dimension reduction, cluster analysis is conducted by calculating the distance between the feature vectors.

This study compares different algorithms: 1) NumericVector method for feature vector extraction and KMeans for clustering, 2) SENet for feature vector extraction and KMeans for clustering, and 3) CapsNet for feature vector extraction and GaussianMixture for clustering. The "NumericVector" method will typically convert examples to numeric vectors, impute missing data, and reduce the dimension using Dimension Reduction. CapsNet and SENet are neural networks for image identification. KMeans and GaussianMixture are algorithms for clustering.

3. Analysis and Results

3.1. DATA PROCESSING AND FEATURE VECTORS EXTRACTION

Due to the lack of data related to Chinese traditional villages, accurate village boundaries cannot be obtained directly. The vector maps can be used to obtain village outlines as well as the roads, waters and terrains. For example, to analyse village patterns related to transportation, image just contains building and road information (Figure 3).

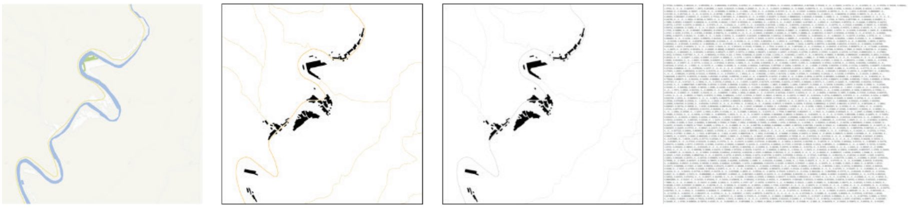


Figure 3. Image processing and feature vectors (partly).

Based on the results of the feature vectors extraction it is possible to search for samples of settlements with similar morphology to the selected one (Figure 4) by calculating the distance between the feature vectors. The clustering results also provide a reference for the analysis of the relationship between village building and their environment.

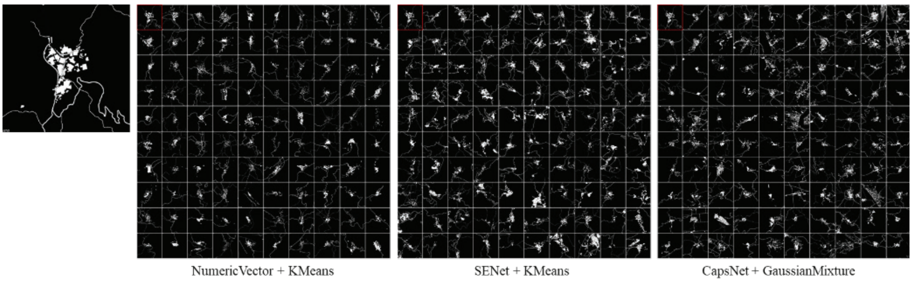


Figure 4. Similarity analysis and research.

3.2. DIMENSION REDUCTION AND CLUSTER ANALYSIS

After dimension reduction, cluster analysis is conducted by calculating the distance between the feature vectors using the pre-trained models in Mathematica. The result as feature plots illustrates unsupervised clustering distribution (Figure 5). SENet - KMeans method is more reliable as the result shows gathering and smooth transitions, while NumericVector – Kmeans method cannot get obvious clusters and the clusters of CapsNet – GaussianMixture method is too discrete compared with the real situation.

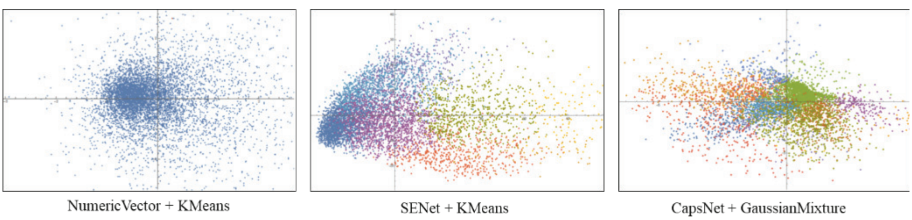


Figure 5. Clusters of feature vectors after dimension reduction.

Based on the SENet - KMeans method, the village images are located in the feature space plot according to their feature distances after visualization (Figure 6).

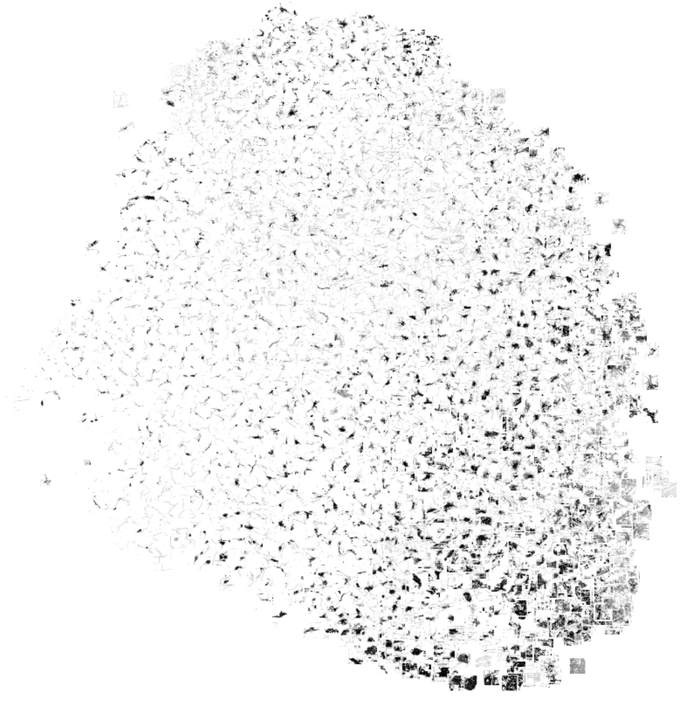
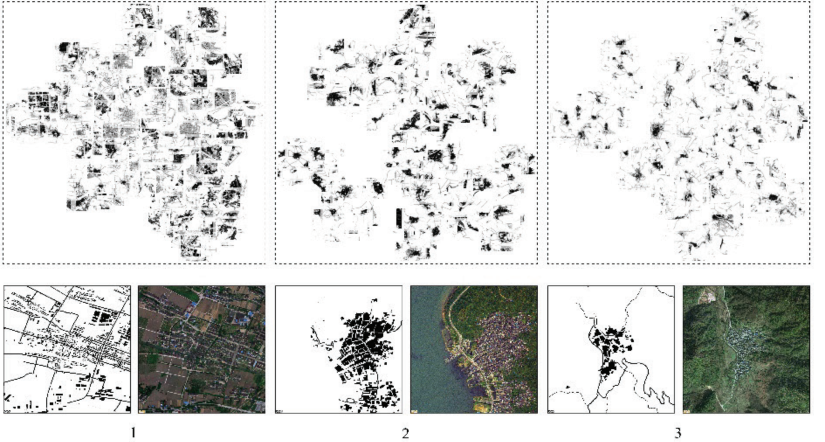


Figure 6. Feature space plot based on the feature distance.

Until now the study reveals five main types (Figure 7) of Chinese traditional village patterns referring to the relation between village shape and roads:



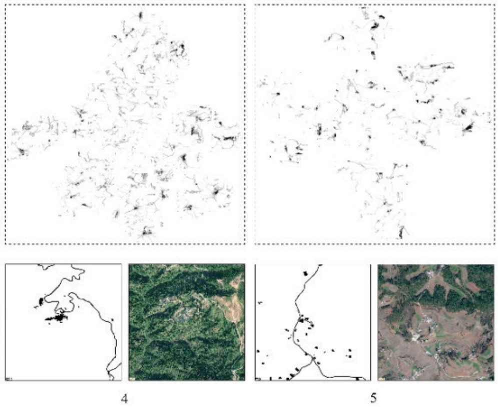


Figure 7. Feature space plots of 5 village clusters and typical village.

- Type 1 village is located in a plain area with a developed road network, surrounded by farms or pools. It is larger but has a blurred boundary. Most of these villages are located in rich areas like southeast coastal provinces and agricultural regions in north.
- Type 2 village is located in a plain or hilly area with a dense river and road network. It is of medium size with a clear boundary. The building density is high. Most of these villages are located in developed provinces in the Yangtze River Delta region.
- Type 3 village is located in a hilly or valley area with several main roads. It is in small size and has a clear boundary restricted to the terrain. Most of these villages are located in developing regions in central and southwest China.
- Type 4 village is located in a valley or mountainous area with one or two roads. It is in small size with dispersed settlements. Most of these villages are located in the mountains in less developed regions.
- Type 5 village is located in a hilly or mountainous area with one or two roads. It has scattered settlements or buildings with little population. Most of these villages are located in the mountains regions or northwest highland.

Compared with the traditional manual classification, previous studies made qualitative generalisations about the morphology of traditional Chinese villages in a general cognition, while classification using machine learning can reflect the transition and gradations of villages responding to the change of tangible and intangible environment. The result is not just a graphic composition of forms like linear, circular, concentrated or scattered shape. It provides a detail view to show the exact features more than a diagram description, which is helpful to protect and continue the villages' unique morphology and culture features in human-centric view.

4. Conclusion and Discussion

The proposed method considers their overall spatial form and other factors such as size, transportation, graphical structure, and density

. At the same time, it explores a framework using machine learning like cluster analysis in the conservation of traditional villages. Compared to existing comparative studies of villages within some limited regions, this paper is based on a less common method using multi-source detail data in such a large scale for clustering.

One reason is that with China's new stage of urbanization, rural development needs higher-level resource allocation and industrial layout perspectives. Although there are differences in form among villages in different regions, there is a reference value and research value in the characteristics of territorial space development. Another reason is that it is easier to collect detailed vector maps for individual villages in a small region and focus on local village features. However, collecting large-scale rural data in China is still difficult but necessary, and digital technology provides the possibility of processing and integrating massive multi-source and multi-modal data.

But the results are influenced by factors such as the robustness of the model, the quality of training data. A more straightforward approach would be to use semantic segmentation to process satellite images then obtaining more accurate morphological information. In the future, one possible direction is to utilize semantic segmentation. In addition to land classification, road extraction and building edge extraction are also necessary. These attempts can better assist in further quantitative analysis of village morphology.

In the design practice, another direction of future work is using deep learning and neural network to predict the boundary expansion and spatial texture continuation of settlement growth according to the terrain and social information. Besides the image-based and text-based recognition and clustering, graph and graph neural network (GNN) analysis and prediction can also be used for computer vision, natural language processing, recommendation systems, and social network analysis in more complex urban and settlement research. The intended research will explore the possibility of data-driven methods used directly for design scheme generation and decision-making.

Acknowledgements

This work was supported by the National Natural Science Foundation of China under Grant 52178008 and Postgraduate Research & Practice Innovation Program of Jiangsu Province under Grant 5001002202.

References

Alvarez-Marin, D., & Ochoa, K. S. (2020). Indexical Cities: Articulating personal models of urban preference with geotagged data. *arXiv preprint*, arXiv:2001.10615, https://doi.org/10.48550/arXiv.2001.10615

Cai, C., Guo, Z., Zhang, B., Wang, X., Li, B., & Tang, P. (2021). Urban morphological feature extraction and multi-dimensional similarity analysis based on deep learning approaches. *Sustainability*, 13 (12), 6859. doi: 10.3390/su13126859.

Cao, J., Wang, J., & Wang, Y. (2016). Interpretation of village space in Yuanjia village, Lique, Shaanxi from a syntactic perspective. In *Proceedings of the Third National Conference on Rural and Town Planning Theory and Practice and the Second Countryside Architecture Seminar* (pp. 281-290).

He, Y., & Deng, W. (2011). Village space based on the dominant surname family: A case study of Suzhuang, Shanxi. Architecture Journal, 519(11), 11-15.

Hillier, B., Leaman, A., Stansall, P., & Bedford, M. (1976). Space syntax. *Environment and Planning B: Planning and design*, 3(2), 147-185, https://doi.org/10.1068/b030147.
Jiang, M., & Kang, J. Y. (2019). Research on the countermeasures of the protection and development of the environmental street space and architectural style and features: Taking Yunnan Wande traditional village as an example. *Research on Heritages and Preservation*, 4(03), 59-62. doi:10.19490/j.cnki.issn2096-0913.2019.03.015.
Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60 (6), 84-90. doi: 10.1145/3065386.
Li, Y., Yao, W., Zhang, Y.K., & Peng, W. (2020). Spatial distribution characteristics of traditional Chinese settlements. *China Cultural Heritage*, 04, 51-59.
Liu, C., Cao, Y., Yang, C., Zhou, Y., & Ai, M. (2020). Pattern identification and analysis for the traditional village using low altitude UAV-borne remote sensing: multifeatured geospatial data to support rural landscape investigation, documentation and management. *Journal of Cultural Heritage*, 44, 185-195. doi: https://doi.org/10.1016/j.culher.2019.12.013.
Moosavi, V. (2017). Urban morphology meets deep learning: Exploring urban forms in one million cities, town and villages across the planet. *arXiv preprint*, arXiv: abs/1709.02939.
Nusrang, M., Aidid, M. K., & Rais, Z. (2022). K-Means Cluster Analysis for Grouping Districts in South Sulawesi Province Based on Village Potential. *ARRUS Journal of Mathematics and Applied Science*, 2(2), 73-82.
Pu, X., Wang, Y., & Huang, Q. (2020). Analysis of quantitative methods of obtaining the boundary shape of rural settlements. *Architecture & Culture,* 12, 189-191. doi:10.19875/j.cnki.jzywh.2020.12.067.
Stojanovski, T. (2022). Urban morphology and artificial intelligence. *Urban Morphology*, 26(1), 78-88, DOI: 10.51347/UM26.0005.
Tang, P., Wang, X., & Shi, X. (2019). Generative design method of the facade of traditional architecture and settlement based on knowledge discovery and digital generation: a case study of Gunanjie Street in China. *International Journal of Architectural Heritage*, 13(5), 679-690.
Traditional Chinese Village Digital Museum. Retrieved December 16, 2021, from http://www.dmctv.cn/
UNESCO. (2009). *Hoi An Protocols for best conservation practice in Asia: professional guidelines for assuring and preserving the authenticity of heritage sites in the context of the cultures of Asia*. UNESCO Bangkok.
Wang, Y., Liang, C., & Li, J. (2019). Detecting village-level regional development differences: A GIS and HLM method. *Growth and Change*, 50(1), 222-246, https://doi.org/10.1111/grow.12275.
Xia, C., Bergquist, R., Lynn, H., Hu, F., Lin, D., Hao, Y., ... & Zhang, Z. (2017). Village-based spatio-temporal cluster analysis of the schistosomiasis risk in the Poyang Lake Region, China. *Parasites & vectors*, 10(1), 1-10, https://doi.org/10.1186/s13071-017-2059-y.
Yang G., & Cai Y. (2020). Correlation between spatial structure and familial society of Wuyan ancient village, Huangyan district of Zhejiang province. *Planners*, 03, 58-64.
Yang, G. (2014). The integrated characteristics of space and social significance of traditional Chinese settlement. *Tongji University Journal Social Science Section*, (3), 60-68.
Yang, X. (2020). Advances in Quantitative Research Methodologies for the Spatial Layout of Rural Settlement in Recent 20 Years. *Urban Planning International*, 35(04), 72-80. doi:10.19830/j.upi.2018.340.
Yang, X., & Pu, F. (2020). Spatial cognitive modeling of the site selection for traditional rural settlements: a case study of Kengzi village, southern China. *Journal of Urban Planning and Development*, 146 (4), 05020026. doi: 10.1061/(asce)up.1943-5444.0000616.

GESTURE MODELING

In Between Nature and Control

CHOR-KHENG LIM1
[1]*Department of Art and Design, YuanZe University.*
[1]*kheng@saturn.yzu.edu.tw, 0000-0003-0990-5191*

Abstract. This research develops a robotic arm-controlled wax material fluidity shape/form-making system that combines AI image recognition and remote control capabilities. AI image recognition is used to capture the designer's hand gestures or movements as an intuitive input method for controlling shape forming. It gives the designer active control of the desired model shape through the temperature, rotation, and movement controls in the process of phase change of wax materials. The system has the additional advantage of remote control, allowing designers to create wax models remotely through intuitive gesture-controlled interactions. This advantage will become a new method of the digital co-creation process.

Keywords. Form-finding, Gesture Modeling, Wax, Phase Change Material, Robotic-arm Controlled.

1. Introduction

As early as before the computer age, one notable example of the use of form finding is the work of Spanish architect Antoni Gaudí, who employed an upside-down force model to explore the geometrical forms of catenary arches. Based on motion dynamics and gravity, this method allowed Gaudí to invent his architectural languages and achieve his goal of perfecting and surpassing the Gothic style, as seen in his projects such as the Basilica and Expiatory Church and the Holy Family and Church of Colònia Güell (Browne, 2008; Huerta, 2006).

The experimental work of Antoni Gaudí and Frei Otto has laid the groundwork for what is now known as form-finding. This methodology relies on natural rules and physical forces to promote the principles of transformation through the relationship between form, material, and structure ((Lopes et al., 2014).

In the digital era, researchers utilized computational design methods in form-finding processes, such as 3D modeling, parametric design, mathematical algorithm, robotic arm-controlled design, etc (Williams, 2022). Zexin and Mei (2017) explored the relationship between architectural drawings and form-finding and argued that architects should incorporate traditional architectural drawings into their form-finding processes. They discussed projection systems and their impact on architectural design and described an experiment using a robotic arm to establish a cylindrical projection form-finding system. Patiño (2019) defined form-finding techniques as strategies that are based on rules and physical principles, and that use the relationship between form, material, and structure to generate transformative results. These techniques typically involve algorithms that take input, process variables, and produce output or results (Torreblanca-Díaz et al., 2019).

Wax material is often used in the making of a model as sculptural shapes due to its material properties being easy to sculpt. In recent years, it is also commonly used in the material application of 3D-printed test models. Many artists also apply the phase-change feature of wax materials, from liquid to solid, in a creative process.

Artist Wu Kai Xun (Wu, 2016) explored the relationship between form and material. He proposed a "Fluidity Formation," forming process using the physical property of phase change with wax materials. This process allows the material itself to have the possibility of creating its shape, providing the new fluidity forms of the architectural space. This technique is based on the material properties of these substances and involves heating the materials to different temperatures to achieve the desired effects. The process of Fluidity Formation consists of three stages. In the first stage, wax is used to create a mold for the final sculpture or structure; in the second one, plaster is poured over the wax mold; and in the third one, the materials are heated until the wax melts, and the plaster hardens. This process results in a unique and dynamic final product, with voids, openings, and transparency created by the flow of the materials. Artist Wu Kai Xu used this technique to create a series of lamps, called the Fluid Lamp Series. These lamps are based on Fluidity Formation principles and utilize unique material qualities of plaster and wax to create dynamic visual effects. When the light passes through the voids and openings created by the material, it creates a dynamic and unpredictable visual experience. Fluidity Formation is an innovative and exciting approach to using fluid materials in architecture and sculpture. Its unique material qualities and unexpected results make it a valuable tool for artists and designers.

Architect Monica Sanga considered casting as a dynamic formation process, using water as an infinitely flexible formwork (Sanga, 2012). She investigated the various formal results of casting through a phase change. By using wax as the casting material, Monica was able to investigate the potential for dynamic and flexible forms. The use of additives in the wax also produced a range of failures and desirable results. The process employed by Monica was inherently experimental and accidental, resulting in unique and one-of-a-kind pieces. Through the use of water as formwork and wax as the casting material, Monica created dynamic and flexible forms that push the boundaries of traditional casting techniques.

In addition, some designers used wax forming as a structural design. In her project Waxploration, German designer Katharina Gross used a mixture of molten wax and brass boxes to create a series of tables with stalactite-like legs (Gross, 2014). The wax mixture she used was a strong and durable variant that included marble dust, and she employed a candle-making technique to create the furniture pieces. This unique and creative approach to furniture design showcases the potential of wax as a sculptural material.

1.1. PROBLEM AND OBJECTIVE

The form generation in the phase-changed wax design processes is based on the characteristics of the material itself, and the form is automatically generated through a natural and abstract fluid-forming process. Many artists have found wax, a phase-changed material, to be fascinating to work with in this process, but the lack of control in the automatic shape-generation process can make it difficult to predict the final shape and can thus lengthen the creation process. Therefore, they need to spend more time to create a form that they are satisfied with.

This process is different from digital modeling, where the appearance of the forms can be controlled according to design conditions. In contrast, in the 3D printing process that also uses wax as a molding material, a 3D model must be created first, which is a more specific and accurate forming design process. In the process of fluidity form finding, both abstract and organic form shapings are required, as well as more concrete form control due to functional requirements.

Therefore, this study aims to incorporate the phase-change forming properties of wax materials into an intuitive and controllable operation for the fluidity shape-making process. By doing so, it is hoped that the fluidity form-finding process in a shape and form design can be more predictable, efficient, and functional. This research develops a wax-based robotic arm-controlled fluidity shape/form-making system that combines AI image recognition and remote control capabilities, allowing artists and designers to create wax models remotely through intuitive gesture control interactions.

2. The Development of the Gesture Modeling System

Some researchers use natural body movements such as hand gesture as the input interface for 3D modeling (Choul, Gao, Huang, & Fuh, 2013; Kang, Zhong, Qin, Wang, & Wright, 2013). Choul et al. (2013) introduced a 3D conceptual design system that aims to enhance the creative process for designers. The system utilizes hand gestures and motion capture as its user interface to facilitate fast and unobstructed expression of design ideas in 3D. The system is designed to be user-friendly and intuitive with hand gestures optimized for ease of use and real-time recognition. The framework and components of the real-time 3D conceptual design and visualization system are presented.

This research develops the gesture-modeling system, which meets the condition of "in between nature and control" based on the three main steps of form-finding techniques proposed by Patiño (2019). Figure 1 shows the framework of the system. The first step of the system involves the use of a camera to capture the user's bodily gestures as input for the system. An AI image recognition system then processes this input in the second step, which uses previously-trained body gesture data to identify the gestures and send control signals to the remote robotic arm via MQTT. The control parameters of the robotic arm are then used to operate the wax model in the third step. This system operates in a loop because the user watches the wax model taking shape while controlling the bodily gestures to achieve the remote, real-time production of the wax model. During this process, an online web interface or video conference can be used for visualization purposes.

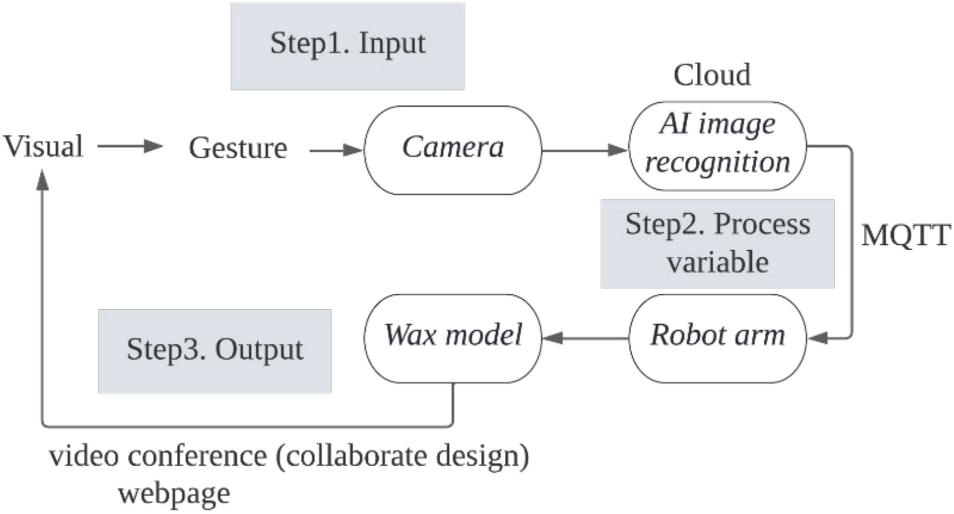


Figure 1. The framework of the Gesture Modeling system

The three steps are described as follows:

2.1. GESTURES AS INPUT PARAMETERS

In this research, two tests and experiments were conducted using gestures as input parameters. The first type was body gestures, and the second, hand gestures. The body gestures involve whole-arm movements such as lifting the left hand, lifting the right hand, placing the left hand in front of the chest, placing the right hand in front of the chest, and crossing the left and right hands. These movements were used to control the movements of the remote robotic arm, including clockwise and counterclockwise rotation, forward and backward movement, and opening and closing an electric fan. This was primarily used to control the shaping of the wax model. The hand gestures involve moving the left and right hands across the black-circle pattern on the table, indicating three electric fan positions (Figure 2), similar to a pulling motion, to control the switching of the remote electric fans and the solidification of the liquid wax to a solid state.

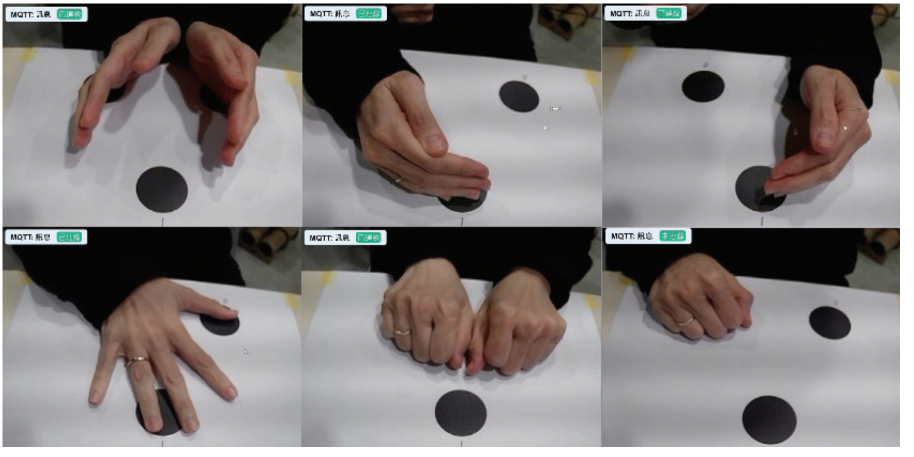


Figure 2. The different hand gestures as input parameters

2.2. AI IMAGE RECOGNITION

In this study, a cloud-based AI pre-trained model is used for gesture recognition. For the body gesture, the Node-RED platform is used to connect to Posenet to recognize the user's arm movements. The operation is fast and accurate, suitable for any user. In addition, for the hand gesture, the image recognition training of the Teachable Machine is used for hand gesture recognition. Because the hand gestures can be easily recognized by working with the table fan position map, even different users can recognize them. In addition, a simple Scratch AI block is used as a cloud computing platform for the hand gesture experiment.

3. Case Studies

Figure 3 shows the design of the robotic arm control devices for producing the wax model. The device has a container filled with 1000cc of wax and 1300cc of water, maintained at a temperature of 75 degrees to keep the wax in a liquid state. Three electric fans are placed around the container to control the cooling. When the surface temperature of the wax drops to 60 degrees, a thin film of solid wax will begin to form. At this point, the robotic arm will move to the surface of the thin film, using the wax-coated element on the base, to slowly lift the thin film. The base which is fixed on the arm can be rotated and moved forward and backward, and the switch of the electric fan can also be controlled remotely.

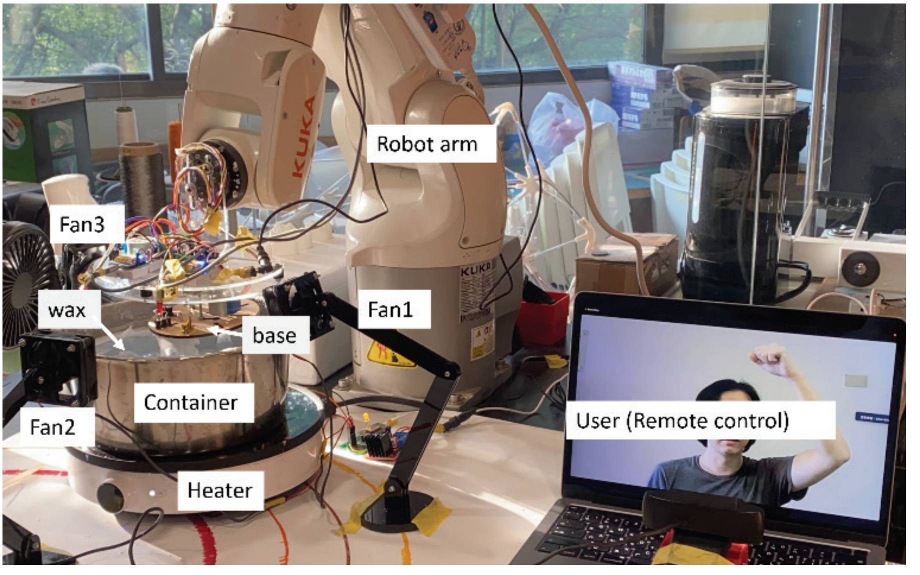


Figure 3. The control system of the robot arm for making a wax model

In this stage, seven different experiments were conducted with various parametric flow control conditions. As shown in Figure 4, based on the input conditions of the first type of body gesture, three parametric flow experiments were performed. The parameters for these three experiments were fixed as follows: 1. the speed at which the robotic arm was raised; 2. the fan remained on; and 3. the number of wax-coated elements on the base was eight. The variable parameters were set as follows: 1. The base can be rotated 360 degrees in both the clockwise and counterclockwise directions; and 2. The base can be translated back and forth.

As shown in Figure 5, based on the second input conditions, namely hand gestures, four parametric flow experiments were performed. The parameters for these four experiments were fixed as follows: 1. the speed at which the robotic arm was raised; and 2. continuous rotation. The variable parameters were set as follows: 1. the number of wax components on the base was either 8 or 3; and 2. fan switch control.

	Controlled factor		Experiment1	Experiment2	Experiment3
Pulling up the thin film	Robot arm	lifting	✓	✓	✓
		rotation	✓	x	✓
		pan	x	✓	✓
Solidification of the liquid wax	Fan	3 (switch on)	✓	✓	✓
Base	Wax-coated element	8	✓	✓	✓
Model				 	

Figure 4. The experiments with parameters of body gestures input

	Controlled factor		Experiment4	Experiment5	Experiment6	Experiment7
Pulling up the thin film	Robot arm	lifting	✓	✓	✓	✓
		rotation (continuous)	✓	✓	✓	✓
Solidification of the liquid wax	Fan *3	switch on continually	✓	✓	x	x
		switch on sequentially	x	x	✓	x
		random controlled	x	x	x	✓
Base	Wax-coated element	8	x	✓	✓	✓
		3	✓	x	x	x
Model						

Figure 5. The experiments with parameters of hand gestures input

3.1. OUTPUT- NATURAL AND CONTROLLED WAX MODEL

With the Gesture Modeling system developed in this study, the results of seven experimental parameter tests show natural and rule-specific wax models. Although the wax models produced in the same experimental conditions are not exactly the same, the same pattern rules can be observed from the produced models. Figure 6 shows the pattern rules that were inferred. The logic and rules inferred from the modeling patterns in this research can effectively allow designers to control parameters, producing the form of the natural fluidly wax models, closer to the desired shape in the mind. At the same time, due to the adjustment of experimental parameters, various spatial and structural-like models can be produced, so they are suitable for artists and architectural designers to use in the early concept design phase.

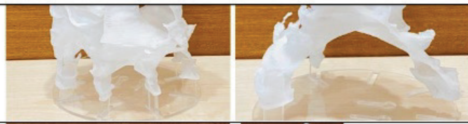
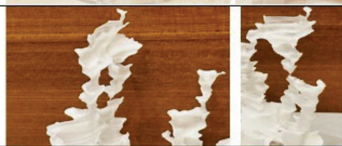
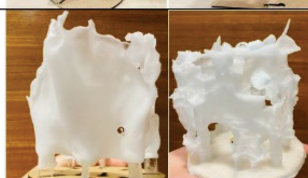
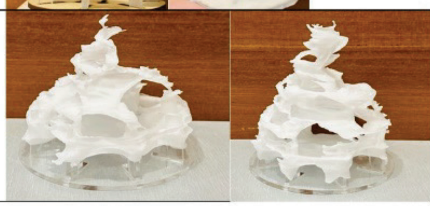
	pattern	Wax model
1	arched feet	
2	Floral-like struts	
3	Spiral-like Line	
4	Hole-like	
5	Funnel- like	
6	Cylindrical Surface	
7	Spiral-like layered surface	

Figure 6. The pattern rules produced from the Gesture Modeling system

4. Results

This research develops a wax-based robotic arm-controlled fluidity shape/form-making system that combines AI image recognition and remote control capabilities. AI image recognition is used to capture the designer's hand gestures or movements as an intuitive input method for controlling shape forming. An intuitively-controlled behavior is like the process of ceramic hand-pulling. The robotic arm is used as a control mechanism to control the forming of wax models. It gives the designer active control over the desired model shape through the temperature, rotation, and movement controls in the process of phase change of wax materials. The control device on the robotic arm will activate the corresponding moving mechanism and temperature control according to the received AI image information to form the shape of the model with a curved wax surface (Figure 7).

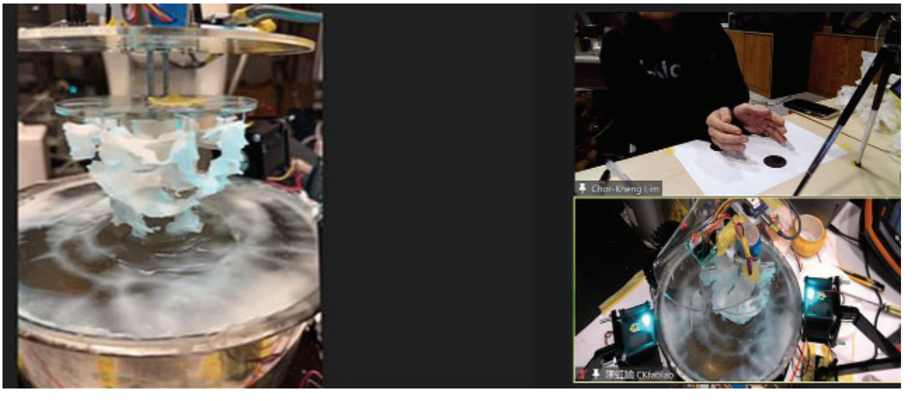


Figure 7. The model-making process using Gesture Modeling system

The system only takes 5-10 minutes to complete a wax fluidity surface model with a height of 30 cm and a diameter of 10 cm. The making process is faster than 3D printing. At the same time, it has a controllable surface to form the main appearance and the organic shape of the wax during the phase change process. In addition, the system has the additional advantage of remote control, allowing designers to create wax models remotely through gesture control interactions. This advantage will become a new method of the digital co-creation process. Since wax is a phase-changed material, it also has the benefits of recycling in the study model stage of the design process.

5. Future Studies

Future research hopes to conduct user testing on the developed system, record the operation process of the designer or artist using the system, analyze usability, and improve system operation and design. In addition, this research also hopes to develop the system into a design tool for creators to create physical forms.

References

Browne, C. (2008) Gaudí's organic geometry. Computers & Graphics 2008, 32, 105-115, https://doi.org/10.1016/j.cag.2007.09.005

Choul, D.; Gao, Z.; Huang, S.; Fuh, C.-S. (2013). HAND GESTURE RECOGNITION, TRACKING, AND PROJECT TO A 3D MODE. In Proceedings of the IPPR Conference Faircloth, B., Welch, R., Sinke, Y., Tamke, M., Nicholas, P., Ayres, P., Eherenbard, E., & Ramsgaard Thomsen, M. (2018). Coupled Modeling and Monitoring of Phase Change Phenomena in Architectural Practice. In T. Rakha, M. Turrin, D. Macumber, F. Meggers, & S. Rockcastle (Eds.), 2018 Proceedings of the Symposium on Simulation for Architecture and Urban Design (pp. 81-88)

Gross, K. (2014) *Katharina Gross adapts candle-making process to create wax tables* . dezeen from https://www.dezeen.com/2014/07/25/katharina-gross-waxploration-wax-tables-candle-making/

Huerta, S. (2011) Structural Design in the Work of Gaudi. Architectural Science Review 2006, 49, 324-339. https://doi.org/10.3763/asre.2006.4943

Kang, J., Zhong, K., Qin, S., Wang, H., & Wright, D. (2013). Instant 3D design concept generation and visualization by real-time hand gesture recognition. Computers in Industry, 64(7), 785-797. https://doi.org/10.1016/j.compind.2013.04.012

Lopes, J.V.; Paio, A.C.; Sousa, J.P.(2014). Parametric Urban Models Based on Frei Otto's Generative Form-Finding Processes. In Proceedings of the Proceedings of the 19th Conference on Computer Aided Architectural Design Research in Asia (CAADRIA). pp. 595-604. https://doi.org/10.52842/conf.caadria.2014.595

Sanga, M. (2012) *Studies in wax casting*. Sanga Architecture and Design from https://www.sanga.design/projects/studies-in-wax-casting.

Torreblanca-Díaz, D.A.; Patiño, E.; Valencia-Escobar, A.; Urdinola, D. (2019). Form-finding methodology as strategy for formative research in industrial design education Experimental techniques for the early creative phases of the product design process. In Proceedings of the Blucher Design Proceedings; pp. 45-54.

Williams, K.(2022) Form-Finding, Architecture and Mathematics. Nexus Network Journal, 24, 1-3, https://doi.org/10.1007/s00004-021-00588-3.

Wu, K-X. (2016) *Fluid Lamp*. architizer from https://architizer.com/projects/fluid-lamp/.

Zexin, S.; Mei, H. (2017) Robotic Form-Finding and Construction Based on the Architectural Projection Logic. IOP Conference Series: Materials Science and Engineering, 216. DOI 10.1088/1757-899X/216/1/012058

OPTIMISING THE CONTROL STRATEGIES FOR PERFORMANCE-DRIVEN DYNAMIC BUILDING FACADES USING MACHINE LEARNING

YUANYUAN LI[1], CHENYU HUANG[2] and JIAWEI YAO[3]
[1]*School of Architecture, Qingdao City University*
[2]*School of Architecture and Art, North China University of Technology*
[3]*College of Architecture and Urban Planning, Tongji University*
[1]*yuan93li@hotmail.com, 0000-0002-6761-8169*
[2]*huangchenyu303@163.com, 0000-0002-6360-638X*
[3]*jiawei.yao@tongji.edu.cn, 0000-0001-7321-3128*

Abstract. The balance between energy consumption and indoor environmental comfort is a continuing research topic in building energy efficiency. The dynamic façades (DF) are considered a practical approach to separate the sun and create more shadows for buildings with curtain walls, reducing the HVAC system's energy consumption. However, the design complexity of the DF leads to a time-consuming simulation process, making it difficult to modify the design parameters in the early design stage efficiently. This paper provides optimized control strategies for four dynamic façade prototypes. We use explainable machine learning to explore the relationship between design parameters of DF and indoor performance, including Energy Use Intensity (EUI) and Daylight Glare Probability (DGP). We deployed the trained model in optimizing the rotation angle of DF per hour on a typical day to minimize the EUI and DGP of the target room. The results show that the rotation angle of DF significantly affects the DGP, whereas the room size affects EUI performance more than rotation angles. Optimized control strategies of DF bring a maxim 13.5% EUI decrease and 51.7% reduction of DGP. Our work provides a generalizable design flow for performance-driven dynamic skin design.

Keywords. Dynamic Façade, Energy Consumption, Indoor Comfort, Computational Simulation, Multi-objective Optimization, Machine Learning

1. Introduction

Global warming has increased the probability of extremely high temperatures in different regions, which requires more energy consumption to maintain indoor comfort for commercial buildings with curtain walls. As a chain reaction, increasing carbon emissions also further aggravate environmental problems. Dynamic façades (DF) can be designed with an environment-driven performance approach, balancing different light indicators to maintain a comfortable indoor light environment. Compared to static building facades, DF can better adapt to complex climate conditions and respond in real-time, shade the exterior of the building, and produce more shadows which will reduce the energy consumption of air conditioning. On the other hand, DF can replace traditional building facades to respond to solar movement and climate change, creating a stable level of visual comfort for the interior.

However, the undecided variables of DF (such as shape, size, materials, and rotation angles) create multiple possibilities in the early design phase, which leads to a complex and time-consuming simulation process for predicting its performance. Therefore it is difficult to modify the design parameters based on the DF performance results in the early design phase. In addition, the existing method of indoor light performance often requires a complex computational simulation process, and even high-performance computers cannot select the best results in a short period through exhaustive computation. If the design team lacks researchers in this direction, it is challenging to conduct performance evaluation and prediction in the early design stage, which often turns the design process into a post-build evaluation. This study introduced interpretable machine learning to improve the speed of DF performance simulation with efficient simulation prediction models and optimize this design multi-objective optimization process to provide recommendations for pre-design decisions.

2. Related works

2.1. DYNAMIC FACADES & CONTROL STRATEGY

As a mediator of positive response between the changing environment and the building interior, the definition of DF has uncertainty, such as "Intelligent Skin" and "Climate adaptive shell". Several studies have highlighted the potential of adaptive skin in maintaining building interior comfort, coping with climate change, and reducing energy consumption (Loonen et al., 2013). DF has multiple potentials to enhance building performance with different design priorities: for example, DF designed with interactivity and bionics as a priority (Romano et al., 2018), DF designed with user comfort as a design priority (Luna-Navarro et al., 2020; Tabadkani et al., 2021), or DF designed with energy efficiency as a priority (Favoino et al., 2014). The design strategy of DF often aims at different goals, how to balance those goals and achieve multi-objective optimization are the main difficulties of DF control strategies. Existing control strategies for the DF strategies include open-loop, closed-loop, and agent model simulation. This paper proposes an agent model simulation approach to find the optimal solution for the DF control strategy.

2.2. ENERGY & LIGHTING SIMULATION

The primary reference for assessing the building performance criteria is the implementation of international green building standards, such as LEED and BREEAM. These green building standards establish quantifiable performance requirements covering indicators such as building energy consumption and indoor comfort, which could be negatively correlated. Therefore multi-objective optimization algorithms are introduced to balance multiple performance indicators for optimal solutions. For example, optimizing the light environment and improving indoor energy consumption simultaneously (Shi et al., 2020), minimizing cooling load, and maximizing daylighting performance in summer (Kim & Clayton, 2020). The operation of DF for multi-objective optimization requires a comprehensive calculation of environmental parameters to achieve an optimal solution for a certain period, creating an intelligent control system to cope with the complexity is an essential research direction (Böke et al., 2019; Favoino et al., 2022). Current researches on DF mainly use the Rhino/Grasshopper platform to build a parametric model (González & Fiorito, 2015), which could be combined with modules such as Ladybug and Honeybee to simulate, evaluate, and optimize the performance of the façade (Roudsari & Pak, 2013). The simulation outcome could provide solid evidence for design assessment in the early design stage, finally achieving the performance-driven design purposes in the pre-design phase.

2.3. MACHINE LEARNING & OPTIMIZATION

For real-world building design problems, designers usually have to deal with conflicting design objectives, e.g., minimum energy consumption versus maximum thermal comfort, and minimum energy consumption versus minimum construction cost. Thus, multi-objective optimization is suitable in the architecture area. Current research in computational optimization has proven to be effective in building energy efficiency (Evins, 2013) (Ascione et al., 2017) and can be applied in several building research directions. Genetic algorithms as a population-based approach are well suited to solve multi-objective optimization problems for designs. (Nguyen et al., 2014).

The design of DF requires a real-time response to the environment which is a time-consuming task, especially when some complex digital models require multiple environmental simulations. Thus the simulation model has often been oversimplified and led to inaccurate modeling. Much of the current research is developing approximation methods through metamodels or alternative models (e.g., ANN), which approximate pre-established performance functions (describing the objective) without reducing the complexity of the problem. Han used Neural networks to predict light environment metrics such as DA and UDI under the influence of dynamic epidermis (Han et al., 2021). By introducing explanation machine learning into the prediction of light environment and energy consumption, the performance of DF will add to the providing interpretable knowledge for parametric movable facades design and increase the efficiency of the dynamic facades optimization process.

3. Method

The primary goal of this paper is to improve the user's light environment comfort and reduce building energy consumption simultaneously. Based on the design characteristics of dynamic facades, four parametric models of dynamic façade are constructed on Rhino/Grasshopper platform with different parameters (unit length, unit width, movable unit parameters, different patterns, and other variables of the facade), each model will use Honeybee, Ladybug, and Energyplus for light and energy simulation. Each variation will generate the light environment assessment indexes: illuminance, Daylight Glare Probability (DGP), and Energy Use Intensity (EUI). Next, the output data will be used for batch simulation to generate a dataset for XGBoost, which can explain the contribution of dynamic facade parameters for the indoor light environment index and energy consumption. The final step is using a machine learning prediction model to build a prediction workflow in Grasshopper, then using genetic algorithm (Wallcei X) for the light indexes and energy consumption optimization (Makki et al., 2018). With the prediction model and optimal approach, it is easy to build the annual control strategy corresponding to the facade. Based on the light environment data for the year, the real-time light indexes like Illuminance, DGP, and room energy consumption will be compared for four prototypes. The conclusions of this paper provide a reference for user-centered interior light design under the influence of dynamic facades and investigate the performance of dynamic facades in reducing energy consumption. The research workflow is shown in the Figure 1.

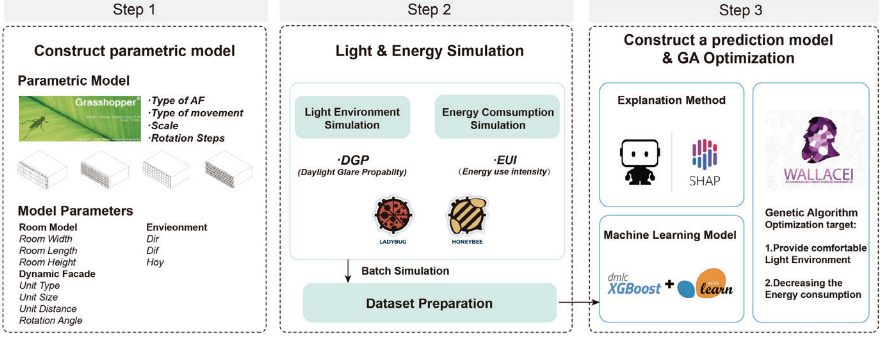


Figure 1. Research workflow

3.1. PARAMETRIC MODEL

The reference room of the DF is an office building type. In order to increase the wide adaptability of the research process, each parameter of the building and the environment is decomposed into several different variables in the preliminary stage to achieve an accurate parametric simulation, and the length, width, and height of the target room are renamed as variables RW, RL, RH in the preliminary simulation process to obtain a large amount of sample space.

The parametric model is mainly selected from the open and closed type DF (façade surface rotates around the axis). The four open and closed prototypes are categorized with the basic shapes: rectangle, triangle, circle, and triangle combined hexagon as the main simulation prototype. The first step is to establish the typical parametric building room model and set the suitable range of length, width, and height; then divide the target façade into grids. The size of the grids is equal to DF unit size (width and height), and finally, place four DF cells of different morphological types in each grid. The designed room is depicted in Figure 2.

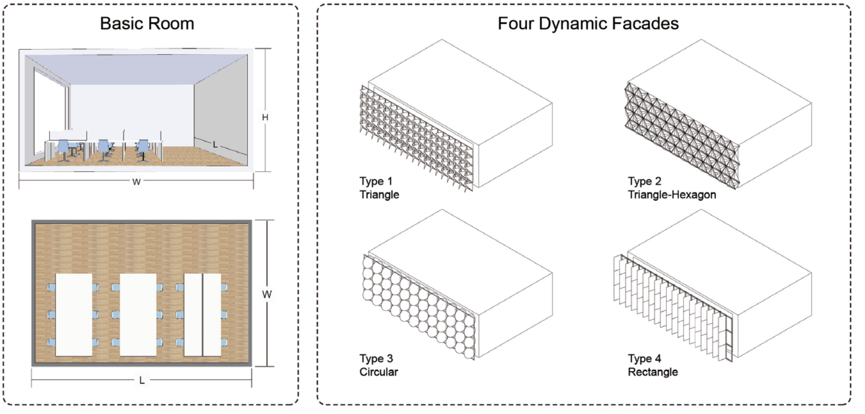


Figure 2. Parametric dynamic façade model

3.2. INDOOR ENVIRONMENT PERFORMANCE SIMULATION

With increasing attention to indoor light comfort, daylighting evaluation has been defined by different criteria: 1) window or glazing area in relation to the room area or facade area. In this paper, the window-to-wall ratio chooses 80% as a simulation to curtain wall; 2) quantity of daylight inside a room; levels for daylighting are generally described as preferred or recommended - either by specific illuminance (lux) levels in a workplace. Considerations of sunlight access and its duration will be influenced by design variations such as orientation, the disposition of rooms and their windows, and the selection of solar shading devices.

Energy simulation is a computer-based analytical process that helps to evaluate architectural energy performance. It can improve energy efficiency by making necessary modifications before the building is constructed. In this paper, simulation is conducted in the Grasshopper platform for the EUI. Simulation indexes for lighting and energy are listed in Table 1.

Light and energy simulation are closely related to changing weather, but the weather is hard to predict. Thus, in this paper, weather data are deconstructed into multiple parameters: time is composed of the month, day, and hour; the location chooses Shanghai and uses EPW file to extract climate data to construct CIE sky.

Table 1. Simulation indexes for lighting and energy

Name	Definition	unit	Suitable range
DGP (Daylight Glare Probability)	DGP is an indication of the percentage of people who would be disturbed by glare.	%	<0.35= imperceptible; 0.35-0.4 = perceptible; 0.4-0.45= disturbing; >0.45= intolerable;
EUI	The total end-use intensity results from the simulation. Specifically, this is the sum of all electricity, fuel, district heating/cooling, etc., divided by the gross floor area.	kWh/m2	

3.3. PARAMETRIC MODEL COMBINED WITH SIMULATION MODULES

The Ladybug-Honeybee plugins combined EnergyPlus, and visual comfort with Radiance, which extends the capabilities of Grasshopper for simulating energy consumption and light condition. There are four main parts to constructing the simulation process. 1） Using the parametric model from 3.1 to generate multiple outcomes with random parameters. 2） Preparing simulation: Using Honeybee to convert the Brep of surface for the Honeybee model, deconstruct the climate condition into quantifiable parameters, and use random modules to select HOY for each simulation. 3) Running four simulation modules for one target room each time: Illuminance simulation, Daylight Autonomy, DGP simulation, and Energy simulation. 4) Using Ladybug fly for batch simulation with random parameters, the outcome data are saved through the TT toolbox. Each simulation costs 722 seconds, and 100 times simulation will take 20 hours. In this study, each prototype of DF runs 1000 times to create a dataset for machine learning.

3.4. MACHINE LEARNING

We used machine learning to model the nonlinear relationship between 11 impact factors and the seven lighting and energy simulation outputs. XGBoost is an optimized distributed gradient boosting library designed to be highly efficient, flexible, and portable. It implements machine learning algorithms under the Gradient Boosting framework. XGBoost provides a parallel tree boosting (also known as GBDT, GBM) that solves many data science problems quickly and accurately. The Boosting method sequentially trains weak learners, in which Boosting iteratively fits a weak learner, aggregates it into the ensemble model, and "updates" the training dataset better to consider the current strengths and weaknesses ensemble when fitting the next base model. Typically, the goal of boosting the approach results in a model with less bias (more accurate).

This paper compares four ensemble models using XGBoost for the Boosting method. Using the hold-out method to divide the simulation dataset, with 70% as the training set and 30% as the test set. The training was performed using the Scikit-learn machine-learning library. Model evaluation was performed using the coefficient of determination (R^2) and mean square error (MSE). The optimal model is determined by model comparison and hyperparameter optimization.

4. Results

4.1. SIMULATION OUTCOME

According to the parametric model set in Part III, the simulation process of epidermal performance was built in the grasshopper platform using the ladybug and honeybee modules. In order to obtain more details of the Illuminance simulation, each room plane was divided into 2500 grid areas of 50*50. The camera points are located at the middle point against the window.

The visualization results (Figure 3) show the detailed indoor light environment of the room with four different types of DF installed at noon on September 21. The dimensions of the target room are 8m(RW)*14m(RL)*4m(RH), and the façade units are all 0.8 wide cell modules with selected opening and closing angles of 60° and 45°, respectively. For the illuminance index of the target room, Type2 and Type3 can make the light distribution in the room more uniform, and the shading effect in the window area is evident; for the index of DGP, each prototype can reduce the uncomfortable glare to different degrees by changing the opening, and closing angles, among Type 2 has the best effect of glare prevention.

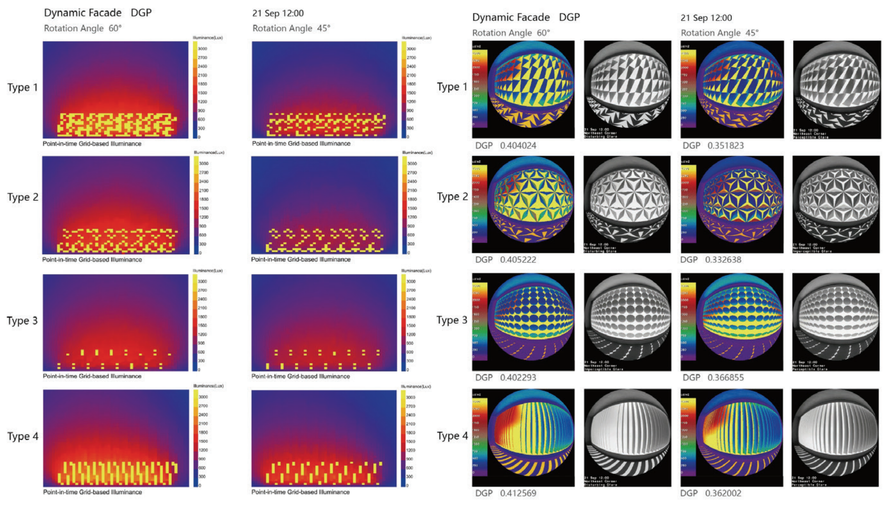


Figure 3. Visualization of light simulation

4.2. MACHINE LEARNING MODEL

Figure 4 plots the Shapley value of each feature for each sample, indicating which features are most important and how the sample size affects the model. The features' ranking still indicates the features' contribution to the model or feature importance. Each point represents a sample, and the color indicates the size of the feature value, with red indicating a higher feature value and blue indicating a smaller feature value. The color differences allow us to match how changes in feature values affect the model's output. In addition, broad regions represent a large number of aggregated samples. The plots show that rotation angle is the most crucial feature to impact the DGP value, proving that DF can improve the light environment comfort. As for EUI, it is affected more by room size, but rotation angle can also be considered essential for decreasing EUI.

4.3. OPTIMIZATION CONTROL STRATEGIES

For the optimization of DF, we use Wallacei to generate parameter combinations. Machine learning models from section 3.4 are imported into GHpython, and EUI and

DGP outcomes could be generated within 1.3 seconds, reducing the time for performance simulation. To evaluate the optimization outcome, each DF prototype will run with the same population setting: Generation Size of 10 and Generation Count of 20. In order to acquire the façade rotation angle for every hour and evaluate the whole day's performance, the first test gene sets and the best average of fitness genes are chosen to make a comparison (Figure 5), which shows that Type 1 and Type 4 can decrease EUI better, and lower the average DGP of one-day simultaneously. Optimized control strategies of DF bring a maxim 13.5% EUI decrease and 51.7% reduction of DGP.

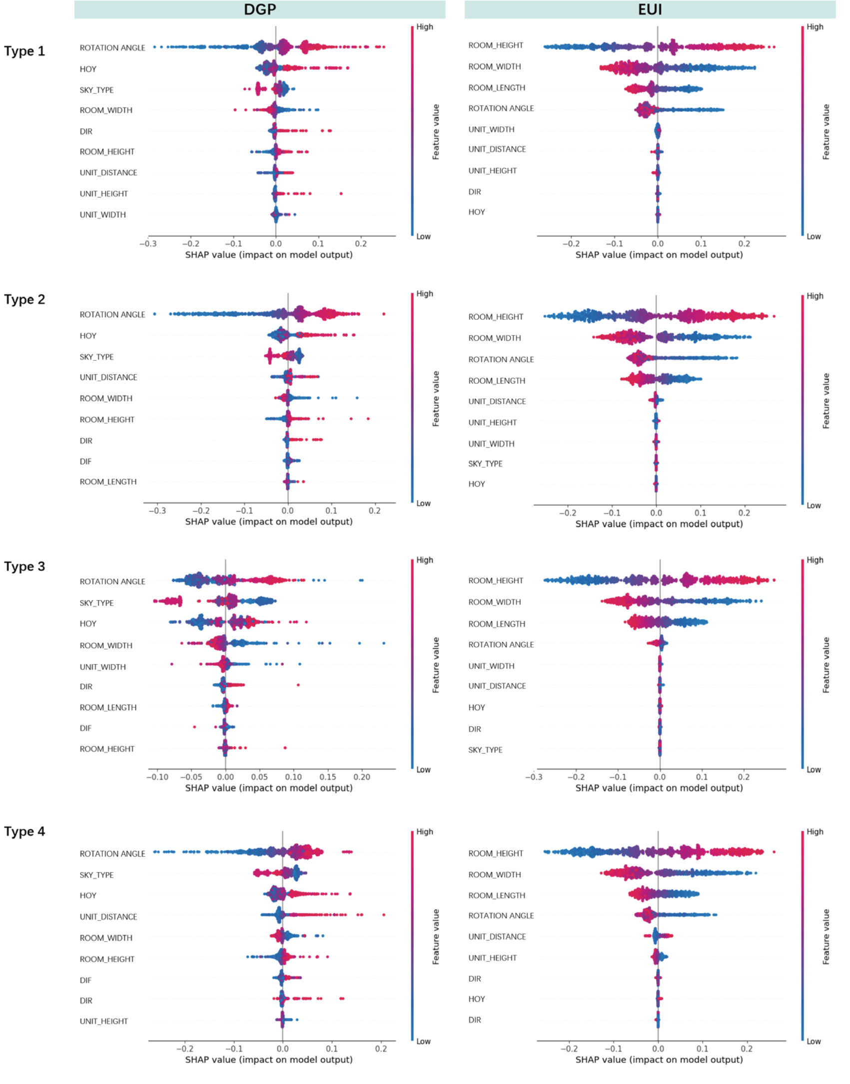


Figure 4. Summary of features importance

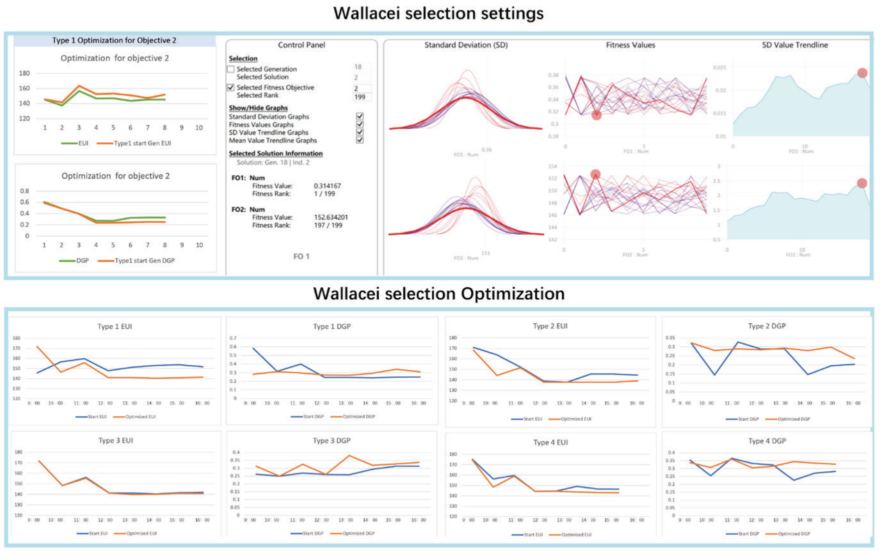


Figure 5. Wallacei Optimization

5. Conclusion

The experiment demonstrated that the two objectives (EUI and DGP) might be negatively correlated, as evidenced by comparing the simulation results generated by the EUI. However, through multi-objective optimization, the relationship between the two indicators can be further balanced, and it is possible to reduce some of the energy consumption while improving the comfort of the indoor environment is accomplished.

This study proves that the proposed workflow can effectively generate operating parameters of the DF for one day and obtain real-time operating parameters, achieving the goal of finding the optimal solution for the DF operation. The next step in this research topic is to generate an optimal control strategy for the whole year and add more parameters to achieve a comprehensive and relatively realistic optimization strategy for the operation of the DF. In addition, we will calculate detailed energy consumption with DF, which is necessary for future sustainable building operations.

Acknowledgements

This research was financially supported by the National Natural Science Foundation of China under Grant NO. 52278041 & 51908410 and the Fundamental Research Funds for the Central Universities.

References

Ascione, F., Bianco, N., de Masi, R. F., Mauro, G. M., & Vanoli, G. P. (2017). Energy retrofit of educational buildings: Transient energy simulations, model calibration and multi-objective optimization towards nearly zero-energy performance. Energy and Buildings, 144, 303–319. https://doi.org/10.1016/j.enbuild.2017.03.056

Böke, J., Knaack, U., & Hemmerling, M. (2019). State-of-the-art of intelligent building envelopes in the context of intelligent technical systems. Intelligent Buildings International, 11:1, 27-45, DOI: 10.1080/17508975.2018.1447437

Evins, R. (2013). A review of computational optimisation methods applied to sustainable building design. In Renewable and Sustainable Energy Reviews (Vol. 22, pp. 230–245). https://doi.org/10.1016/j.rser.2013.02.004

Favoino, F., Baracani, M., Giovannini, L., Gennaro, G., & Goia, F. (2022). Embedding intelligence to control adaptive building envelopes. Rethinking Building Skins, 155–179. https://doi.org/10.1016/B978-0-12-822477-9.00007-3

González, J., & Fiorito, F. (2015). Daylight design of office buildings: Optimisation of external solar shadings by using combined simulation methods. Buildings, 5(2), 560–580. https://doi.org/10.3390/buildings5020560

Han, Y., Shen, L., & Sun, C. (2021). Developing a parametric morphable annual daylight prediction model with improved generalization capability for the early stages of office building design. Building and Environment, 200(March), 107932. https://doi.org/10.1016/j.buildenv.2021.107932

Kim, H., & Clayton, M. J. (2020). A multi-objective optimization approach for climate-adaptive building envelope design using parametric behavior maps. Building and Environment, 185. https://doi.org/10.1016/j.buildenv.2020.107292

Loonen, R. C. G. M., Trčka, M., Cóstola, D., & Hensen, J. L. M. (2013). Climate adaptive building shells: State-of-the-art and future challenges. Renewable and Sustainable Energy Reviews, 25, 483–493. https://doi.org/10.1016/j.rser.2013.04.016

Luna-Navarro, A., Loonen, R., Juaristi, M., Monge-Barrio, A., Attia, S., & Overend, M. (2020). Occupant-Facade interaction: a review and classification scheme. Building and Environment, 177(March), 106880. https://doi.org/10.1016/j.buildenv.2020.106880

Makki, M., Showkatbakhsh, M., & Song, Y. (2018). Wallacei: An evolutionary and Analytic Engine for Grasshopper 3D. Wallacei. https://www.wallacei.com/

Nguyen, A. T., Reiter, S., & Rigo, P. (2014). A review on simulation-based optimization methods applied to building performance analysis. In Applied Energy (Vol. 113, pp. 1043–1058). Elsevier Ltd. https://doi.org/10.1016/j.apenergy.2013.08.061

Romano, R., Aelenei, L., Aelenei, D., & Mazzucchelli, E. S. (2018). What is an adaptive façade? Analysis of recent terms and definitions from an international perspective. Journal of Facade Design and Engineering, 6(3), 065–076. https://doi.org/10.7480/jfde.2018.3.2478

Roudsari, M.S., & Pak, M. (2013). Ladybug: A parametric environmental plugin for Grasshopper to help designers create an environmentally-conscious design. Proceedings of BS 2013: 13th Conference of the International Building Performance Simulation Association, 3128–3135 DOI:10.26868/25222708.2013.2499

Shi, X., Abel, T., & Wang, L. (2020). Influence of two motion types on solar transmittance and daylight performance of dynamic façades. Solar Energy, 201(March), 561–580. https://doi.org/10.1016/j.solener.2020.03.017

Tabadkani, A., Roetzel, A., Li, H. X., & Tsangrassoulis, A. (2021). A review of occupant-centric control strategies for adaptive facades. Automation in Construction, 122(May), 103464. https://doi.org/10.1016/j.autcon.2020.103464

AN ATTENTION ECONOMY THEME PARK

The Voluntary Rigging of Digital Personas

DANIEL ESCOBAR[1], CARLOS NAVARRO[2], and EVANGELIA PAPASPYROU[3]
[1]*OLA Research.* [2]*SCI-Arc.* [3]*AUTH.*
[1]*descobar01490@gmail.com, 0000-0003-3772-3723,*
[2]*carlosaul@gmail.com, 0000-0002-2976-0674*
[3]*papaspirou.eva@gmail.com@gmail.com, 0000-0002-5913-4818*

Abstract. Planetary scale computation is evolving the way we digitize the physical urban space. The following research aims to provide an architectural response to the accelerating digitization of our physical world and societal life processes of economy and communications. It acknowledges the legitimate bias in the perceptual value of territories favored by the new Attention Economy of the Metaverse and Blockchain-based Virtual Environments. It proposes the analogy of a theme park derived from the distorted collective vision of today's reality, of a reduced collection of favored attraction locations. The research provides first a review of contemporary studies related to the operation of the Attention Economy in the Metaverse, Web3 platforms, and Gamified Virtual Environments, as well as studies on recent architectural expressions or typologies of these spaces. A series of methodologies are described next to convey the impact of recent advances in Artificial Intelligence (AI) on the creation of digital personas and worldmaking for this type of economy. The methodologies comprise a three-stage workflow based on data mining and curation, processing through AI-aided generative methods, and implementation with game engine environments, ultimately discussed regarding simulation and creative agency.

Keywords. attention economy, virtual environments, embodiment, digital twin, multimodal AI

1. 1. Introduction

The following research aims to provide an architectural response to the accelerating digitization of our physical world and societal life processes of economy and communication.

This phenomenon, rooted in the evolution of the Internet and the World Wide Web, is today manifested as both the accumulation of Cyber-physical systems such as the Internet of Things (IoT) and the hybrid physical-digital communicative framing of all

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 209-218.

kinds of social interactions (e.g., Web 3).

Cyber-physical systems, such as smart grids, autonomous automobile systems, or IoT technologies, can operate within a planetary-scale computation causing perceptual distortions of our physical world. This has been studied by Benjamin Bratton via the idea of The Stack, as an accidental megastructure made of different genres of computation aligned in layers. He argues that planetary-scale computation both distorts and reforms modern jurisdiction and political geography and produces new forms of these in its own image (Bratton, 2016)

Today, with the advent of the Metaverse, this altered reality, inherent to smart cities or hybrid territories, is shifting its cyber-physical latency towards a more spatial-visual communicative framing incentivized by a new economy of embodiment that operates between the virtual and physical worlds (Schumacher, 2022).

This research acknowledges the legitimate bias in the perceptual value of territories favored by the new Attention Economy of the Metaverse and Blockchain-based Virtual Environments and proposes the analogy of a theme park as an architectural response to the distorted collective vision of today's reality, of a reduced collection of favored attraction locations. See Figure 1. The theme park typology identifies users' attention as a valuable commodity that triggers a new economy of exchange modalities of incentivized attention.

A series of methodologies are described next to convey the impact of recent advances in Artificial Intelligence (AI) on the creation of digital personas and worldmaking for this type of economy. The studied methods are ultimately analyzed in terms of their effectiveness and possible improvements for future work with more recent AI-aided procedures.

Figure 1. The Attention Economy Theme Park

2. 2. Background

The initial research focuses on understanding the attention economy of contemporary Virtual Environments (VE) operating in the context of Web3, Blockchain, and immersive NFTs. Understanding contemporary motivations for virtual embodiment, exchange modalities and direction, and gamification led to identifying a symbiotic link between attention economy-incentivized VEs and the digital representation of their architecture or spatial design.

2.1. INCENTIVIZED VE

Recent research on this subject focuses mainly on the turn to more Web3-based behavior in social media and platforms such as Blockchain-based Video Games (BVGs) powered by the new Crypto Creator Economy and Virtual Embodiment in the Metaverse.

2.1.1. Blockchain-based Video Games

Guidi and Michienzi explain the economic motivation of BVGs, such as the famous case of Descentraland, which evolved Web2 App-based Social Media to Web3 Wallet-based Blockchain Social Media (BSOM), adding an economic layer to user interactions, giving users control of the urban planning of the VE and the chance to create and monetize their content (Guidi and Michienzi, 2022).

Min and others propose that Blockchain is the foundation of Decentralized Applications (DApps) and digital games, which can be categorized based on the nature of the Blockchain technology. Users and creators are stimulated to engage in the game economy while progressively accumulating Rule Transparency, Asset Ownership, Assets Reusability, and User-Generated Content (UGC) as they move into the Ethereum and other proof-of-work POW and proof-of-stake POS-based Blockchains (Min et al., 2019).

2.1.2. Crypto Creator Economy

BlockchainBrett refers to the Crypto Economy as the foundation of the Creator Economy, where audiences become owners as collectors of content. This economy develops through Engagement Protocols (Emote NFTs, Artist-to-Fan Compensations, or Content-backed Currencies) within Web3 On-Chain Ownership-based relationships through Wallet-to-Wallet messaging. It exploits the Lifecycle of Content NFTs regarding access, distribution, and ownership control by creators. Creators control their content and community and can use composable content to co-create as media Legos (BlockchainBrett, 2022).

2.1.3. Virtual Embodiment

Lam and others explain how automatic capture of human motion to recreate in virtual worlds focuses only on head and hand movements. It argues that full body tracking and avatar reconstruction requires high-resource network environments and proposes that delivery efficiency can improve in low-resource network environments by compressing data (Lam et al., 2022).

This research reveals that Incentivized VEs engage with the attention economy through gamification and wallet-based interactions on Web3 platforms, benefiting creators and collectors. Embodiment in the metaverse requires a more cost-efficient solution than full-body tracking. And Blockchain-based asset ownership stimulates UGC, i.e., the Holly+ project based on Machine Learning Identity Spawning with an alternate scheme of ownership control (Herndon, 2021).

2.2. ARCHITECTURAL SEMIOLOGY OF INCENTIVIZED VE

Related research on the matter explores the typological look or digital representation of the architecture where the incentivized behavior occurs, ranging from Blockchain-based Architecture to Worldmaking of Gamified VEs, and AI technologies to realize VEs in the Metaverse.

2.2.1. Blockchain-based Architecture

Grasser and Parger's recent experiment proposes the use of the Blockchain block as an open container that can store multimedia content and act as a resilient Web3 asset, suggesting a decentralized approach to architecture, as a design method and for blocks or assets assemblage (Grasser and Parger, 2022).

2.2.2. Worldmaking of Gamified VE

Ozden and others provide a review of urban gamification examples such as Second Life, SimCity, and Virtual Smart City Hero, and acknowledge the importance of the Virtual Reality (VR) experience of smart cities. Through gamification and choreographing of users inside a Digital Twin (DT) city, they propose to simulate energy consumption and environmental sustainability (Ozden et al., 2022).

Barsan-Pipu and others focus on how relevant empathic data can be acquired in real-time by exposing subjects within a dynamic VR-based Virtual Procedural Environment (VPE) and assessing their emotional responses while controlling the actual generative parameters via a live feedback loop. Their procedural system based on three-dimensional noise fractal algorithms can produce evolving VPEs with either raw/coarse (voxel) or smoothed results (Barsan-Pipu et al., 2020).

2.2.3. AI-Based VE in the Metaverse

Aloqaily and others present the metaverse as a VE that features real and imitational objects such as avatars and other forms of digital objects. They state that DTs can evolve synchronously throughout the lifetime of physical entities. They propose that AI will enhance Metaverse services in the future and envision a DT-6G Metaverse that can be realized through the cooperation of various technologies including DT, XR, 6G, and the blockchain, where significant amounts of data captured through XR devices are processed by AI to build DTs which are then deployed from the Blockchain (Aloqaily et al., 2022).

This research demonstrates that current approaches for a Blockchain-based architecture are mostly based on participatory design and speculative fabrication methods, with yet unexplored design expressions or semiology, seemingly suitable for the mereological approaches of the current digital era. Worldmaking of gamified VEs is suitable for sustainability and performance simulation in DT sites. Blockchain, AI, and XR will propel the metaverse and DTs for users and sites in the future, opening possibilities for a new architectural semiology powered by VPE feedback loops in combination with XR and high-resource networks.

3. Methods

This section introduces resourceful methodologies for embodiment and worldmaking for the new attention economy of the metaverse. They leverage UGC and asset ownership of BVGs economies to stimulate the interaction of users and creators. They study gamification for simulation in DT sites, as well as the potential of AI and VPEs to generate a new architectural semiology for hybrid reality.

The case study proposes the architectural analogy of a theme park, where attention economy incentivized behavior is to be experienced by procedurally generated avatar agents: a DT avatar, a social media influencer, and an NFT character, all inhabiting a composite world that assemblies DT sites from Lima, Los Angeles, and Athens.

The methods rely on collaborative Human-AI content generation and consider three types of workflows for the creation of multimodal assets: The first one consists of data mining and curation for both 3D content and natural language processing. The second one takes the curated data for content generation aided by deep learning algorithms. The third part consists of bringing all the content together in a game engine platform where users and agents can interact in a virtual environment.

3.1. DATA MINING AND CURATION

Two types of datasets are used for starting content generation. One is visual and the other one is a text dataset. For the visual data, images of faces or characters are gathered from either personal photos, public online images, or AI-generated faces from a StyleGAN2, or Stable Diffusion. These images serve as the base for creating embodied agents/avatars.

In the development of the 3D environment, videos from YouTube and on-site recordings from different locations around the world are used to be later processed with photogrammetry tools. The text data is mined from writings of several theoretical architecture texts and a set of conversations from podcasts. Thus, providing a variety of perspectives and narratives that can be learned by AI in the next part of the workflow. The conversations are extracted from YouTube videos using a speech-to-text deep learning model. This way we can use the speech data input jointly with the writings.

3.2. DATA PROCESSING

3.2.1. Procedural 3D Reconstruction

To generate a virtual 3D environment that can be loaded into a game engine, the set of images and videos from sites is processed into photogrammetry models using applications such as Reality Capture, Meshroom, and Polycam. These applications take as input images or videos, preferably with overlapping frames, and output a 3D mesh model of the object or site. The 3D mesh model is exported as an obj file format and can be imported into open-source procedural 3D modeling software Blender for more processing or animation. See Figure 2.

To construct the agents as embodied avatars, an image of a posed body is generated for later 3D translation. Photoshop can be used to rapidly combine selected front-facing StyleGAN output faces with public online images of different posed bodies.

3.2.2. AI-Aided Generative Bodies

In this section, a workflow is described on how to convert static images into 3D objects using pre-trained deep learning models for fast predictions. The exported image can be used as input for PIFuHD or ICON models. These are both deep learning algorithms that use an input image to predict a 3D mesh body that can be used for animation. These algorithms are accessible to run in the web-based coding platform Google Colab.

Figure 2. Digital Twin Site

Alternatively, a more AI-aided approach can be done to generate the face and body. On the one hand, a pre-trained object detection and segmentation algorithm, YOLO V5 or MaskRCNN, tuned for selecting the human category can be used to extract any person from an image. It will export the segmented person as a PNG with white background and a mask that can be used for processing. The segmented person can be used as input for the posed body prediction algorithms mentioned above. See Figure 3.

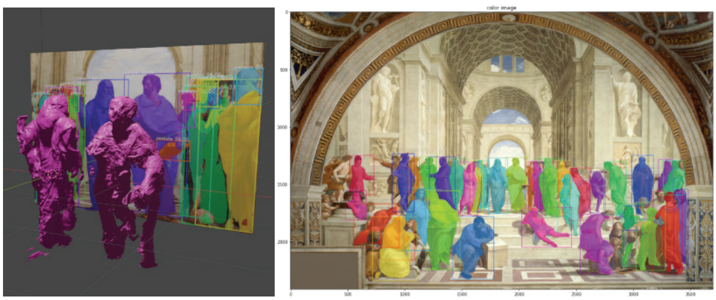


Figure 3. Image to Body Generation for Avatar

After completing the avatar's body generation, animation rigging can be done easily using the web application Mixamo. The 3D mesh body can be uploaded into Mixamo and by selecting several keypoints, an animation like walking or running can be applied to the avatar and downloaded as a rigged armature with keypoints to animate in Blender. Once the animated avatar is loaded in Blender, a texture painting process can be done to map the original Photoshop body onto the avatar. Thus, concluding a full texturization and the completion of an embodied avatar that can be used as an AI agent in later steps.

3.2.3. AI-Aided Generative Speech

For this section, a workflow that uses text and natural language processing models (NLP) is described both as a functional mapping tool, from speech-to-text (STT) and text-to-speech (TTS), and as a generative tool.

For processing conversation audio extracted from YouTube videos, a pre-trained STT model, Wave2Vec, was used to extract text from audio. The input dataset is extracted from podcast conversations using PyDub library, which can download a video from YouTube and return an audio file in MP4 format. This is then processed into a series of audio clips 60 seconds long using a Librosa library for Python. The audio clips are loaded as an input batch into Wav2Vec. The output is a series of text that are concatenated back into the full conversation and saved as a TXT file format.

With the text processed into the correct format, it can then be fed into a pre-trained GPT2 model for learning. This process called fine-tuning allows for the training of a new dataset by using a set of pre-trained weights to find better patterns and achieve faster convergence. Once trained, the model can then be used as a generative tool by inputting a starting prompt like some words or a sentence, and the model will return text output that displays what it has learned from the original dataset. See Figure 4.

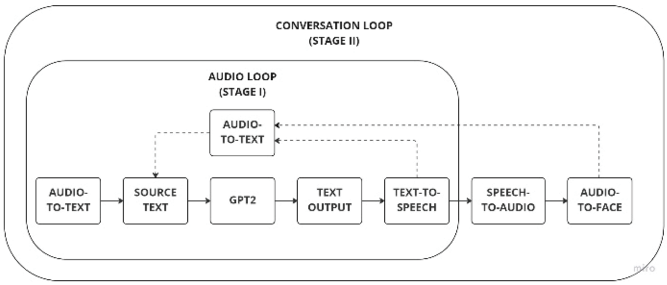


Figure 4. Audio-to-Text Translation Loops

After this process, input prompts are used to evaluate what the model has learned from the texts. These texts are then assigned to the avatars to generate their voice and perspectives about topics. To use that text as speech for the avatars, a TTS pre-trained model, MelGan, is used. It inputs a text file that gets encoded into a vector and inference is run to return a person speaking the text. The speech audio is then saved as a WAV file format that can be used to animate the way the avatar speaks in a later step.

3.2.4. 3.2.4. AI-Aided Audio-To-Face

Emotion and animated faces help us communicate. To build better AI emotive agents we can now map our facial movements to meshes.Using the Nvidia Omniverse, the Audio2Face application can be used to animate and rig a face that produces realistic emotive expressions and mouth movements when speech audio is used as input. For this part of the process, a 3D model of the avatar is imported into the application and the face is rigged according to a set of key points.

These key points are used to pair a pre-trained face model provided by the application which allows mapping an audio track to facial movement as the 3D model talks. After training is complete, the 3D model can be exported to Blender and animated with speech. After this process, the avatar can now talk animatedly based on text output from GPT2. This process is repeated for every character. See Figure 5.

Figure 5. Audio 2 Face Rigging for Avatar

3.3. GAME ENGINE ENVIRONMENT

To create an interactive virtual immersive experience, PlayCanvas is used as a web-based multiplayer game engine. Environments created within here can be easily accessed via a web browser. Scripting in Javascript is used to control cameras and UI interactions for users. The game engine serves as the best platform to place and animate all the assets generated in the previous section. See Figure 6.

At this step, the 3D photogrammetry models of the site are loaded. They are arranged in such a way that a continuous circulation could be achieved. This results in spaces that act as attractions. The organization provides a new hybrid experience of physical sites recontextualized in a digital environment. Each space acts as a destination in a proposed digital theme park. 3D UIs are activated with specific semilogical features when an avatar agent performs a valuable action, and a transaction occurs.

Figure 6. PlayCanvas VE

The avatars are loaded onto the environment and treated as AI agents where they come packed with topics of conversation that are generated by a fine-tuned GPT2 model. This allows them to have serendipitous interactions with each other or other human users. They are also allowed to roam freely on the walkable surfaces of the environment. Finally, the environment is inhabited by AI agents and humans can access it as an interactive experience where users can observe how the agents move and interact with the digital twin scans of real physical spaces.

4. Results and Discussion

Under the umbrella of creative AI, the convergence of these methods leverages the potential of pre-trained models for fast prototyping and iteration of ideas for creating digital personas and worldmaking. This case study acts as a simulation test ground where multimodal generative AI workflows are guided by incentivized attention. The semiology of 3D UIs is affected by valuable transactions of AI agents, which can develop agency beyond traditional non-playable characters.

Worldmaking and character development in virtual environments and video games require a huge undertaking that is both time-consuming and costly. The methods described in this paper allow for an implementation that is open-sourced and can be done through a variety of free software.

4.1. FUTURE WORK

Future focus of this type of work can include 3D reconstruction with more recent AI-aided methods such as Neural Radiance Fields (NeRFs) for digital reconstructions of sites. Other future work can include optimized processes for creating digital personas with a complete AI-aided generative approach using text-to-image with StableDiffusion. Image outputs can be fed into posed body mesh reconstruction algorithms. Another field of work could be Natural Language Processing (NLP). New language models such as GPT3, Whisper, or ChatGPT from OpenAI, can facilitate the communication between agents and humans in both the digital and physical.

Recent projects by DeepMind and Meta explore the use of simulation platforms for research in Embodied AI. According to Meta, there is a paradigm shift from internet AI based on static datasets to embodied AI where agents act in realistic environments (Savva et al., 2019). DeepMind's projects explore how agents can learn to generalize a skill set in open-ended play, by creating procedural 3D worlds to simulate a variety of environments where agents are prompted with goals through text (Team, O.E.L et al., 2021). Their results show that agents are able to learn skills that are transferable to unseen games (Team, A.A. et al., 2018). Being able to create agents that are trained in the digital and can act in the physical brings new possibilities to design, construction, and robotics. This could lead to changes in design policymaking, bringing forward the role of agency and authorship within the creative process, with the rise of more autonomous bots and robots collecting data from our built public and private spaces.

5. Conclusion

This paper introduces feasible advances to contemporary research on the operation of attention economy-incentivized VEs and their architectural expressions. It acknowledges that the advent of the Metaverse triggers the emergent manifestations of space to pulse in the dipole physicality-virtuality, introducing more hybrid permutations. The case study examines gamification for simulation in DT sites, as well as the potential of AI and VPEs to generate a new architectural semiology for the hybrid reality. The methods propose cost-efficient alternatives for embodiment and worldmaking leveraging UGC and asset ownership of Blockchain-based game economies to stimulate the interaction of users and creators. The research puts forth the necessary balancing of human agency for crafted reconstructions and AI agency for the simulation of behaviors and performance in VEs. Traditional human-centric visions of creativity get consequently challenged when the role of human agency is reshaped by the augmenting of skills with AI.

Acknowledgments

We sincerely thank the Current Team: Provides Ng, Eli Joteva, Ya Nzi, and Artem Konevskikh for the sharing of introductory procedural and AI-aided techniques.

References

Aloqaily, M., Bouachir, O., Karray, F., Ridhawi, I. A., & Saddik, A. E. (2022). Integrating Digital Twin and Advanced Intelligent Technologies to realize the metaverse. *IEEE Consumer Electronics Magazine*, 1–8. https://doi.org/10.1109/mce.2022.3212570

Barsan-Pipu , C., Sleiman, N., & Moldovan, T. (2020). Affective Computing for Generating Virtual Procedural Environments Using Game Technologies. *ACADIA 2020: Distributed Proximities / Volume I: Technical Papers [Proceedings of the 40th Annual Conference of the Association of Computer Aided Design in Architecture (ACADIA)* .

BlockchainBrett. (2022). *The Crypto Creator Economy*. Creators GMI. Retrieved from https://creator.mirror.xyz/Gt_4Uy1K_9LBElWuFQdnknzazZ1-eVQTCvRNMZlY-u0

Bratton, B. H. (2016). *The stack - on software and sovereignty*. MIT Press.

Grasser, A., & Parger, A. (2022). Blockchain architectures, the potential of web3 for decentralized participatory architecture - collaborative objects on the blockchain. *Proceedings of the 40th International Conference on Education and Research in Computer Aided Architectural Design in Europe (ECAADe) [Volume 1]*. https://doi.org/10.52842/conf.ecaade.2022.1.431

Guidi, B., & Michienzi, A. (2022). Social Games and blockchain: Exploring the metaverse of Decentraland. *2022 IEEE 42nd International Conference on Distributed Computing Systems Workshops (ICDCSW)*. https://doi.org/10.1109/icdcsw56584.2022.00045

Herndon, H. (2021). *Holly+*. Holly+- Holly Herndon. Retrieved from https://holly.mirror.xyz/54ds2IiOnvthjGFkokFCoaI4EabytH9xjAYy1irHy94

Lam, K. Y., Yang, L., Alhilal, A., Lee, L.-H., Tyson, G., & Hui, P. (2022). Human-avatar interaction in Metaverse. *Proceedings of the 4th ACM International Conference on Multimedia in Asia*. https://doi.org/10.1145/3551626.3564936

Min, T., Wang, H., Guo, Y., & Cai, W. (2019). Blockchain Games: A Survey. *2019 IEEE Conference on Games (CoG)*. https://doi.org/10.1109/cig.2019.8848111

Ozden, S., Arslantürk, E., Senem, M. O., & Imdat, A. (2022). Gamification in Urban Planning: Experiencing The Future City. *Conference: ASCAAD 2022: The 10th International Conf. of the Arab Society for Computation in Architecture, Art and Design*.

Savva, M., Kadian, A., Maksymets, O., Zhao, Y., Wijmans, E., Jain, B., Straub, J., Liu, J., Koltun, V., Malik, J., Parikh, D., & Batra, D. (2019). *Habitat: A platform for embodied AI research*. arXiv.org. Retrieved from https://arxiv.org/abs/1904.01201

Schumacher, P. (2022). The metaverse as opportunity for Architecture and Society: Design Drivers, core competencies. *Architectural Intelligence*, *1*(1). https://doi.org/10.1007/s44223-022-00010-z

Team, A. A., Bauer, J., Baumli, K., Baveja, S., Behbahani, F., Bhoopchand, A., Bradley-Schmieg, N., … Zhang, L. (2023). *Human-timescale adaptation in an open-ended task space*. arXiv.org, https://doi.org/10.48550/arXiv.2301.07608.

Team, O. E. L., Stooke, A., Mahajan, A., Barros, C., Deck, C., Bauer, J., Sygnowski, J., … Czarnecki, W. M. (2021). *Open-ended learning leads to generally capable agents*. arXiv.org, https://doi.org/10.48550/arXiv.2107.12808.

Generative, Algorithmic & Evolutionary Design

COMPARING DESIGN STRATEGIES

A System For Optimization-based Design

LIKAI WANG[1], PATRICK JANSSEN[2], TUNG DO PHUONG BUI[3], and KIAN WEE CHEN[4]
[1]*Xi'an Jiaotong-Liverpool University.*
[1]*wang.likai@outlook.com, 0000-0003-4054-649X*
[2]*White Lioness Technologies*
[2]*patrick@janssen.name, 0000-0002-2013-7122*
[3]*National University of Singapore.*
[3]*akibdpt@nus.edu.sg, 0000-0003-3951-8192*
[4]*Princeton University.*
[4]*chenkianwee@gmail.com, 0000-0003-4397-7324*

Abstract. The paper presents a system for supporting optimization-based design exploration, with a special focus on comparing design strategies for early design stages. The system is developed to facilitate designers to conduct multiple optimization runs so as to compare the strengths and weaknesses among different design strategies. The system connects Rhino-Grasshopper to a web-based evaluation server. In Rhino-Grasshopper, designers can encode different design strategies and input these parametrized strategies into the optimization workflow that connects to the evaluation server. The evaluation server records the data sent from Rhino-Grasshopper and provides various interfaces to visualize and analyze optimization results. To demonstrate its utility, a case study is presented to demonstrate the utility of the system in urban design.

Keywords. Computational Design, Design Optimization, Building Performance, Design Exploration, Design Tool.

1. Introduction

Over the past decade, a significant number of studies have been conducted investigating the application of computational design optimization to architectural and urban design. In these studies, the optimization is typically based on a specific design strategy and is aimed to help the designer identify high-performing design variants within the selected strategy. However, this application paradigm also limits the usage of computational design optimization in early design stages. Early-stage design often involves parallel exploration and comparison of multiple alternative design strategies, such as building typologies (Xu et al., 2019), building façade schemes (Wang, Zhang, et al., 2022), and building formal characteristics (Wang, 2022). In this regard, incorporating computational design optimization into such parallel design exploration processes can further enhance its practical utility in architectural and urban design.

Focusing on the above-mentioned need, this paper presents an interactive design system that supports optimization-based parallel design exploration. The system uses Rhino-Grasshopper as the parametric modeller, connected to a web-based evaluation server. During the optimization process, designs are generated in Rhino-Grasshopper and sent to the evaluation server, where they are evaluated and added to the design population. Once the optimization is completed, designers can quickly view design options in the stored population. The system also allows designers to compare design options across different populations with various data analysis and visualization interfaces. As such, it provides designers with an integrated and friendly environment to perform optimization using various parametrized design strategies, which further enables the comparison of the strengths and weaknesses of these strategies.

To put this study into context, this paper first provides a brief overview of the relevant studies. The main body of the paper describes the system and demonstrates its utility using a case study. We conclude the paper by discussing its relevance to computational design optimization.

1.1. COMPUTATIONAL OPTIMIZATION FOR ARCHITECTURAL AND URBAN DESIGN

Computational design optimization has been widely adopted in research for addressing complex performance-related design problems by searching for high-performing solutions within a design space delineated by a parametric model. In addition, to facilitate designers to gain a deeper understanding of the optimization result, several tools were also developed, such as Design Explorer, Performance Map (Wortmann, 2017), and Biomerpher (Harding & Brandt-Olsen, 2018), which offer various approaches to visualize or organize the data obtained from the optimization. In these studies, the case-study design example typically emphasizes helping designers identify critical design parameters and reveal latent relationships between design parameters and building performance. However, more complex architectural design activities are often absent in most of these studies, such as design reflection, modification, as well as parallel and iterative exploration.

More recently, a few studies have been conducted focusing on integrating computational design optimization into early-stage design exploration and design development processes (Janssen et al., 2022; Wang, 2022). In these studies, designers are assumed to conduct multiple optimization runs with different parametrized design strategies, in order to compare their strengths and weaknesses. Such studies demonstrate the potential of how computational design optimization can be intertwined into a broader design exploration or development process. When it comes to utility, although running multiple optimizations with different parametrized strategies is technically feasible, this often entails a large amount of data analysis and visualization work. The tedious data processing work could discourage designers to adopt such optimization-based design exploration in practice. Therefore, a platform to assist designers in managing and organizing data from multiple optimization runs can further promote the application of such optimization-based design exploration.

In regards to data management for parametric design, certain recent studies propose platforms to facilitate designers to compare different parametric models (Cristie & Joyce, 2021; Woodbury et al., 2017). These platforms are able to record designers' editing in the parametric model and display changes in the generated design. As such, it facilitates designers to document their design development process and offers a more interactive parametric design environment. These studies highlight the importance of integrated platforms for data management and organization in complex parametric design activities. Nevertheless, these existing platforms only support documenting changes in parametric models, and no equivalent platforms can be used in optimization-based design exploration. Thus, the system presented in this study addresses the lack of an infrastructure that can allow designers to fluidly compare and inspect results from multiple optimization runs.

2. System for Parallel Optimization-based Design Exploration

To address the challenge mentioned above, the presented design system provides designers with an integrated platform that can manage and analyse results obtained from multiple optimization runs. The development of the system is aimed to facilitate designers to perform optimization-based design exploration for multiple design strategies. As such, the strengths and weaknesses of different design strategies can be investigated. In contrast, conventional design optimization applications typically focus on one design strategy, and as a result, it is unlikely to help designers identify flaws and deficiencies in the strategy.

2.1. SYSTEM OVERVIEW

Fig. 1 illustrates the framework of the presented system. The system connects Rhino-Grasshopper with an evaluation server to establish a loop for design optimization. The evolutionary algorithm is executed in Rhino-Grasshopper. The evaluation server performs simulation and design evaluations and provides various user interfaces for information communication. There are two components connecting Rhino-Grasshopper to the evaluation server: one for sending building geometry to the server, and one for receiving performance scores from the server. During the optimization, the evaluation server records the information of each design variant in a database, including generative parameters, performance scores, an identity number (ID), and an image of the design variant.

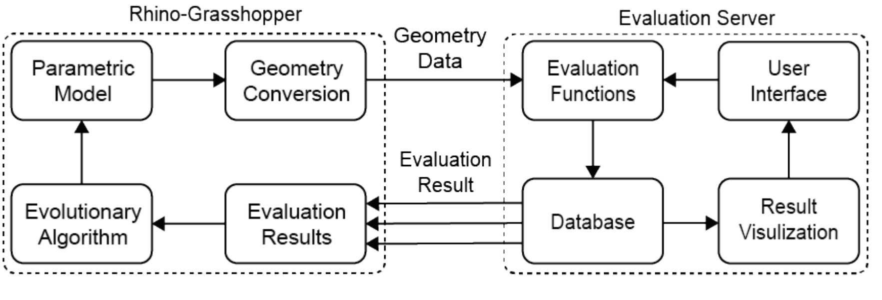


Figure 1. System Framework

2.2. OPTIMIZATION PROJECT

In this system, optimization projects are used as a container for storing and managing optimization settings and results. In each optimization project, designers are required to specify design evaluation settings, surrounding urban contexts, and other critical design conditions such as locations and weather data resources (Fig. 2 left).

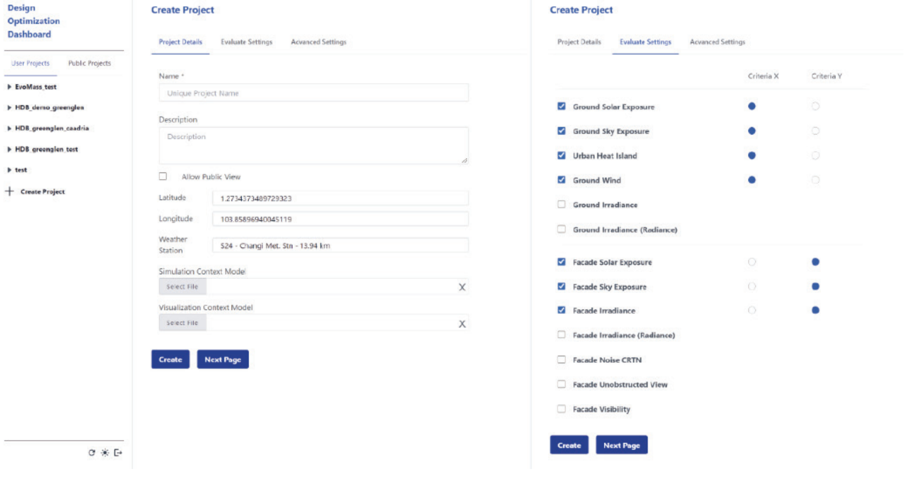


Figure 2. "Create Project" page (left) and "Evaluation Settings" page (right)

When creating an optimization project, designers need to define the evaluation metrics on the server (Fig. 2 right). Currently, the system provides a set of commonly-used evaluation metrics for measuring the performance of generated designs based on ground planes and building facades (Wang, Janssen, Do, et al., 2022). The division of ground planes and building facades results in a bi-objective optimization. This division is predefined in the system to reduce the demand for setting feasible optimization objectives.

The building simulation and performance evaluations are conducted by the evaluation server, using cloud-based parallel computing to shorten the time. This differentiates this system from ordinary simulation-based design optimization workflow in Rhino-Grasshopper, in which simulations are conducted using local computational resources.

2.3. OPTIMIZATION RUN

After the creation of an optimization project, the next step is to establish an optimization workflow that connects Rhino-Grasshopper to the evaluation server. In Rhino-Grasshopper, the parametric model needs to be connected to the two components mentioned earlier (Fig. 3). Once the design optimization workflow is established, the evolutionary component will retrieve all existing optimization projects in the evaluation server, from which the designer can select one project and then start the design optimization process.

Due to the use of parallel computing, most optimization runs can be finished within one hour for 200 to 1000 iterations of design generations and evaluations, which is conditional on how many evaluation metrics are included in the optimization project. This makes it possible for designers to conduct multiple design optimization runs in a reasonable amount of time.

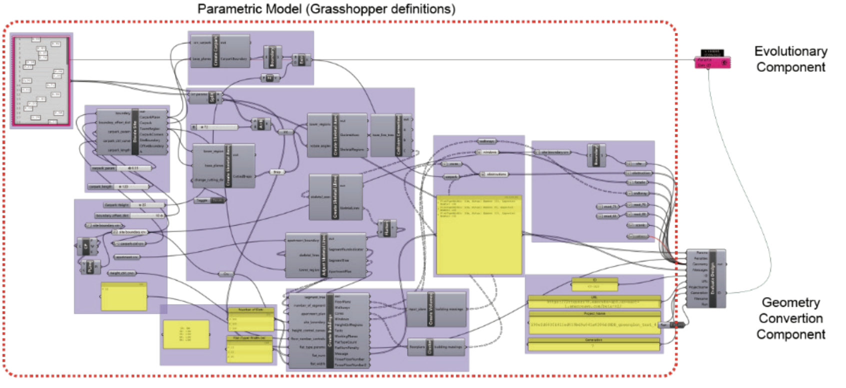


Figure 3. Design Optimization Workflow in Rhino-Grasshopper

Once the optimization is completed, designers can view the optimization result restored in the evaluation server through the user interface. The user interface offers several options to visualize the optimization result including a scatter plot, a parallel plot, and a list of generated designs. In addition, the scatter and parallel plots also have interactive features that allow designers to screen out designs of interest. Designers can select an area in the scatter plot or define numerical ranges in the parallel plot, and then the designs not within the selected area or range will be filtered out. This function facilitates designers to identify commonalities in the selected designs, which enhances the information extraction from the optimization result.

Beyond conducting a single optimization run, a major difference in using this system is that designers can perform multiple optimization runs based on different parametric models or a parametric model with different meta-parameters. When there is more than one optimization run, every run is listed under the selected optimization project (see Fig. 2 left). In addition, essential information about each optimization run is also visualized altogether, which allows designers to easily compare and identify the overall strengths and weaknesses of each parameterized design strategy. These features will be further demonstrated in the following case study.

3. Case Study

This case study demonstrates how the presented system can be applied to urban design and how the proposed design optimization mode with multiple optimization runs can help designers better understand the design problem and make more informed design decisions. The case study describes a residential precinct development project in Singapore, which consists of a standalone carpark building and multiple high-rise residential buildings.

In this case-study project, the position of the carpark building plays a decisive role in the overall configuration of building layouts and the performance of the residential building. In addition, its position is also subject to various design constraints beyond performance such as circulation and noise control. Therefore, choosing different carpark building positions can be viewed as different design strategies incorporating these above-mentioned factors. As a result, we assume that the designer specifies several possible positions for the carpark building and attempt to investigate the impact of these positions on the overall performance of the precinct design.

In this case study, a skeletal parametric model was used for design generation (Wang et al., 2023; Wang, Janssen, & Chen, 2022). Based on the input carpark building position, the parametric model can generate residential buildings according to a set of skeletal axes. The variation of the skeletal axes in terms of length and direction results in high variability of building arrangements and precinct configurations. The use of this parametric model offers sufficient design differentiations for the optimization to search for high-performing solutions. Fig. 4 illustrates the design variation for two different carpark building positions.

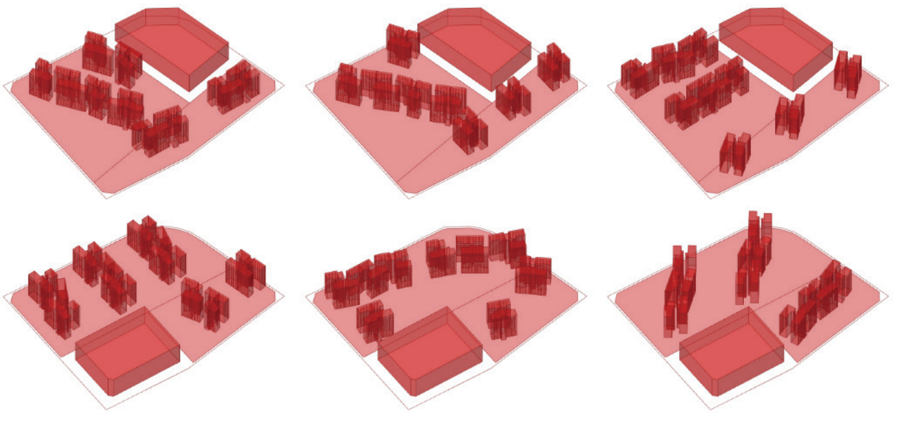


Figure 4. Generated designs with different carpark building positions

For the optimization, seven evaluation metrics are considered, including façade solar exposure, façade sky exposure, façade irradiance, façade noise, ground solar exposure, urban heat island, and ground wind permeability. The first four metrics are combined as a score to evaluate the performance in terms of building façades, while the last three are combined as another score to evaluate the performance in terms of ground planes. Regarding design generation, four carpark positions are defined, and each occupies one of four corners of the building plot. Lastly, for each optimization, a design population with 30 individuals is used and is evolved for eight generations.

3.1. OPTIMIZATION RESULT

With four different carpark positions, each is used for one optimization run. Fig. 5 demonstrates the optimization results in the scatter plot. In the scatter plot, the optimized designs from each design population are represented as dots and shown with different colors. In addition, Pareto fronts are also drawn according to each optimized design population. The x-axis indicates the performance related to the ground planes, and the y-axis indicates the performance related to the building façade.

Regarding design exploration, the result demonstrated in the scatter plot provides an overall comparison of the design populations based on different carpark positions. Among the four populations, Design Population 3, shown with green, has a carpark on the northeast corner resulting in a prominent advantage compared to the other three design populations. Following it, Design Population 1 (yellow) with the carpark on the southwest corner and Design Population 0 (blue) with the carpark on the northwest corner can also achieve acceptable performance. In addition, Design Population 1 can surpass Design Population 3 on the y-axis (façade performance). In contrast, Design Population 2 (purple) has the lowest overall performance, while it should be highlighted that one of the designs achieves the highest score in terms of ground performance (the x-axis).

Figure 5. Scatter Plot (screen capture)

The visualization of the design populations helps designers quickly grasp the overall trade-off between precinct configurations based on different carpark positions. Following this analysis, designers can further explore regions of interest to extract more detailed design information from the optimization. For example, designers can select a region in the scatter plot as shown in the red rectangle in Fig. 5. The selected designs are then displayed in the below design list, in which designers can investigate the formal feature from the images. According to these designs, it can be found that the designs in this region typically have north-south-oriented residential buildings as this orientation primarily prevents the buildings from receiving excessive solar radiation on afternoons.

Apart from the scatter plot, designers can also conduct numerical analysis using the parallel plot. Fig. 6 shows the parallel plot of the optimization results, which can reveal the correlation between each evaluation metric and the overall performance of the design. In addition, when using the parallel plot, designers can screen out designs based on different evaluation metrics, and the designs within the selected range will be displayed in the below design list as shown in Fig. 6 bottom. For example, in Fig. 6, the designs with high ground wind permeability are selected (pink lines), and the listed designs typically have east-west oriented residential buildings.

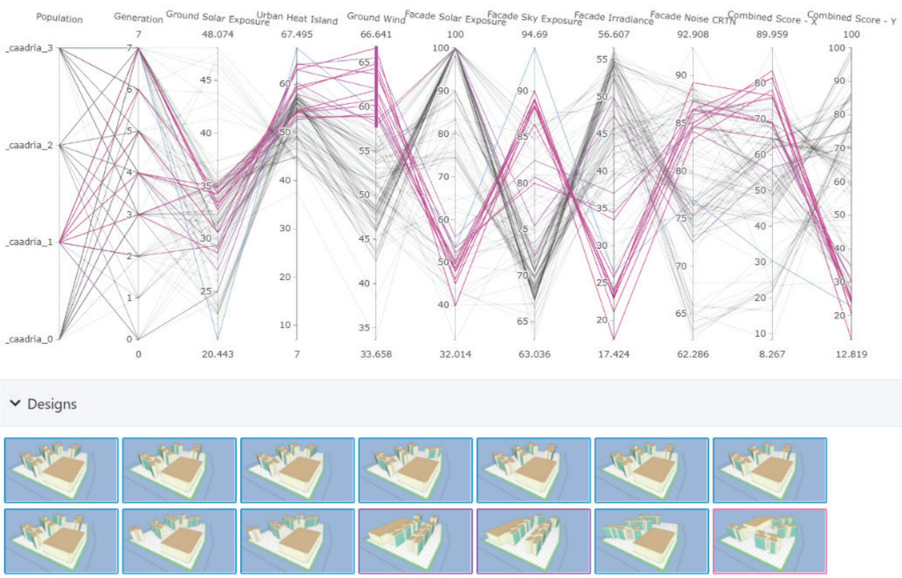


Figure 6. Parallel Plot (screen capture)

Lastly, designers can further delve into the detail of each design or compare two designs of interest. Fig. 7 shows the comparison viewer of the system. In this view, designers can compare two designs from any of the design populations. In addition, designers can also run the simulation and visualize the simulation result in this viewer to obtain a more intuitive contrast between the two designs. In Fig. 7, two designs with different carpark positions and residential building orientations are displayed with the wind permeability simulation visualized. The visualization of the simulation result shows that the west-east-oriented residential towers have less impact on wind flow and are more conducive to achieving desirable wind permeability.

To summarize, the above analysis is not purely aimed to identify the best solution for the design problem but allows designers to gain an understanding of the design problem at different levels and from different angles. Therefore, the analysis also helps designers be aware of the latent weaknesses when choosing other carpark positions (e.g. Population 2), which may then lead to adopting other measures to mitigate the negative effect. Thus, this also differs from optimization purely focused on numerical improvement and design refinement.

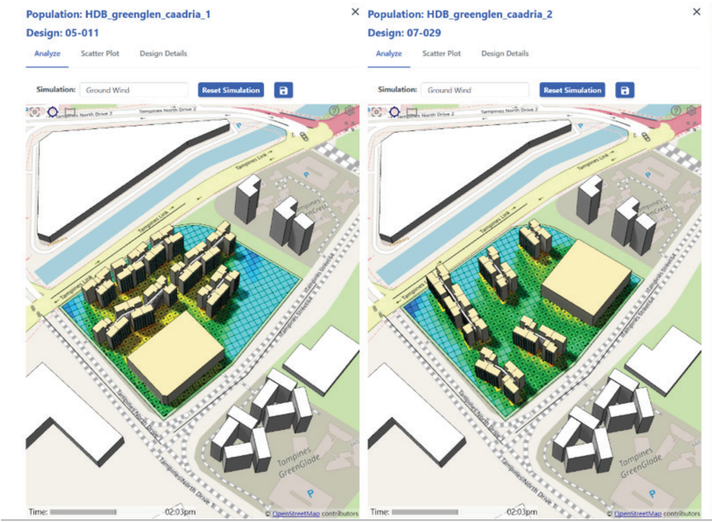


Figure 7. Design Comparison View (screen capture)

4. Discussion and Conclusion

The above case study displays the main features and functions of the presented system. Compared with other existing tools or platforms, the presented system provides an integrated environment for performance simulation, computational optimization, and data management. Moreover, the data organization structure is specially designed for optimization-based design exploration, in which the optimization is not merely focused on one optimization run based on one design strategy but allows for broader design exploration of a variety of alternative design strategies. Nevertheless, a challenge of using this system is that designers have to propose multiple design strategies and conduct parametric modelling, which can be intelligently demanding for some designers.

The case study also demonstrates the utility of the system and the corresponding design exploration approach in early-stage urban design. The emphasis on multiple optimization runs encourages designers to explore different alternative strategies at the outset of the design process. This facilitates designers to obtain a broader view of the design problem. In contrast, when using conventional computational design optimization approaches, the numerical improvement achieved by the optimization may leave designers with false confidence in the performance strength of the selected design strategy or design scheme, and the excessive focus on just one design strategy can also reinforce design fixation and stifle creativity.

The above discussion also highlights the need for recalibration of the research and tool development of computational design optimization for architectural and urban design. Computational design optimization is initially designed for solving well-defined problems, which makes it easier to apply to engineering design problems focusing on quantitative improvement. However, the poorly-defined and "wicked" nature of architectural and urban design makes conventional optimization paradigms often fail to effectively support architects' design exploration and synthesis activities that are more focused on qualitative aspects characterizing the design problem.

To conclude, this paper presents a system supporting optimization-based design exploration. The system offers a data management structure for organizing results from multiple optimization runs. As demonstrated by the case study, the system facilitates a parallel design exploration for multiple design strategies, which helps designers identify strengths and weaknesses within these strategies and obtain a broader view of the design problem at early design stages. With the case study, the paper discusses the utility of the presented system in early-stage architectural design and highlights the current mismatch between computational design optimization research and architectural design.

References

Cristie, V., & Joyce, S. C. (2021). Versioning for parametric design exploration process. *Automation in Construction*, 103802.

Harding, J., & Brandt-Olsen, C. (2018). Biomorpher : Interactive Evolution for Parametric Design. *International Journal of Architectural Computing,* 144–163. https://doi.org/10.1177/1478077118778579

Janssen, P., Bui, T. D. P., & Wang, L. (2022). Möbius evolver: Competitive exploration of urban massing strategies. In I. As, P. Basu, & P. B. T.-A. I. in U. P. and D. Talwar (Eds.), *Artificial Intelligence in Urban Planning and Design* (pp. 293–321). Elsevier.

Wang, L. (2022). Optimization-aided design: two approaches for reflective exploration of design search space. *International Journal of Architectural Computing,* 20(4), 758–776. https://doi.org/10.1177/14780771221134958

Wang, L., Janssen, P., & Chen, K. W. (2023). Evolutionary Optimization of Benchmarks: Parametric Typologies for Generating Typical Designs. In J. S. Gero (Ed.), *Design Computing and Cognition'22* (pp. 699–717). Springer International Publishing.

Wang, L., Janssen, P., & Chen, K. W. (2022). Evolutionary Design of Residential Precincts, A Skeletal Modeling Approach for Generating Building Layout Configurations. *Proceedings of the 27th International Conference of the Association for Computer- Aided Architectural Design Research in Asia* (CAADRIA) 2022, April, 415–424.

Wang, L., Janssen, P., Do, T., Bui, P., & Chen, K. W. (2022). A Rapid Design Optimization Framework. *Proceedings of the 40th Conference on Education and Research in Computer Aided Architectural Design in Europe* (ECAADe 2022), 1, 619–628.

Wang, L., Zhang, H., Liu, X., & Ji, G. (2022). Exploring the synergy of building massing and façade design through evolutionary optimization. *Frontiers of Architectural Research,* 11(4). https://doi.org/10.1016/j.foar.2022.02.002

Woodbury, R., Mohiuddin, A., Cichy, M., & Mueller, V. (2017). Interactive design galleries: A general approach to interacting with design alternatives. Design Studies, 52, 40–72.

Wortmann, T. (2017). Surveying design spaces with performance maps: A multivariate visualization method for parametric design and architectural design optimization. International Journal of Architectural Computing, 15(1), 38–53. https://doi.org/10.1177/1478077117691600

Xu, X., Yin, C., Wang, W., Xu, N., Hong, T., & Li, Q. (2019). Revealing urban morphology and outdoor comfort through genetic algorithm-driven urban block design in dry and hot regions of China. Sustainability (Switzerland), 11(13), 3683.

PAINTERLY EXPANSION

MARC GIBSON[1]
[1] *School of Architecture & Urban Design, RMIT Architecture | Tectonic Formation Lab*
[1] *marc.gibson@rmit.edu.au , 0000-0002-2886-4500*

Abstract. Painterly Expansion seeks to develop a novel workflow that interfaces digital sculpting processes and algorithmic geometry production to generate complex spatial and surface qualities. The research speculates that the combination of human intuition and algorithmic processes will generate novel hybridised geometric outcomes. The workflow seeks to mask indexical qualities inherent in isolated processes of mesh based folding algorithms and procedural deformation techniques whilst exploring opportunities to extract emergent geometric qualities. The methodology is evaluated recursively with each iteration split into three stages, that encompass the layering of algorithmic, parametric, and manual modelling processes. The purpose of shifting between modelling processes is to counter the linear application of indexical qualities that originate from a singular process with the intention of producing emergent topological and ornamental effects. The algorithmic approach explores the application of differential growth logic in combination with procedural retopology and manual sculpting.

Keywords. Digital Sculpting, Mesh Expansion, Recursive Algorithm, Topology Generation, Intricate Ornamental Form.

1. Introduction

Painterly Expansion is an exploration of integrating bottom-up algorithmic workflows with gestural digital sculpting techniques by developing a live link between software packages. A key objective of this process is to enhance the occurrence of hybridized aesthetics as a result of integrating a human authored design intent into generative processes that would otherwise operate iteratively in a closed manner. The iterative process of moving between an algorithmic growth process and manual sculpting is an intentional attempt to subvert the occurrence of aesthetic indexical traits that originate from a singular process. The design procedure seeks to generate a form that has the potential to operate spatially and ornamentally through large-scale folds and procedural expression of materiality. The base geometry is generated with a half edge mesh expansion algorithm that was designed to interface with the authored output of a 3D sculpted form. The secondary direct input is the vertex colour painted by the designer

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 231-240.

which dictates the location & intensity of the growth. The directionality of the expansion algorithm also attempts to balance the internal logic of local curvature, global normal directionality, and approximated volumetric depth. The form is generated iteratively by passing from Zbrush to Grasshopper to be sculpted & reformed. The interaction between processes is an attempt to give rise to new qualities and reduce the indexicality back to a singular process.

This 'Painterly' builds upon the concept of 'Behavioral Formation' explored by Roland Snooks with regard to integration of 'the architect's design intention within emergent processes' (Snooks, 2021, p. 7) and seeks to further imbue user design interaction through gestural modelling techniques. The generative workflow utilises layers of methods that operate recursively to derive a topology with high levels of ornamental complexity. The recursive iteration is a notional “cycle” where all specified operations/conditions are completed. The flexibility of a recursive algorithm is that the conditions can be influenced by previous iterations and have an impact on the outcome of future cycles. This can lead to the emergence of qualities that can be generated from sophisticated interactions between rules. The painterly expansion research looks to leverage emergent qualities derived from algorithmic processes and embeds the workflow within manual polygonal and sculptural workflows.

A consideration of this research is in the qualitative reading of geometry by evaluating the presence of 'indexical qualities' in a resulting geometry. This evaluation refers to aesthetic markers that are representative of the underlying methodology or software utilized in its formation. This research refers to 'emergent qualities' as features of geometry without a clear indexical trait and speculates this occurrence can be reinforced through the layering of complex systems and authored subjective input. Geometry produced through 3D modelling software can embed visual traits that indicate the technical methodology utilised (Figure 5). This is present in the foundational representation of geometry using a standardized class of Mesh, Surface or SubD. Where a Mesh ‘Type’ or ‘Class’ of geometry is constructed from a hierarchy of elements such as “Vertices”, “Edges”, and “Faces”. The inherent ‘Type’ topology impacts the way the user generates form. A similar ideological expression occurs when considering the traits of algorithmic structures such as multi-agent systems to describe swarm behaviour where the parameters of a single boid impact the larger structure. In this instance, the research is particularly interested in the emergent qualities that start to occur when the algorithmic tendencies of a multi-agent growth system are iteratively undermined or exaggerated by digital sculpting processes.

2. Zbrush Live Link

Digital sculpting software such as Zbrush operates on the high-resolution point (vertex) position of mesh-based geometry. The operation acts as a deformation simulating the physical displacement of matter. Typically, this allows the designer to develop more gestural geometric patterns and forms that are otherwise difficult to generate in polygonal modelling software. The mesh deformation and retopology mechanisms with Zbrush are used as a 'break' in the process to shed the indexical mesh structure of a previous iteration. This sculpting process is paired with flexible retopology or 'remeshing' techniques such as ‘dynamesh’ that regenerate the foundational order, position, and edge directionality of the underlying geometry. Retopology is not restricted to sculpting applications and can be solved using a number of techniques such as voxel based projection and ball pivot surface reconstruction. The technique used for generating the mesh has an impact on the surface quality and polygonal edge direction. Zbrush contains a number of procedural operations housed in the 'deformation' class that displace vertex positions. The location and gradient of deformation effects can be controlled via a process of 'masking'. A mask acts on a per vertex basis as a float value to dampen the impact of manual brushing or deformation operations.

The 'live link' software Chameleon was developed as an application that interfaces Grasshopper and ZBrush to exchange geometry and parametrically set up command sequences (Gibson, 2020). The application provides the user access to the geometry construction classes of ZBrush whilst maintaining the root editability of a parametric model. The plugin is written in C# .NET Framework utilizing RhinoCommon and Grasshopper Libraries. The algorithm compiles ZScript called via the Grasshopper node-based user interface. The methodology for ZScript compilation and scheduling via Windows Operating System Subprocesses was inspired by Nik Willmore's Python Interface (Wilmore, 2017). The novel aspects of this proposed research do not originate from the scheduling of geometry and commands to Zbrush but instead are centred on the workflow of moving between an algorithmic output and manual intervention through digital sculpting. The 'live link' is simply the mechanism that reduces friction in the movement of data between applications rather than the source of specific innovation in this instance.

The core development of Chameleon includes multi-threaded mechanisms for optimised conversion of geometry types and commands into a simplified wavefront .obj format. This hybridized process is an attempt to reduce the friction typically experienced when pausing and exporting the 3d model between applications. A secondary utility of Chameleon is accessing the procedural deformation methods inside Zbrush for refining the surface smoothness by geometric features. The Zbrush Live Link workflow was utilised in the 'Cloud Affects' 2019 Shenzhen Biennale Installation (Snooks and Samartzis, 2019). In collaboration with Tectonic Formation Lab the latest release of Chameleon is also integrated into the design and print optimization processes for the 3d printed pavilion 'Unclear Cloud' displayed at the National Gallery of Victoria Sampling the Future exhibition in 2021 (Snooks and Samartzis, 2020).

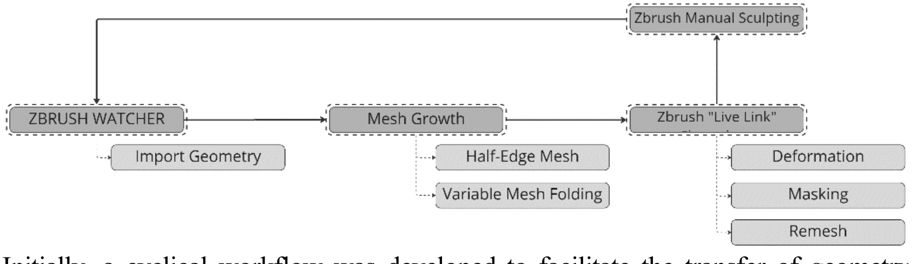
3. Mesh Expansion Context

The core generative algorithm takes inspiration from the biological process of 'differential growth' which describes the cellular growth of flora in response to both internal and environmental stimuli (Barlow, 1989). The resulting forms that follow this process can lead to self-similar folds in the edge condition of leaves and blooming flowers such as Lilium Casa Blanca (Liang, Haiyi, and L Mahadevan, 2011). The expansion methodology offers the potential for spatial applications due to the inherent collision avoidance built into the logic. The algorithmic approach was not specifically selected for the direct geometric translation of natural phenomena but for the procedural process of subdivision as a method for mesh expansion. A parallel motivation for selecting Differential growth as an underlying process was due to the indexical legibility of the collision parameters present in topological folds.

The use of differential growth is present in computational design strategies and has been implemented by NERVOUS in the FloraForm jewellery collection (Rosenkrantz, 2015). The FloraForm process utilises the Half Edge Mesh Data Structure as a vehicle for subdividing mesh face topology. The governing principles determining the subdivision is translated by comparing the distance of vertices. A force is applied to each vertex per iteration based on the nominal distance between vertices noted as 'stretch' and relative face angle encapsulated as a 'bend' force. The expansion is restricted to the edge condition of a modelled mesh geometry resulting in an aesthetically representative outcome of the natural differential growth process. A similar methodology is present in Long Nguyen's University of Stuttgart Institute for Computational Design and Construction C# workshop series (Nguyen, 2018). The mesh expansion utilises an optimized Half-Edge Mesh that is not restricted to the edge condition and instead operates on internal points in each iteration.

4. Painterly Expansion Cyclical Workflow

Figure 1. Painterly expansion cyclical workflow



Initially, a cyclical workflow was developed to facilitate the transfer of geometry between mesh folding processes and digital sculpting. The effect and methodology are captured in three stages to evaluate the effect of each layered process. The motivation is to subvert and curate typical indexical traits present in the underlying mesh folding algorithm by procedural and manual sculpting methods. The emergent aesthetic qualities are captured under the descriptor of the 'painterly' to denote the presence of an author's hand. Figure 1 depicts the painterly expansion operates off a recursive concept where the design geometry procedurally passes through modelling processes. The initial stage 'Mesh Growth' begins with a seed mesh based geometry that is folded for a specified quantity of iterations. The geometry is automatically sent to Zbrush and procedurally masked, deformed and remeshed. Once the procedural operations have concluded the user can manually sculpt the geometry in Zbrush. After this the user runs a macro that automatically returns the geometry and commands back to grasshopper. A file system watcher picks up on the activity and begins the next cycle of the expansion process. The workflow is captured in 3 sections below as to compare the effect of each stage in the process. These sections are separated to demonstrate the outcome of the isolated algorithmic process, combination of algorithmic and procedural and full application of the painterly expansion workflow.

4.1. ALGORITHMIC WORKFLOW

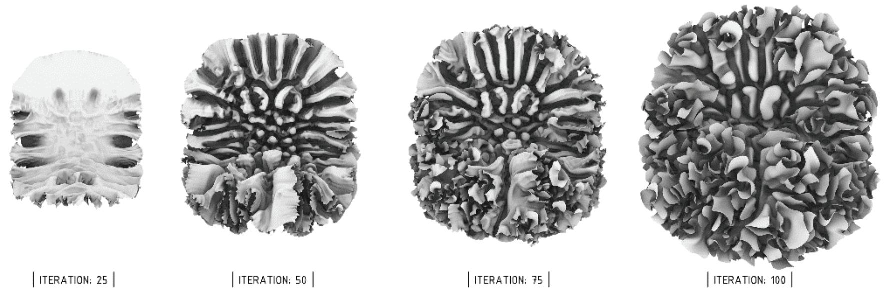


Figure 2. Isolated mesh expansion algorithm geometric output

The painterly expansion growth mechanism is built upon the Plankton implementation of a Half Edge Mesh Data Structure and fundamentally draws from the methodology demonstrated by Long Nguyen (Nguyen, 2018). The Plankton library utilises the RhinoCommon geometry classes which is underpinning the mesh class used in Daniel Piker's Kangaroo3D Grasshopper Application (Piker, 2020). Plankton was selected as it operates in conjunction with the methodology to allow for a change in the internal mesh topology which would otherwise require lengthy computation between iterations. This is due to a standard mesh structure requiring a reordering of faces, edges vertices if any mesh element is removed or subdivided. Each vertex receives a 'collision force' that has a magnitude and direction determined by neighbouring vertices. This force is applied to all mesh vertices and is not restricted by edge conditions. Figure 2 depicts the mesh growth process which was uniformly weighted on all vertices with no secondary procedural or sculpted operations. The mesh folding considers local bending in two stages of separate ranges. This introduced two scales of creasing that iteratively increased toward iteration 100.

4.2. ALGORITHMIC AND PROCEDUAL WORKFLOW

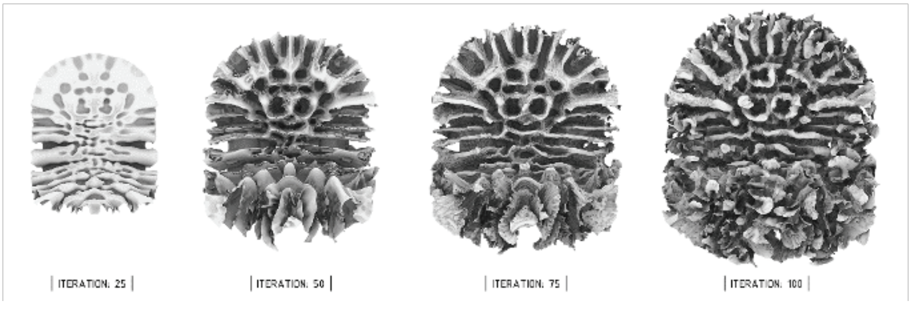


Figure 3. Isolated mesh expansion algorithm & parametric Zbrush geometric output

The procedural operations in Zbrush operate at the conclusion of each algorithmic cycle and do not directly intervene in the expansion logic when processing. Figure 3 depicts the procedural extent of the painterly expansion process without user intervention in Zbrush. The procedural commands mask features of the geometry by curvature to localize the effect of deformation. As a result certain features can become exaggerated in a positive feedback loop. This is apparent by the two near symmetrical indentations at the centre top of the form in Figure 3. The operations also reinforce the presence of horizontal striations when compared to Figure 2. The legibility of these defined folds could be subverted by introducing more effective remeshing & smoothing procedures. Figure 1. An image with a caption

4.3. APPLICATION OF PAINTERLY EXPANSION WORKFLOW

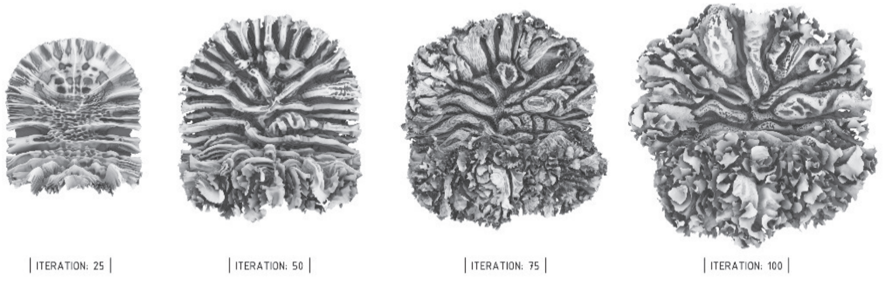


Figure 4. Full application of painterly expansion workflow

Figure 4 demonstrates the full cyclical output of the painterly procedure. The expansion operated iteratively where the growth algorithm would run for a set series of loops and then be automatically sent to Zbrush for procedural processing and manual sculpting. Upon import into Zbrush the geometry is masked by local curvature as a method of isolating sharp creases in the geometry. A base black and white colour are applied to highlight these regions. Once procedural commands are concluded the user is able to intervene by manually sculpting. The geometry is smoothed and remeshed each iteration which reduced the appearance of mesh artefacts such as the indentations present in iteration 25. The remeshing process alters the continuity of the mesh face edge flow allowing new folds to emerge across the surface as visible in iterations 50 to 75 in Figure 4.

5. Process Evaluation

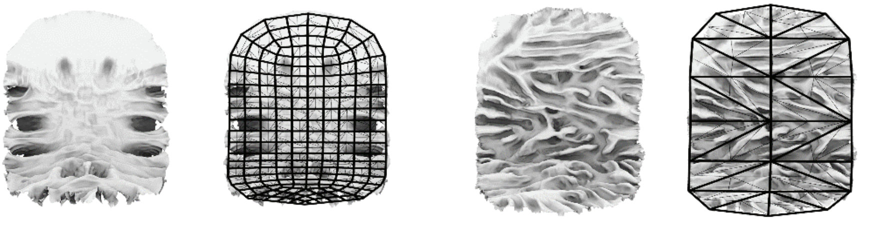


Figure 5. Mesh tessellation comparative impact on mesh expansion

The initial remeshing phase has a relatively evident impact on the presence and directionality of mesh folds. This is visible in Figure 5 where the quad faces lead to horizontal striations whilst the presence of mesh triangulation introduces diagonal streaks. This could be remedied at the mesh folding stage where the bending angle considers a larger range mesh vertices rather than immediately connected vertices. This methodology requires a relatively high population of vertices to sample each iteration which can have an impact on performance. Another method was to iteratively remesh the geometry applying a priority to areas of high curvature. The result of this method can lead to better edge flow of mesh faces that align more closely to the underlying topology. A potential drawback is that when operating procedurally in absence of manual intervention the remeshing process can reinforce existing pathways in a positive feedback loop which can reduce the potential for emergent patterns to occur.

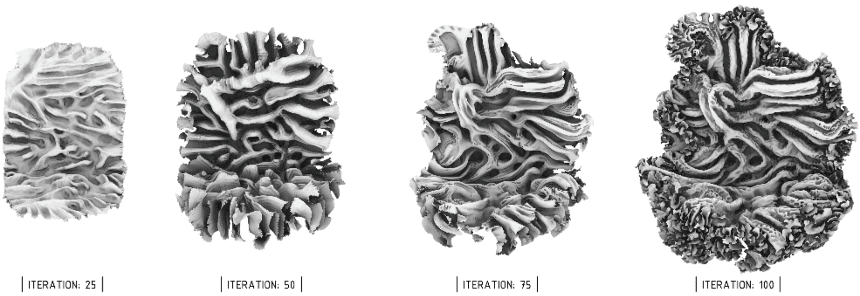


Figure 6. Full application of painterly expansion workflow with triangulated mesh input

Figure 6 demonstrates the painterly procedure applied to a triangulated base mesh. The expansion initially operated on a lower resolution mesh which allowed control over the initial directionality of folding. The folds increase in detail and topological complexity as iterations progress as visible between iterations 50 & 75.

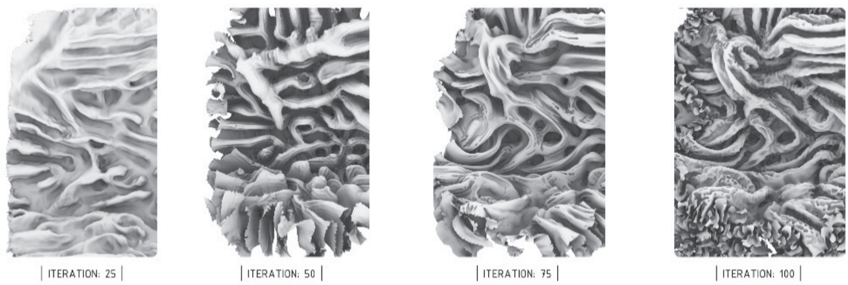


Figure 7. Detail view comparing painterly expansion iterations

Large scale sculptural moves were intentionally avoided in Zbrush. Instead, fine grain detail was reshaped in internal regions to encourage the presence of corrugations along the larger folds. This is apparent in Figure 7 where the central knot is iteratively inflated, smoothed and remeshed between iteration 25 and iteration 75. The geometry is iteratively smoothed and remeshed to better control quad edge flow to reinforce folding. The legibility of indexical mesh folding geometries appears to shift scale as the geometry increases in resolution evident between iteration 50 and iteration 100 in Figure 7. The resolution of the geometry is a by-product of both iterative remeshing and increased topological complexity. The addition of manual sculpting appears to reduce the uniformity of mesh folding when compared to the isolated algorithmic & procedural examples in Figure 2 and Figure 3.

6. Process Development

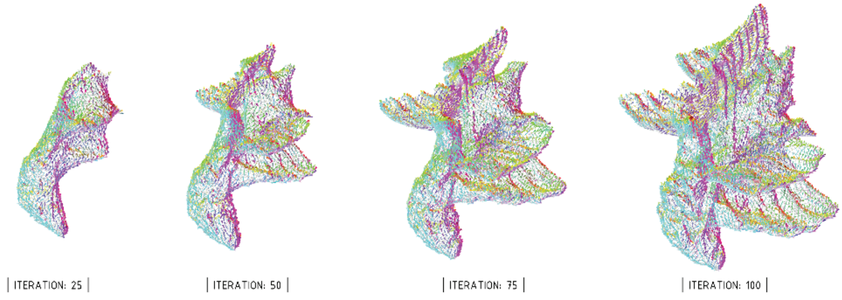


Figure 8. Vertex agent directionality per painterly expansion iteration

An alternate expansion process was developed to further augment the interaction between manual and algorithmic processes by adding a visual mechanism to control the application and location mesh folding. Figure 8 depicts the directionality of each agent per iteration. One consideration was to store vertex colour within the mesh as an effective variable that can be translated as a float to impact the magnitude of force applied to specific regions. A secondary motivation was to use the vertex colour as a visual key for the user to determine where to intervene and further subvert the consistent application of folds across the geometry. The vertex colour class houses red, green, blue, and alpha values deposited by the user in Zbrush. Figure 9 shows the colour input used as a weighted value to control the methods used by the painterly expansion algorithm.

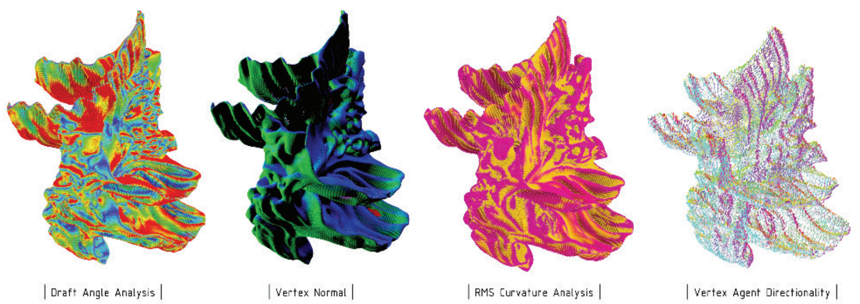


Figure 9. Local mesh analytics utilised in painterly expansion

The vertex colour regions denoted in white received a multiplier to the expansion search radius, an increase in normal influenced vectors and an increase in bend resistance. This was a mechanism to encourage larger topological extrusions whilst allowing the more localized fine grain folding to generate corrugations along surfaces. A secondary multiplier was applied to coloured ridge conditions as an effective 'attractor' to increase the tendency of outward growth. To prevent self-intersection each vertex maintained float values that represented local Root Mean Squared Curvature, relative normal direction, force direction and relative draft angle displayed in Figure 9. As this process did not focus on typical contour print requirements the weighting assigned to the vertex draft angle was effectively zero. A potential side effect of utilizing draft angles can lead to increased vertical expansion relative to the print plane. Building fabrication constraints into the algorithmic and sculpting process can be another way of further developing the painterly expansion technique.

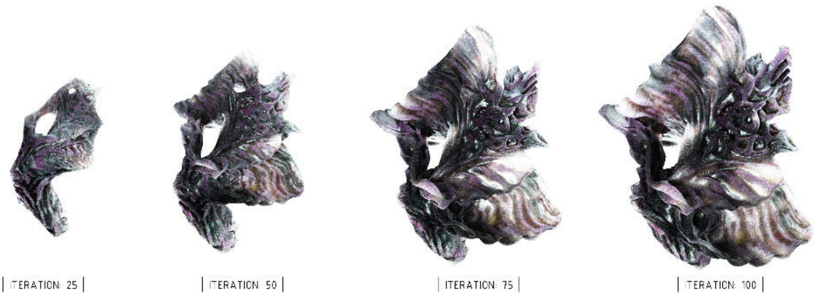


Figure 10. Rendered output demonstrating emergent topological folds & material transitions

7. Conclusion

The expanded form bares trademarks of the differential growth process however the geometry in Figure 10 demonstrates new qualities such as pitted knots and twisted corrugations of the sculpted centre mass from iteration 75 to 100. The inflated presence of these knots appears to occur as a result of a positive feedback loop when the geometry is masked by curvature and then deformed outward. The subsequent iterations create sharper ridge lines that the curvature analysis highlights during the masking process. The presence and intensity of these highlights can be manually subverted when the user intervenes in the process. Alternatively, a more sophisticated masking response could be introduced that considers the relative thickness of a mesh to prevent over sharpening thin regions. Critically the process still appears to be dominated by an outward force which unintentionally generates the appearance of a clear centroid in Figure 10. This could be remedied by using a more asymmetric base geometry or applying more varied directional forces to the vertices. The vertex colour force weighting appears to be an effective solution for directing fine grain detail in the generative algorithm however this relegates control of the macroscopic form to the algorithm. The painterly process was cautiously restricted in the remeshing procedures to maintain consistent vertex order and UV coordinates for material application at the visual representation stage. Further experimentation in balancing the impact of top-down modelling in relation to algorithmic process should be explored. Painterly expansion is an exploration of generative form making by embedding gestural sculpting techniques into recursive algorithmic processes to produce intricate forms. The integration of manual sculpting techniques within multi-agent algorithmics generates novel opportunities to further integrate the intent of the designer.

Acknowledgements

Thank you to Dr Roland Snooks for his invaluable advice and mentorship.

References

Barlow, Peter W. (1989) Differential Growth in plants—A Phenomenon That Occurs at All Levels of Organization. Environmental and experimental botany 29.1 (1989) (pp.1–5). https://doi.org/10.1016/0098-8472(89)90034-8

Gibson, M. (2020). Chameleon: Zbrush Live Link. Retrieved November 21, 2022 from https://www.food4rhino.com/app/chameleon

Liang, H., & Mahadevan, L. (2011). Growth, geometry, and mechanics of a blooming lily. Proceedings of the National Academy of Sciences, 108(14), 5516–5521. doi:10.1073/pnas.1007808108

Nguyen, L. (2018) University of Stuttgart Institute for Computational Design and Construction. Workshop Live-Streaming: C# Scripting and Plugin Development for Grasshopper. Retrieved November 21, 2022, from https://www.icd.uni-stuttgart.de/teaching/workshops/workshop-live-streaming-c-scripting-and-plugin-development-for-grasshopper/

Piker, D. (n.d.). Kangaroo3D: Interactive Physics/Constraint Solver. Retrieved November 21, 2022, from http://kangaroo3d.com/

Rosenkrantz, J., Louis-Rosenberg, J. (2015).'Floraform – an exploration of differential growth' Retrieved November 21, 2022, from https://n-e-r-v-o-u-s.com/blog/?p=6721

Snooks, R. (2021) Behavioral Formation: Volatile Design Processes and the Emergence of a Strange Specificity. Actar.

Snooks, R., Samartzis, P., Boman, C., Mohamed, H., Bao, N., Bufton, G., Gibson, M., Parry, C., Wang, D. (2020). Cloud Affects, Shenzhen Biennale, Shenzhen, China.

Snooks, R., Samartzis, P., Boman, C., Mohamed, H., Bao, N., Gibson, M. (2021). Unclear Cloud, Sampling The Future, National Gallery of Victoria, Victoria, Australia.

Willmore, N. Nik’s (2017) Zbrush Commander. Retrieved November 21, 2022, from https://www.grasshopper3d.com/forum/attachment/download?id=2985220%3AUploadedFile%3A1857873

RESEARCH ON PLAN GENERATION DESIGN OF PARKING LOT BASED ON INTEGER PROGRAMMING

MENGTING ZHANG[1], PEIWEN LI[1,2] and GANG YU[3]
[1,2,3]*Harbin Institute of Technology, Shenzhen.*
[2]*XKOOL Technology Co., Ltd.*
[1]*lexi_zmt@163.com, 0000-0002-3963-5121*
[2]*lpwseu@163.com, 0000-0003-2463-8968*
[3]*yuarch@hit.edu.cn, 0000-0002-9352-4385*

Abstract. As an auxiliary space in engineering projects, the process of parking lot design is often under-appreciated, which is highly repetitive, time-consuming, and leaves little space for creativity. Thus, optimizing existing design methods and improving design efficiency is an aspiration of the industry. Based on integer programming, this paper focuses on the surface parking lot as the main research object and explores the possibility of changing the parking lot plan design from manual to computer-generated. According to the research of relevant norms, this paper classifies the parking lot plan in terms of organisation mode and summarises 27 kinds of parking lot plan modes by analysing the state of the inner ring, traffic organisation, and parking angle. The model can efficiently solve the problems related to parking lot arrangements, such as whether to arrange parking spaces in the inner-ring of the contour and how to choose the internal parking angle, which saves manpower and improves the design efficiency and quality.

Keywords. Parking Lot Plan Generation, Integer Programming, Generative Design, Visualization, Optimisation Algorithm

1. Introduction

Currently, parking lot design relies primarily on manual labour with fluctuating outcomes, and indicators such as parking capacity are closely linked to the designer's professional competency. Parking design, as an auxiliary space for construction work, provides a less creative space, which is not only time-consuming but also difficult to optimise. However, the parking design has formed a mature workflow, either parking lane planning or parking angle selection can be calculated to achieve the most optimal results. Therefore, using the computer to generate parking lot plans can improve design efficiency, and promote the development of disciplines and industries.

In the field of computer generation, various complex algorithms have been applied to different parking lot optimisation scenarios. For example, (Song & He, 2004) used genetic algorithms to calculate the maximum number of parking spots based on manual layouts, (Wang & Xie, 2019) used grid-segmentation and greedy-algorithm to solve

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 241-250.

the optimal layout of parking plans, and (Huang et al., 2020) proposed a parking space planning system based on pattern search algorithms. These algorithms have a more complicated solution process and thus inevitably cost more time. Whereas, mathematical modelling of problems can effectively reduce the time required to solve them. (Babushkina et al., 2021) developed a linear programming model to find the optimal ratio between the number of parking spaces and the available area through multi-criteria optimisation. (Crisostomo & Baldovino, 2019) also applied linear programming to model a specific park lot and solved it using the simplex method, but neither of them gave a specific solution for the layout. Furthermore, fewer studies have considered the greening component of surface parking and have not tied the algorithm to the visualisation design.

In this paper, we propose the use of integer programming methods to establish and solve relational equations for the layout of parking spaces in order to achieve the automated generation of surface parking designs. Integer programming, which refers to integer linear programming (ILP) in this paper, is a mathematical optimisation method, the essence of which is to find the optimal solution of functions under constraint. In comparison to evolutionary models that require multiple iterations, exactly compiled ILP models have the characteristic of efficiently obtaining globally optimal solutions. Based on existing parking standards, this study summarises the organisation pattern of the parking lot layout and uses it as a clue to build and solve the ILP model, while building a visualisation model using Grasshopper.

2. Organization Patterns for Parking Lot

The current layout of the parking lot is only categorised by the mode of parking angle, including vertical (perpendicular to the aisle), parallel (parallel to the aisle), diagonal (at a certain angle to the aisle), and staggered (containing two or more types of vertical, parallel and diagonal). The parking angle mode, however, can only describe the angle of the parking slot, without presenting features of traffic organisation and the parking lot's division. This paper decomposes the parking layout into the inner-ring state, traffic organisation and parking angle mode, which are classified and combined into the corresponding organisation mode.

Primarily, based on the basic design idea of the car park, a ring road is usually set along the contour of the site to ensure the connectivity of parking aisles inside. The placement of parking spaces between the ring road and the site boundary, coined as inner-ring parking in this article, has an effect on the total number of parking spaces, which is an important feature for describing the parking layout, including the full inner-ring, the semi-inner-ring (part of the inner-ring area is arranged with parking spaces), and no inner-ring parking. Besides, according to the relationship between parking aisles' direction and the longer side of the parking lot, the plan can be categorised into the horizontal arrangement, the vertical arrangement, and the hybrid arrangement. For the purposes of this paper, horizontal arrangement means that the direction of arranging the parking spots is parallel to the longest edge, and aisles of vertical arrangement are perpendicular to the longest side. Meanwhile, it is common for both horizontal and vertical arrangements to occur at the same site, which in some cases may lead to a more efficient parking layout. Finally, since parallel parking is usually not the dominant form of a parking lot, only perpendicular parking, diagonal parking, and staggered parking are considered in this study.

On the basis of the above classification, categories in the three model languages are arranged and combined to produce 27 organizational patterns (Table 1) that can be used to describe the majority of parking lot plans.

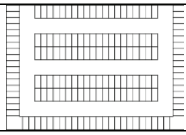
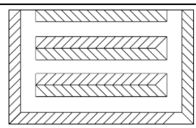
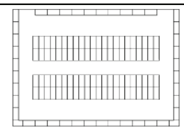
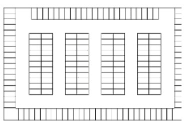
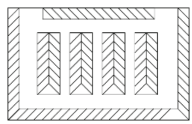
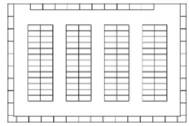
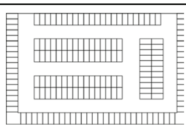
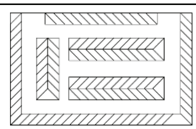
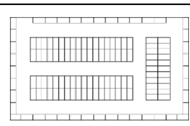
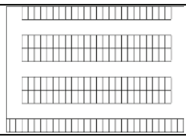
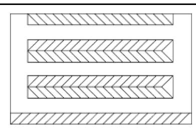
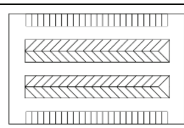
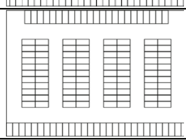
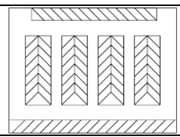
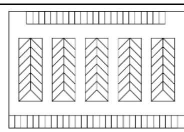
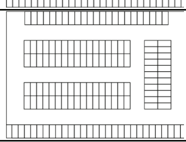
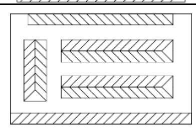
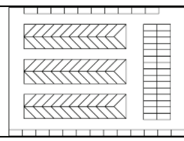
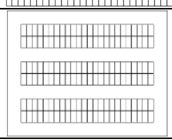
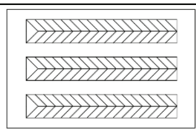
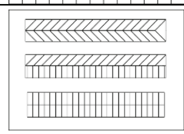
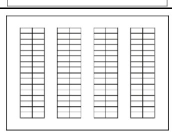
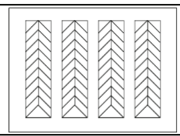
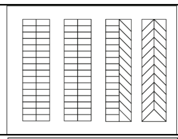
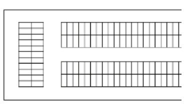
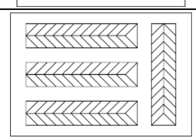
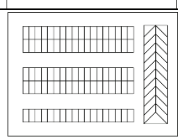
the inner-ring state	traffic organisation	parking angle		
		vertical angle	diagonal angle	staggered angle
full inner-ring parking	horizontal arrangement			
	vertical arrangement			
	hybrid arrangement			
semi-inner-ring parking	horizontal arrangement			
	vertical arrangement			
	hybrid arrangement			
no inner-ring parking	horizontal arrangement			
	vertical arrangement			
	hybrid arrangement			

Table 1 Organisation patterns of parking lot plan

3. ILP Modelling

3.1. CORRELATIONAL RESEARCH

While each organisation pattern presents different possibilities for parking lot layouts depending on the actual situation, generic design methods can be developed to suit different scenarios. (Abdelfatah & Taha, 2014) constructed an integer programming model to find the best-fitting parking angle in a rectangular parcel, but did not consider the mixing of different parking angles which can achieve higher parking efficiency. An integer programming formulation, proposed by(Yildirim et al., 2019), could solve a hybrid problem with multiple angles, but sets a large number of decision variables, which affected the efficiency of the solution to some extent. These studies can only describe the semi-inner-ring parking mode and did not consider greening factors in their models. This paper thus builds on the study of Yildirim et al. to construct an ILP model, that focuses on rectangular sites, for all organizational patterns, considering the relationship between greenness and parking spaces.

3.2. CONSTRAINTS AND ELEMENTARY PARAMETERS

The main constraints of the design of the parking lot are as follows: (1) The driveways and parking spaces should be arranged inside the site. In the case of a rectangular parking lot, its long side width is set to R_L and its short side width is set to R_W. (2) Driveways and parking lanes are expected to be parallel or perpendicular to the site boundary. (3) All paths need to form loops, avoiding the occurrence of end roads. The width of the driveway, W_d, is set to 6 meters by default. (4) The parking module must parallel to the parking lanes and the angle of parking spots should conform to the industry standard, including 0°, 30°, 45°, 60°, and 90°. (5) For surface parking, a linear green belt is typically used to plant trees for shade providing, and its width, W_g, is not less than 1.5m.

Based on the parking organisation patterns described in the previous section, 20 parking configuration modes can be summarized, including 5 single-row modes (index $i = 1,2,3,4,5$) and 15 double-row modes (individual combinations of 5 parking angles and combinations in pairs, $C_5^1 + C_5^2$, index $i = 6,7,...,20$). Figure 1 shows the relevant parameters that correspond to the configuration mode. For a single-row mode, the width W_i consists of the width of parking slot a and the minimum access width required for it. While for the double-row mode, W_i is increased by the width of parking lot b, with the access width W_{ab_i} taken to be the larger one between the minimum access width required for a and b. The length of the parking slots is L_{a_i} and L_{b_i}, which are used to calculate their quantity under length constraint of site length

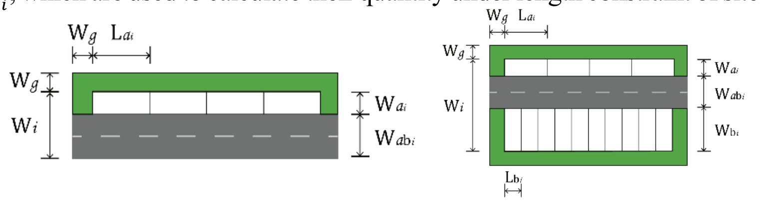


Figure 1 Parameters of parking configurations

3.3. GENERATIVE MODEL.

3.3.1. Horizontal Arrangement with Semi-inner-ring and No Inner-ring Parking

The goal of the model is to maximise the sum of parking spaces' quantity for the chosen configuration. Considering the relationship between the location of the green belt and the parking spots, however, 20 modes that increase the greenery width directly will inevitably generate redundant or insufficient green belts (Figure 2). Thus, in this study, single-row spaces are divided into those with and without green belts. The top and bottom inner-ring spaces are the optional single-row with greenness, and in their opposite side, there are two fixed single-row spaces, one of which is without the green belt. The top and bottom driveways, with a width of W_d, are also fixed to ensure the selectivity of double-row modes (Figure 3).

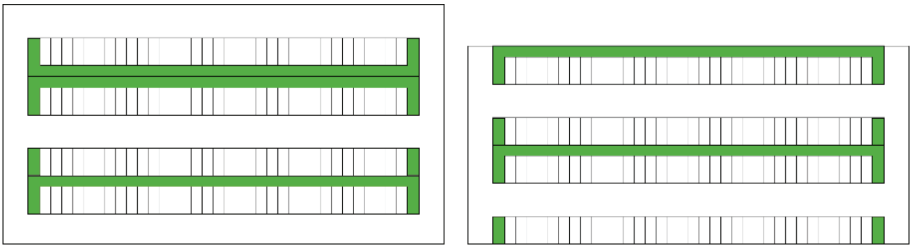


Figure 2 Redundant or insufficient green belts

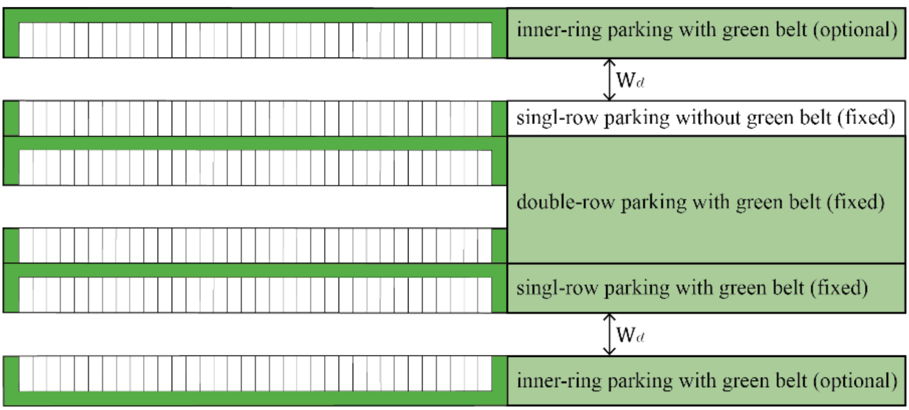


Figure 3 Optional and fixed parking modes

In order to solve parking problem with horizontal arrangement in rectangular site, the following ILP model is established:

$$\text{Maximise} \sum_{i=6}^{20} S_i \times x_i + \sum_{i=1}^{5} S_i \times f_i + \sum_{i=1}^{5} S_i \times m_i + \sum_{i=1}^{5} S_i \times n_i \quad \text{(Formula 1)}$$

$$\text{subject to } \sum_{i=5}^{20} (W_i + W_g) \times x_i + \sum_{i=1}^{5} (W_{a_i} + W_g) \times f_i + \sum_{i=1}^{5} (W_{a_i} + W_g) \times m_i + \sum_{i=1}^{5} W_{a_i} \times n_i \leq R_W - 2W_d \quad \text{(Formula 2)}$$

$$\sum_{i=1}^{5} f_i \leq 2 \quad \text{(Formula 3)}$$

$$\sum_{i=1}^{5} m_i = 1 \quad \text{(Formula 4)}$$

$$\sum_{i=1}^{5} n_i = 1 \quad \text{(Formula 5)}$$
$$x_i \in Z^+ \quad i = 6,7,\ldots,20 \quad \text{(Formula 6)}$$
$$f_i \in Z^+ \quad i = 1,2,3,4,5 \quad \text{(Formula 7)}$$
$$m_i \in Z^+ \quad i = 1,2,3,4,5 \quad \text{(Formula 8)}$$
$$n_i \in Z^+ \quad i = 1,2,3,4,5 \quad \text{(Formula 9)}$$

The objective function (Formula 1) is maximising the total number of parking spaces and the decision variable x_i refers to the number of ith configuration, with its value coefficient S_i, which is the number of parking spaces of each mode, calculated under the constraint of R_L dynamically. f_i denotes the number of rows in the upper and lower inner-ring area, which is limited to no more than two rows by Formula 3. m_i and n_i are the numbers of rows of single-row modes with and without greenery, respectively, Formula 4 and 5 restrict them to a fixed number. Formula 2 ensures that the sum of the width of the modes selected within the width of the site, R_W, while Formula 6 to 9 constrain all variables to integers. It's worth stating that this model can solve not only the question with horizontal arrangement, but also questions with vertical arrangement, simply by permuting the two values, R_L and R_W. For the inner-ring area, however, the model only gives the results for the selection of the upper and lower spaces. Further optimization is required to describe the full inner-ring mode.

3.3.2. Full Inner-ring Parking

The parking lot plan can be decomposed into inner-ring parking and inside parking. The top and bottom inner-ring parking depended on the results of the first model, whereas the left and right sides of the inner-ring area must be solved by the second model. Based on the first solution, the number of rows of is counted for each of the 5 angles, denoted x, with its decision variable b_i referring to the number of parking spots of ith angle. e_i indicates the number of single-row modes used in the left and right sides, with the value coefficient N_i denoting the parking spaces' quantity subject to the width of the rectangular (Formula 10). The following ILP model may can be established:

$$\text{Maximize } \sum_{i=1}^{5} X_i \times b_i + \sum_{i=1}^{5} N_i \times e_i \quad \text{(Formula 10)}$$
$$\text{subject to } L_{a_1} \times b_1 + \sum_{i=1}^{5} (W_{a_i} + W_g) \times e_i \leq R_L - K_1 \quad \text{(Formula 11)}$$
$$L_{a_2} \times b_2 + \sum_{i=1}^{5} (W_{a_i} + W_g) \times e_i \leq R_L - K_2 \quad \text{(Formula 12)}$$
$$L_{a_3} \times b_3 + \sum_{i=1}^{5} (W_{a_i} + W_g) \times e_i \leq R_L - K_3 \quad \text{(Formula 13)}$$
$$L_{a_4} \times b_4 + \sum_{i=1}^{5} (W_{a_i} + W_g) \times e_i \leq R_L - K_4 \quad \text{(Formula 14)}$$
$$L_{a_5} \times b_5 + \sum_{i=1}^{5} (W_{a_i} + W_g) \times e_i \leq R_L - K_5 \quad \text{(Formula 15)}$$
$$\sum_{i=1}^{5} e_i \leq 2 \quad \text{(Formula 16)}$$
$$b_i \in Z^+ \quad i = 1,2,3,4,5 \quad \text{(Formula 17)}$$
$$e_i \in Z^+ \quad i = 1,2,3,4,5 \quad \text{(Formula 18)}$$

The above constraints, Formula 10 to Formula 15, ensure that the total length for each angle plus chosen width of the left and right inner-ring parking modes, restricted to no more than 2 columns by Formula 16, do not exceed R_L, which need to subtract a constant K_i, including the width of driveways and greenbelts. This model complements the selection of left and right inner-ring parking, and can solve the problem in horizontal arrangement or vertical arrangement with full inner-ring parking.

By continuing to deepen the model and decomposing the inside region of the plan into horizontal rows and vertical columns, a plan in hybrid arrangement can be solved. Through several experiments and analysis of the generated results, however, it was found that the results did not differ from the second model in the most of cases. Hence, it can be concluded that hybrid arrangement can not improve the efficiency of the parking lot layout, despite that it is useful in the site with different boundary, like L-plan and other concave polygons.

4. Visualisation and Generation Experiments

4.1. VISUALISATION

In the field of architecture, models have always been a common tool for observing designs and modifying them. With the introduction of parameters, it becomes simpler for designers to control the generation of plan, effectively solving the drawbacks of traditional processes that a slight move in one part may affect the situation as a whole. In contrast to deep learning, which operates on pixels, integer programming generates computable results that, when combined with visualization tools, are capable of translating numerical results into vectorized models. Built on the Rhino platform, this study uses the Grasshopper plugin for the surface parking plan to visualise the generated results.

Generation of the parking plan can be divided into three stages, including parameter setting, inner-ring parking generation and inside parking generation. The basic parameters such as the rectangular boundary curve and width of the driveway are input, and the ILP algorithm is written in python and solved by using the Cplex library to obtain the model solution. Distinguish between the inner-ring parking and the inside blocks and output the corresponding generation results respectively. The algorithm framework is shown in Figure 4 and Figure 5 shows the specific visual generation process, which takes the horizontal arrangement as an example.

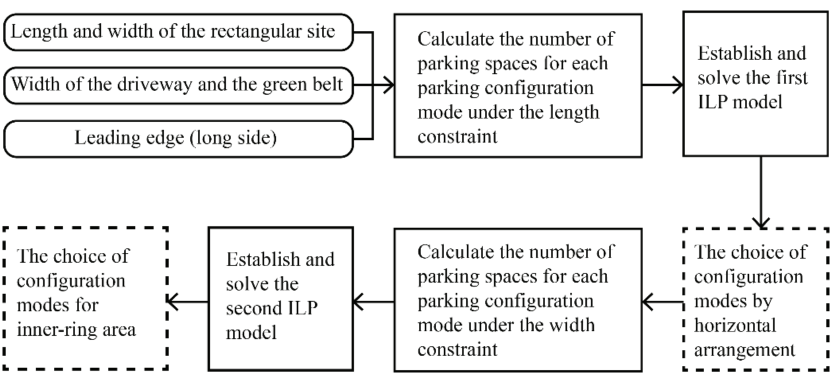


Figure 4 Generation process

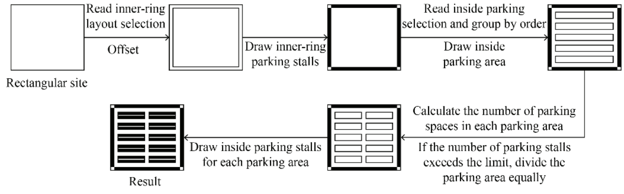


Figure 5 Visual generation process

4.2. GENERATION EXPERIMENTS

In this study, the following three rectangular boundaries were tested in horizontal and vertical arrangement, counting the number of parking spaces, the average area of each parking spot and the time cost in the generation process. Figure 6 shows the plan examples generated by the visualization tool, and relevant data are shown Table 2.

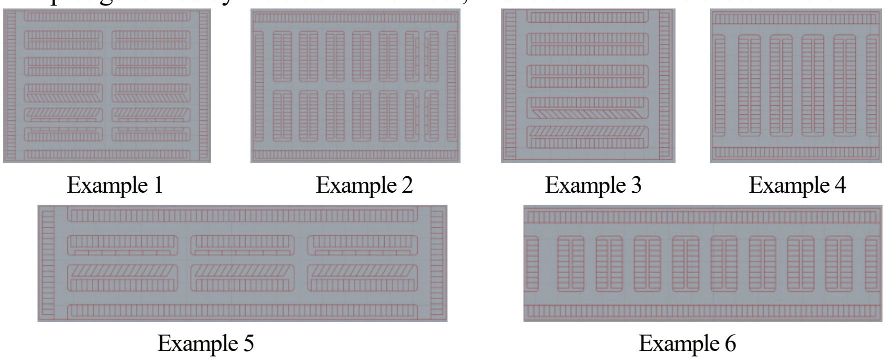


Figure 6 Examples of generated plan

Example of generated plan	Boundary Size	Arrangement	Number of parking spaces	Average area	Generation time
Example 1	133m × 97m	Horizontal arrangement	484	26.7 ㎡	456ms
Example 2		Vertical arrangement	468	27.6 ㎡	460ms
Example 3	90m × 80m	Horizontal arrangement	276	26.1 ㎡	318ms
Example 4		Vertical arrangement	282	25.5 ㎡	303ms
Example 5	180m × 50m	Horizontal arrangement	338	26.6 ㎡	386ms
Example 6		Vertical arrangement	326	27.6 ㎡	356ms

Table 2 Experiment of generation

The purpose of the three models, with aspect ratios close to 1:1.5, 1:1 and 1:3.5 respectively, was to obtain a comparison of the effectiveness of parking lot layout in horizontal and vertical arrangement with moderate, similar, and highly variable aspect ratios. Surprisingly, the difference between the two is not distinct in either case. Even with aspect ratios close to 1, the vertical arrangement can generate more parking spaces. The slight difference in length and width of the boundary led to different parking modes selected, which is evidence that the model is able to choose the most appropriate layout under constraints.

Following the experiments above, the conclusion was that the average area of the model-generated parking spaces remained less than 28 square meters. In addition, if no more than 500 parking spaces are generated, results can be generated in under 500 milliseconds, which has a significant advantage over complex optimization algorithms such as genetic algorithms in terms of the computation time needed for this method.

5. Discussion

This paper studies the computer generation of rectangular parking lots, discusses the classification of parking lots, forms the generation algorithm based on ILP, and constructs the generation model. Unlike the current generation software, such as TestFit, the purpose of this algorithm is to find the optimal or better layout, and the number of parking spaces that can be generated is determined by the boundary. The TestFit is results-oriented and calculates the required minimum size according to the different number of target parking spaces. The solution of this study is more suitable for the generation of parking spaces at fixed sites.

This study takes only rectangular boundaries as the research object, but this method can be extended to non-rectangular cases. In most linear programming solutions, the width of the rectangle determines the choice of the angle of the horizontal, in other words, the length of the rectangle determines the choice of the angle of the vertical arrangement. Therefore, in the case of non-rectangles, such as irregular convex polygons, the basic parking space layout can be achieved by constructing the smallest rectangle on the leading edge, and then the parking space layout of the boundary can be determined by calculation. Therefore, solving the layout of parking spaces on rectangular boundaries through ILP can become the basic logic for solving other boundaries.

In addition, in real situations, the parking lot generation model should also consider more specific landscape design and lighting facilities and other systems. For example, taking tree spacing as a parameter, trees will be arranged at the same time as the green belt is generated. Lighting systems, charging pile facilities, monitor systems and other security measures should also be considered as basic parameters to participate in the algorithm calculation. For example, according to the user's demand for new energy vehicles' parking space, consider the boundary that is most suitable for placing charging piles, reserve a space of 0.2 meters to 0.3 meters, and set vertical parking space to participate in the calculation. Another example is to combine the lighting system with the green belt, set the installation interval length according to different lamp specifications, and generate a complete parking lot layout.

6. Conclusion and Future Work

This study summarises the common plan organisation patterns in 27 according to existing standards related to parking lot design. Based on this, constraints of parking layout are analysed, a multilevel integer programming model is built, and the Grasshopper plugin is used to achieve the visual generation of the solution. For sites with rectangular boundaries, this model is capable of generating solutions effectively, solving the issue of whether to arrange parking spaces in the inner-ring area and the selection of parking modes inside the site. The model improves the number of parking spaces on the site, in contrast, it greatly reduces the time required in the generation, and systematically accounts for the relationship between green belts and parking spaces.

The model is currently restricted to rectangular solutions and has difficulty in deriving results for hybrid arrangement, in addition to not considering possible obstacles in realistic parking lots. In future research, irregular site boundaries and obstacle restrictions should be included.

References

Abdelfatah, A. S., & Taha, M. A. (2014). Parking Capacity Optimization Using Linear Programming. Journal of Traffic and Logistics Engineering, 2(3) http://doi.org/10.12720/jtle.2.3.176-181

Babushkina, A., Petrochenko, M., Kukina, A., & Astafieva, N. (2021). Optimizing parking lot design by Generative design approach. E3S Web of Conferences, 263, 4042. http://doi.org/10.1051/e3sconf/202126304042

Crisostomo, C., & Baldovino, R. G. (2019Parking Space Optimization Using Simplex Method Linear Programming. Paper presented at the 2019 International Conference on Mechatronics, Robotics and Systems Engineering (MoRSE).

Huang, X., Gao, Y., & Duan, X. (2020An Autonomous Parking Space Planning System Based on Pattern Searching Algorithm. Paper presented at the 2020 IEEE 8th International Conference on Smart City and Informatization (iSCI).

Song, Z. Z., & He, W. Z. (2004). Optimized Parking Lot Design of Trade Center Based on Genetic Algorithm. Mathematics In Practice and Theory(1), 19-23.

Wang, B., & Xie, K. (2019). Design of Open Parking Lot Based on Grid-Segmentation and Greedy-Algorithm. World Scientific Research Journal, 5(9), 268-277.https://doi.org/ 10.6911/WSRJ.201909_5(9).0030

Yildirim, G., Yildirim, B., & Satir, B. (2019). Design of a Rectangular Parking Lot Using a Cutting-Stock Formulation. Cankaya University Journal of Science and Engineering, 16(1), 1-16.

GENETIC ALGORITHM-BASED BUILDING GEOMETRIC OPENING CONFIGURATIONS OPTIMIZATION FOR ENHANCING VENTILATION PERFORMANCE IN THE HIGH-DENSITY URBAN DISTRICT

ZHICHAO LIN[1], SHI YIN[2*], WEI LIAO[3] and YIQIANG XIAO[4]
[1,2,3,4] *State Key Laboratory of Subtropical Building Science, School of Architecture, South China University of Technology, Guangzhou, China, 510640.*
[1]*lin-zhichao@outlook.com, 0000-0001-5803-6311*
[2]*archyinshi@scut.edu.cn, 0000-0002-5825-5749*
[3]*wlager@scut.edu.cn, 0000-0002-3911-107X*
[4]*yqxiao@scut.edu.cn, 0000-0003-2946-6261*
[*]*Corresponding author*

Abstract. The quality of the outdoor environment relates to urban ventilation performance. Poor wind conditions in high-density urban districts may lead to severe air pollution and deteriorate outdoor thermal comfort. The increase of openings in building geometry is one of the effective passive design strategies for enhancing the porosity of urban morphology and benefitting urban ventilation. However, the outdoor wind environment correlates with the opening configurations of building geometry complicatedly. For seeking the optimal opening configurations, a decision support tool is urgently needed. Our study proposes a genetic algorithm-based optimization workflow for the opening configurations of building geometry design by integrating Computational Fluid Dynamics simulation and parametric design. A residential block in Shenzhen, China is chosen as an example to show this workflow. The results demonstrate that when the porosity is 15%, the pedestrian-level mean wind speed, the wind speed dispersion, and the pressure difference between the front and rear of the downstream building can be optimized by 20.00%, 19.35%, and 183.33% on maximum. When the porosity is increased to 30%, these values are 42.22%, 16.13%, and 483.33%. The resultant opening distribution probability maps can support building design at an early stage to achieve a comfortable urban environment.

Keywords. Urban Ventilation, Building Openings, Building Porosity, Genetic Algorithm Optimization, Computational Fluid Dynamics.

1. Introduction

Urban ventilation performance is closely related to urban environmental quality. High-density development leads to severe urban heat island effects, largely affects low-level airflow, and worsens the urban environment. To mitigate these problems, numerous urban planning strategies have been proposed and investigated to alter urban morphologies and building configurations. Methods including low building packing and frontal area densities (Buccolieri et al., 2010; Di Sabatino et al., 2007) and low street aspect ratios (Baratian-Ghorghi and Kaye, 2013) were proposed to increase the ventilation rate and breathability of a city (Hang et al., 2012). However, these strategies appear less attractive in urban areas as they significantly lose valuable development floor areas.

This problem can be ameliorated by increasing the porosity ratio of buildings by adding building openings, which has been investigated in numerous previous studies (Yuan and Ng, 2012; Yuan et al., 2014). Openings can accelerate the wind velocity in street canyons by enhancing convection and diffusion processes to improve ventilation performance. Their findings have been incorporated into city or country planning policies to guide sustainable building design.

However, only increasing building porosity might not always achieve a comfortable wind environment. Wind conditions are associated with both the position and size of openings (Fan et al., 2017; van Druenen et al., 2019). Inappropriate opening configuration may result in extreme wind conditions. The effect of opening configurations on the wind environment is complicated, so it is difficult to get the ideal result only by limiting the porosity ratio.

For evaluating urban ventilation performance comprehensively, Computational Fluid Dynamics (CFD) simulation is the most common tool in early-stage architectural design. However, CFD simulation is time costly and the optimization process usually requires a large number of cases for supporting the final decision. It is vital to use efficient optimization approaches to reduce the demands of computational resources. In this perspective, the Genetic Algorithm (GA) can be applied as they are designed to approximate the global optimum solution while reducing the computational time. Many researchers have employed GA to find the optimal solution with the best ventilation performance in the process of urban or architectural design (Kaseb et al., 2020; Lee, 2007).

This study proposes a GA-based optimization workflow for the opening configurations of building geometry design by integrating CFD simulation and parametric design. A residential block in Shenzhen, China is taken as an example for generating the optimal opening configurations by the GA-based optimization workflow. Finally, the opening distribution probability maps will be developed as references for architects to improve urban ventilation performance at early-stage design.

2. Methodology

2.1. STUDY AREA

The block of Buji Xinsan Village, Shenzhen, China is selected as the study area (Figure

1). Shenzhen belongs to the Subtropical Monsoon Climate. Due to the influence of monsoon, southeast wind prevails in summer accompanied by high temperatures. As a typical high-density city in hot and humid areas, the dense building layout obstructs the airflow and deteriorates urban ventilation. Therefore, natural ventilation in summer is highly demanded enhancing thermal desperation and improving environmental quality.

The selected residential block is a typical compact residential neighborhood in Shenzhen mainly consisting of detached houses, and plate-type buildings with long interfaces. These plate-type buildings are wind barrier walls for the downwind blocks, that seriously deteriorate the surrounding environmental quality. As shown in Figure 1, we selected one of the plate-type buildings as the optimization target.

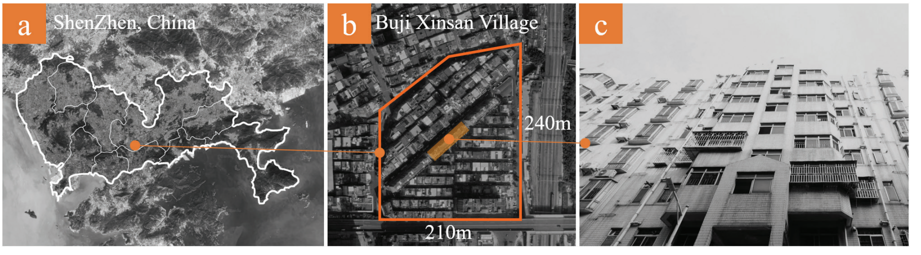


Figure 1. Study objective. a: the map of Shenzhen, China; b: the location of the selected residential block; c: the picture of the target building

2.2. WORKFLOW

The workflow consists of five steps (Figure 2). To begin with, acquire site information to establish a simplified block model; Secondly, set up the parametric model of the target building that can produce various solutions with different opening configurations by adjusting the module IDs; Then, a CFD simulation framework is needed, which simulates the ventilation performance of the solutions, and obtain three evaluation metrics as the optimization objectives; After that, build an optimization framework, and apply Non-dominated Sorting GA (NSGA-II) to search within a large number of solutions to find the Pareto optimal solutions; Finally, draw the opening distribution probability map according to the Pareto optimal solutions.

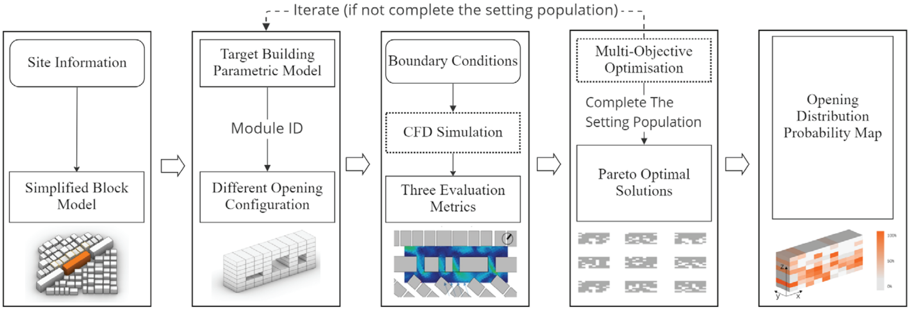


Figure 2. Workflow

2.3. MODEL GENERATION

2.3.1. Block Modeling

The block model is limited to a rectangular range of 150m * 150m centered on the target building and built using 3D modeling software, Rhino 7 (Figure 2). To reduce the time cost of CFD simulation, the buildings are simplified into cubes, and the surrounding vegetation and structures are ignored. The width and depth of the detached houses are 8 to 14m, and the heights are 9 to 15m (3 to 5 floors). The width and depth of the plate-type buildings are 63m and 14m respectively, and the height is 21m (7 floors). The building distances are 1 to 6m.

2.3.2. Opening Configurations

The target building is converted into a volume composed of identical cube modules. Different building porosity ratios can be achieved by controlling the number of removed modules. Changing the positions of removed modules will generate various opening configurations. In this paper, 2 scenarios with different porosity are studied, namely scenario 1 (S1) with 15% porosity and scenario 2 (S2) with 30% porosity.

The area of the original building is about 5000 ㎡. Each floor of the building is composed of 9 modules with a size of 7 * 14 * 3 m. In S1, the total number of initial modules is 63, approximately 7 floors, 9 modules per floor, and 9 to 10 modules need to be removed to meet the porosity of 15%.; In S2, the total number of initial modules is 72, approximately 8 layers, 9 modules per layer, and 22 modules need to be removed to meet the porosity of 30%. Each module is assigned an ID in sequence. The IDs of removed modules are given by the gene pool battery of Grasshopper. Taking S1 as an example, the principle of opening configuration generation and corresponding grasshopper screenshots are shown in Figure 3.

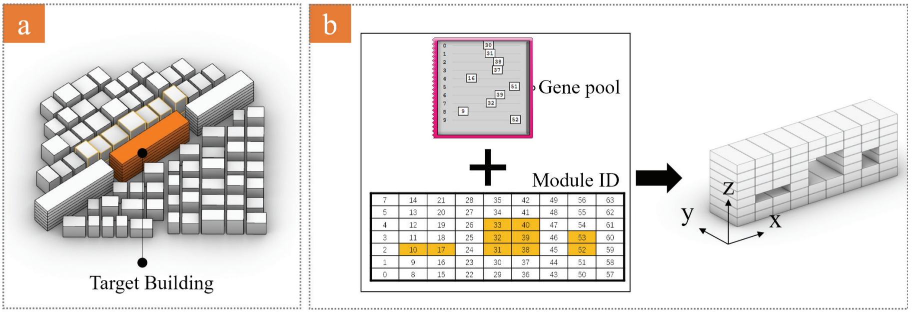


Figure 3. Parametric model generation. a: Simplified residential block model and target building; b: the principle of opening configuration generation

2.4. WIND ENVIRONMENT SIMULATION AND EVALUATION

2.4.1. Boundary conditions

The weather information in summer including June, July, and August are applied as the boundary conditions in the CFD simulations. The weather data file of Shenzhen is obtained through Energyplus (https://energyplus.net/weather), and the wind speed and direction are obtained using a Grasshopper (GH) plug-in Ladybug (Sadeghipour Roudsari & Song, 2019). The average wind speed in summer is 1.90m/s, and the highest frequency wind direction is 115 °.

2.4.2. CFD simulation

Wind environment simulation is carried out on the Butterfly (Sadeghipour Roudsari & Song, 2019), which is a plug-in on GH. It adopts OpenFOAM (OpenFOAM-v1912, 2019) to create and run CFD simulations. OpenFOAM is the most strictly verified open-source CFD engine available and can run multiple advanced simulation and turbulence models. To reduce the CFD simulation time, Reynolds Average Simulation (RAS) turbulence model is adopted, and the simulation grid quality and calculation steps are decreased accordingly. Detailed settings are shown in Table 1.

Parameter name	Setting details
Turbulence model	RAS model RNG k- ω
Top boundary, Lateral and Outlet boundary	5H, 5H and 15H(H=tallest building's height)
Cell size for the area of interest	2m
Cell expansion ratio	1.3
Cells shape	Unstructured grid
Cells count	560952
Iteration	500

Table1. Settings for CFD simulation

2.4.3. Evaluation metrics

We select pedestrian-level mean wind speed (Um), wind speed dispersion (σ), and the pressure difference between the front and rear of the downstream building (ΔP) as the evaluation metrics. Um is the mean value of wind speed values at each point, which is the most commonly adopted metric for evaluating wind environment; σ is based on the basic knowledge of statistics, which is used to reflect the degree of difference between wind speed values at each point, it is a reflection of the uniform distribution of wind speed; ΔP refers to the average air pressure difference between the windward and leeward building surfaces, which can measure the ventilation potential. Figure 4 shows the evaluation range of each metric. Table 2 shows The evaluation scope, formula, and standard.

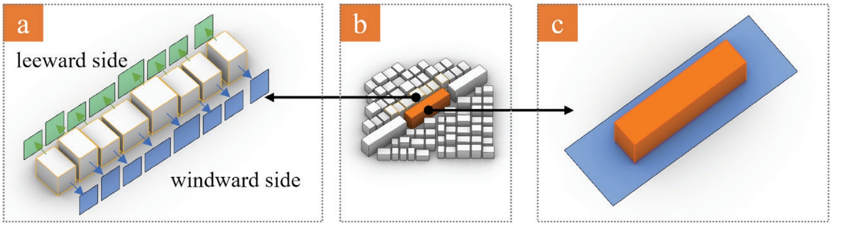


Figure 4. The schematic diagram for evaluation metrics. a: the evaluation range of ΔP; b: the target block and c: the evaluation range of Um and σ (blue area)

Metric	Definition	Formula	Evaluation Standard
Um	The mean value of wind speed at each point at a height of 1.5m	$\frac{1}{n}\sum_{i=1}^{n} U_i$	In the range of U<5m/s, the larger the value, the better
σ	The degree of difference between the wind speed and the average wind speed at each point at a height of 1.5m	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(U_i - U_m)^2}$	The smaller, the better
ΔP	The average air pressure difference between windward and leeward side building exterior surfaces	$\frac{1}{n}\sum_{i=1}^{n}(P_{1i} - P_{2i})$	The higher, the better

Table 2. Evaluation metrics

Where n represents the amount of data, U is the wind speed at the sampling point, P1 is the air pressure on the windward side of the building, P2 is the air pressure on the leeward side of the building

2.5. MULTI-OBJECTIVE OPTIMISATION

The optimization is operated on a GH plugin, Wallacei (Makki et al., 2021), which employs the NSGA-II as the primary evolutionary algorithm. NSGA-II is one of the most popular multi-objective GA. Compared with other algorithms, the NSGA-II can reduce degrees, fast convergence speed, and increase convergence (Deb et al., 2002).

In this study, the gene pool controlling opening distribution is taken as the gene. The three metrics for evaluating the ventilation performance, i.e., Um, σ, and ΔP, are set as the optimization objectives. Because Wallacei can only solve the minimization problem, Um and ΔP are multiplied by a negative value. The specific settings are shown in Table 3.

Parameter name	Setting details
Generation Size	20
Generation Count	20
Crossover Probability	0.9
Mutation Probability	1/n
Crossover Distribution Index	20
Mutation Distribution Index	20

Table 3. GA solver settings

3. RESULT

The optimization solutions are carried out on a computer configured as AMD Ryzen 9 3950X, 16 Cores, 32 Logical Processors. The optimization of the two scenarios has gone through 20 individuals per generation and 20 generations of operations (400 cases in total). Each scenario takes about 85 hours.

3.1. RESULT FOR S1

S1 (porosity of 15%) obtained 20 Pareto optimal solutions (Figure 5), and its metrics ranges are Um: 0.43·0.54m/s, σ: 0.25 to 0.31, ΔP: 0.08 to 0.17Pa. Only the σ value of some solutions with openings has increased compared to the simulation results of the original building without openings (Um=0.45m/s, σ=0.31, ΔP=0.06Pa). The three metrics increased by 0.09m/s, 0.06, and 0.11Pa on maximum respectively, equivalent to 20.00%, 19.35%, and 183.33%. The opening distribution probability map is generated by overlapping all Pareto solutions together. The higher the opening probability, the better ventilation performance can be achieved. 5 solutions with good performance from the Pareto optimal solution of S1 are demonstrated. A coordinate axis is applied for marking the position of each module, that the closest one to the origin point is X=1, Z=1. The openings are mainly in the position of X=0 and X=3 to 7 with Z=2 to 5

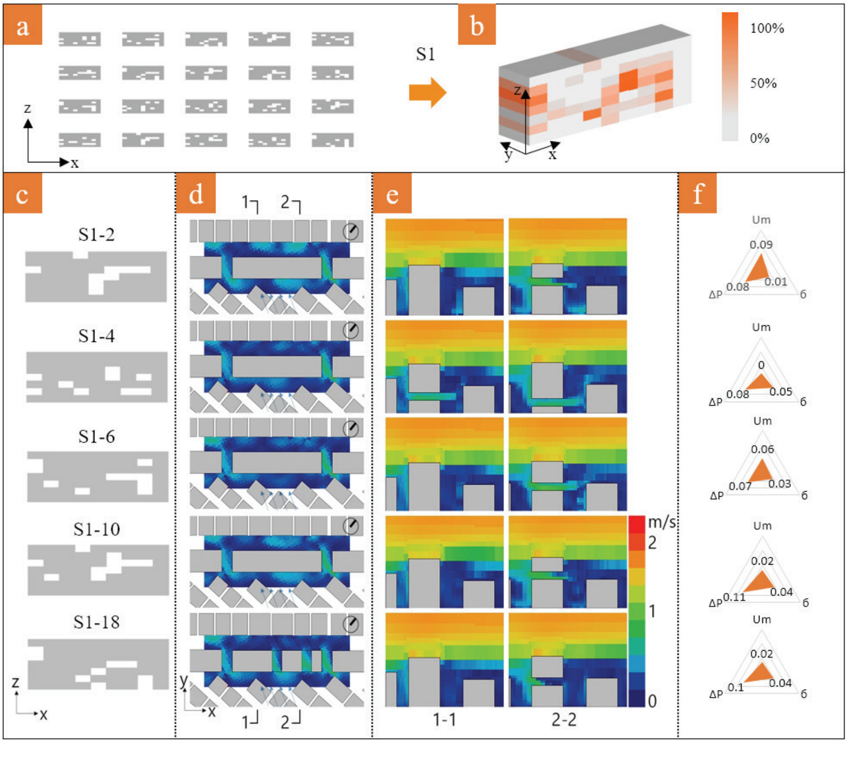


Figure 5. Result of S1; a: the opening configurations of Pareto solutions; b: the opening distribution probability map; c: the selected opening configurations; d: the pedestrian-level wind speed map; e: the profile wind speed map; f: the radar map (optimization value compared to original building without openings)

3.2. RESULT FOR S2

S2 (porosity of 30%) obtained 24 Pareto optimal solutions (Figure 6), and its metrics range is Um: 0.50 to 0.64m/s, σ: 0.26 to 0.33, ΔP: 0.09 to 0.35Pa. Three metrics increased by 0.19m/s, 0.05, and 0.29Pa on maximum, equivalent to 42.22%, 16.13%, and 483.33%, compared to the simulation results of the original building without openings. Again the opening distribution probability map and select 5 good performance solutions are shown in Figure 6. The openings are mainly in the range of Z=2 to 5.

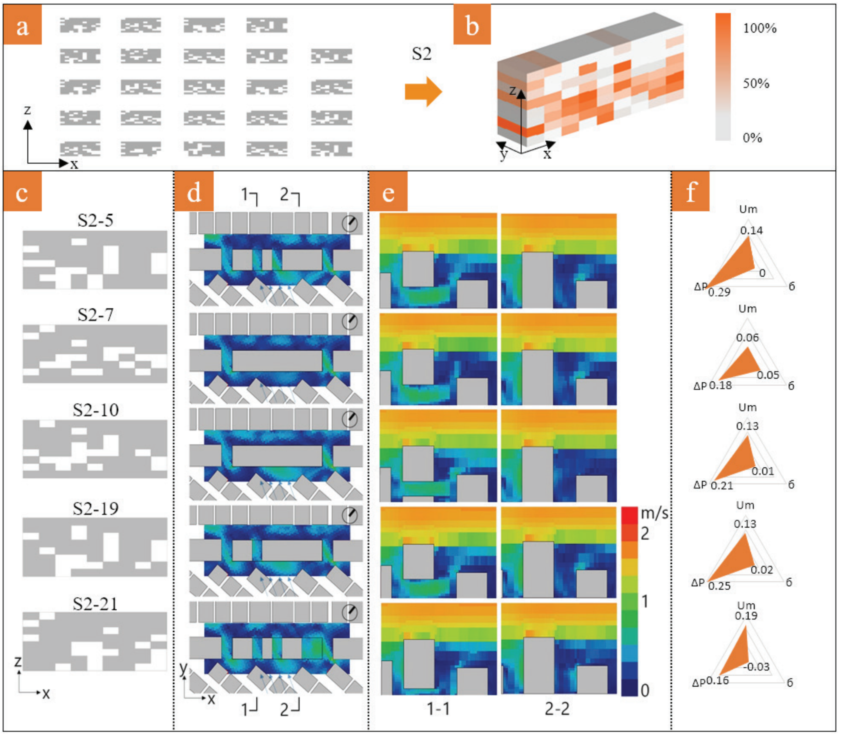


Figure 6. Result of S2; a: the opening configurations of Pareto solutions; b: the opening distribution probability map; c: the selected opening configurations; d: the pedestrian-level wind speed map; e: the profile wind speed map; f: the radar map (optimization value compared to original building without openings)

4. DISCUSSION

This study proposes a GA-based optimization workflow for the opening configurations of building geometry design by integrating CFD simulation and parametric design. A residential block in Shenzhen, China is selected as a case to show this workflow.

According to the results, the opening distributions of S1 and S2 are similar, and the openings are mainly distributed on Z=2 to 5. The potential reason is that the optimal distribution relates to the height of the downstream building. The openings on Z=2 to 5 can make the wind blow to the windward side of the downstream building, increasing its surface air pressure. And the downstream buildings block the airflow, then most of the airflow sinks by gravity. It leads to an increase in the mean wind speed at the pedestrian level. Nevertheless, the results of S1 and S2 also have some differences. The most significant difference is that the openings in S1 are less distributed within the range of X=2 to 4. There may be two reasons for this, one is that the number of individuals in the GA is not large enough, leading to local convergence of the results; the other is that different numbers of openings correspond to different optimal configurations.

This workflow provides opening distribution probability maps for early-stage architectural design. Both the opening sizes and positions with the best ventilation performance are visualized clearly. Following the instructions from these maps, architects can adjust the building geometry design according to specific requirements and a comfortable wind environment can be ensured. Additionally, the proposed method can be a reference for other building geometry parameter optimization problems to improve the built environment.

This study has some limitations. First, CFD simulation for each case is still too expensive in terms of computation resources; Second, only one season is considered in this study, so the optimization results may not be applicable to other seasons; Third, the three metrics adopted in this study are not comprehensive enough to describe the urban ventilation performance; Forth, the positions of openings in this study are completely random. In fact, some positions such as vertical transportation core are not suitable for openings.

Our future research will focus on overcoming these limitations. Firstly, Fast Fluid Dynamics Simulation (FFD) on the GPU can be a solution to reduce the simulation time; Secondly, the situation of different seasons will be considered as well, and each solution will be simulated and evaluated for multiple seasons to improve the applicability of this workflow; Then, the evaluation of air pollutants will be added, such as air change rate and age of air, to evaluate urban ventilation performance more comprehensively; Lastly, for the generation of opening configurations, the priority of opening positions can be given before optimization to meet different functional requirements.

Acknowledgments

This work was jointly supported by the Hong Kong Scholars Program, Guangdong Basic and Applied Basic Research Foundation [No. 2019A1515110577 and No. 2022A1515010412], and Fellowship of China Postdoctoral Science Foundation [No. 2020M672633].

References

Baratian-Ghorghi, Z., & Kaye, N. B. (2013). The effect of canyon aspect ratio on flushing of dense pollutants from an isolated street canyon. *Sci Total Environ,* 443, 112-122. https://doi.org/10.1016/j.scitotenv.2012.10.064

Buccolieri, R., Sandberg, M., & Di Sabatino, S. (2010). City breathability and its link to pollutant concentration distribution within urban-like geometries. *Atmospheric Environment,* 44(15), 1894-1903. https://doi.org/10.1016/j.atmosenv.2010.02.022

Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. A. M. T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE transactions on evolutionary computation,* 6(2), 182-197.

Di Sabatino, S., Buccolieri, R., Pulvirenti, B., & Britter, R. (2007). Simulations of pollutant dispersion within idealised urban-type geometries with CFD and integral models. *Atmospheric Environment,* 41(37), 8316-8329. https://doi.org/10.1016/j.atmosenv.2007.06.052

Fan, M., Chau, C. K., Chan, E. H. W., & Jia, J. (2017). A decision support tool for evaluating the air quality and wind comfort induced by different opening configurations for buildings in canyons. *Sci Total Environ,* 574, 569-582. https://doi.org/10.1016/j.scitotenv.2016.09.083

Hang, J., Li, Y., Sandberg, M., Buccolieri, R., & Di Sabatino, S. (2012). The influence of building height variability on pollutant dispersion and pedestrian ventilation in idealized high-rise urban areas. *Building and Environment,* 56, 346-360. https://doi.org/10.1016/j.buildenv.2012.03.023

Kaseb, Z., Hafezi, M., Tahbaz, M., & Delfani, S. (2020). A framework for pedestrian-level wind conditions improvement in urban areas: CFD simulation and optimization. *Building and Environment,* 184. https://doi.org/10.1016/j.buildenv.2020.107191

Lee, J. H. (2007). Optimization of indoor climate conditioning with passive and active methods using GA and CFD. *Building and Environment,* 42(9), 3333-3340. https://doi.org/10.1016/j.buildenv.2006.08.029

Makki, M., Showkatbakhsh, M., & Song, Y. (2021). Homepage. *Wallacei: Evolutionary Engine for Grasshopper3D.* Retrieved November 24, 2022, from: https://www.wallacei.com/

OpenFOAM-v1912. (2019). Homepage. *Openfoam.* Retrieved November 24, 2022, from: https://www.openfoam.com/

Sadeghipour Roudsari, M., & Mackey, C. (2019). Tools and ladybug. *Ladybug Tools.* Retrieved November 24, 2022, from https://www.ladybug.tools/

Sadeghipour Roudsari, M., & Mackey, C. (2019). Tools and butterfly. *Ladybug Tools.* Retrieved November 24, 2022, from https://www.ladybug.tools/

van Druenen, T., van Hooff, T., Montazeri, H., & Blocken, B. (2019). CFD evaluation of building geometry modifications to reduce pedestrian-level wind speed. *Building and Environment,* 163. https://doi.org/10.1016/j.buildenv.2019.106293

Yuan, C., & Ng, E. (2012). Building porosity for better urban ventilation in high-density cities - A computational parametric study. *Build Environ,* 50, 176-189. https://doi.org/10.1016/j.buildenv.2011.10.023

Yuan, C., Ng, E., & Norford, L. K. (2014). Improving air quality in high-density cities by understanding the relationship between air pollutant dispersion and urban morphologies. *Build Environ,* 71, 245-258. https://doi.org/10.1016/j.buildenv.2013.10.008

AN INVERSE MODELING METHOD TO ESTIMATE UNCERTAIN SPATIAL CONFIGURATIONS FROM 2D INFORMATION AND TIME-BASED VISUAL DISCRIMINATIONS

PIERRE CUTELLIC[1]
[1]*ETH Zürich.*
[1]*cutellic@arch.ethz.ch, ORCID: 0000-0002-7224-9222*

Abstract. This paper focuses on a specific aspect of human visual discrimination from computationally generated solutions for CAAD ends. The bottleneck at work here concern informational ratios of discriminative rates over generative ones. The amount of information that can be brought to a particular sensory modality for human perception is subject to bandwidth and dimensional limitations. This problem is well known in Brain-Computer Interfaces, where the flow of relevant information must be maintained through such interaction for applicative ends and adoption of use in many fields of human activity. While architectural modeling conveys a high level of complexity in its processes, let alone in the presentation of its generated design solutions, promises in applicative potentials of such interfaces must be made aware of these fundamental issues and need developments of appropriate sophistication. This paper addresses this informational bottleneck by introducing a method to retrieve spatial information from the rapid serial visual presentation of generated pictures. This method will be explained and defined as inverse modeling, based on inverse graphics, and its relation to human visual processing.

Keywords. Neurodesign, Design Computing and Cognition, Brain-Computer Interfaces, Generative Design, Computer Vision.

1. Introduction

Virtuous coupling of both human and machine intelligence in guiding the diversity of sought performances in AEC technologies supports the right approach for more desirable outcomes, fluency, and literacy in both control and communication. However, numerous questions and bottlenecks exist at a general level of this hybrid model. This paper focuses on a specific aspect of human visual discrimination from computationally generated solutions. The bottleneck at work here concern informational ratios of discriminative rates over generative ones. The amount of information that can be brought to a particular sensory modality for human perception is subject to bandwidth and dimensional limitations (Wickens, 1974; Venturino and Eggemeier, 1987; Klingberg, 2000; Marois and Ivanoff, 2005; Riener, 2017)—exceeding these rates cause well-known fatigue and disengagement. They can be

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 261-270.

measured in perceived mental efforts at various levels from physiological and behavioral signals as a cognitive load (Sweller and al., 1998; Xie and Salvendy, 2000; Paas and al., 2003). This might lead to counterproductive effects when the speed and quantity of solutions produced by a computer cannot be processed at a similar pace to their generation through human-computer interaction. This particular problem is well known in the field of Brain-Computer Interfaces (BCI), where the flow of relevant information must be maintained through such interaction for applicative ends and adoption of use in many areas of human activity (Lotte and Jeunet, 2018; Perkidis and Millan, 2020; Gramann and al., 2021). While architectural modeling conveys a high level of complexity in its processes, let alone in the presentation of its generated design solutions, promises in applicative potentials of such interfaces cannot escape from these fundamental issues and need developments of appropriate sophistication. Based on previous research (Cutellic, 2019, 2021) and the case study of an ongoing project tailoring BCI methods for architectural modeling at early design phases, this paper will address this informational bottleneck by introducing a technique to retrieve spatial information from the Rapid Serial Visual Presentation (RSVP) of generated pictures (Spence and Witkowski, 2013; Lees et al., 2018). The visual task at hand consists of focusing on sequences of generated pictures and discriminating visually those which seem relevant to the concept of a “ROOM” (or loosely describing an enclosing space). The visual discrimination is then further correlated with peculiar neural phenomena of study within the computational loop of a BCI (Figure 1).

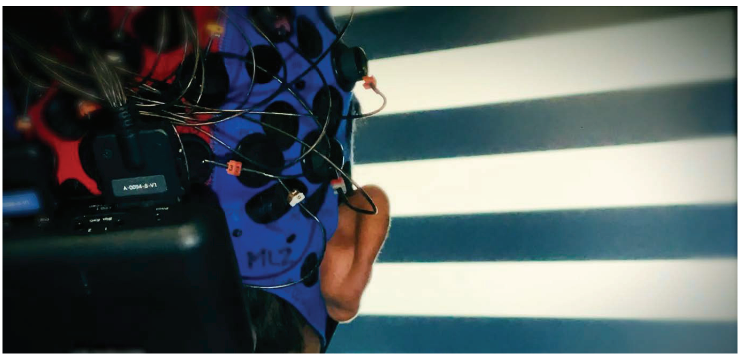


Figure 1. Photo of the conducted BCI experiment and its designed RSVP task for visual discrimination of spatial configurations.

These pictures are necessarily framed by three constraints defined by minimal spatial cues: the monocular viewpoint to which they are presented, the aforementioned informational transfer rates, and maintaining a high degree of uncertainty to increase variance in the results of such interaction. The increased degree of uncertainty comes with increased intractability in modeling spatially complex 3D aggregates from 2D pictures containing low-level features. This method will be explained and defined as an inverse modeling method based on inverse graphics and their relation to human visual processing. Such a model of reference is usually called Vision-as-Inverse-Graphics (Kulkarni and Whitney, 2015) or scene Analysis-By-Synthesis (Grenander, 1976, 1978), where prior geometric information generally serves as ground truth.

2. Background

One of the most basic and well-known feature extraction processes performed by the early visual system, and with computational equivalence, must deal with the segmentation of textures to extract shape information (Gibson, 1950; Marr et al., 1982). The so-called Shape-From-Texture problem in vision research focuses on modeling the recovery of 3D information of object surfaces from distorted texture segments which have been projected with specific angles onto a 2D Picture Plane, such as the retinal one (Gibson, 1950). This distortion, or Gradient, is generally measured by assuming the pattern distribution across the texture segment, whether a non-distorted reference is provided. Without reference, stochastic methods must be applied (Kanatani and Tsai-Chan, 1989; Clerc and Mallat, 2002). In the case of low-level and abstract geometry processing, such as for planar surfaces, distortions can be assumed to be linearly distributed and homogeneous. This assumption becomes even more practical when projections are considered within a Euclidean visual space because of the linear kind of its transformations (Erkelens, 2017), and deterministic methods can then be applied. From a perceptual point of view, perspective projection models are considered the most ideal, convenient, and realistic compared to transformations applied in physical and visual ones, which may depict both distance and size estimation in monocular vision with invariance (Erkelens, 2017). When a texture contains a dominant oriented structure, perspective convergence can be defined, and a mathematical relationship can be established between that 2D convergence in the texture and the parametric rotations of a planar surface in 3D by exploiting local symmetries (Saunders and Knill, 2001). Such convergence in texture gradients has been shown to represent further effective perceptual spatial cues for the extended generalization of ruled surfaces (Andersen, 1998; Gilliam, 1968; Todd et al., 2005). Given the abstract character of the visual objects to be detected in the designed experiment (a room), sets of abstract and basic geometry, such as planar surfaces, represent the most practical approach to describing 3D spatial segments. Using the perspective projection model, their projected distortion can be assumed uniform and reflected as homogeneous pattern distributions. Since only simple linear deformations are supposed to be present in the generated visual stimuli, deterministic numerical estimations of inverse monocular perspective projections of the deformed planar surfaces will be used.

3. Methods

It is assumed that given a picture domain, providing a projection plane, and a singular viewpoint, given by the observer position, all kinds of randomly positioned and rotated planar surfaces defining a spatial enclosure within a field of view would produce straight intersections that delimit the viewed portions of these planes, and that such projection on the picture domain would map to the geometric properties of Voronoi space partition diagrams (Voronoi, 1909; Aurenhammer and Klein, 2000). Accordingly, a generative 2D domain is set to provide for the subsequent multiple texture segments that compose the picture (Figure 2). To deduce the segments from the diagram, all diagram cells must be closed, and their centroids must be randomly generated within an inset of the picture domain and an offset of the previously generated points for perceptually visible boundaries within the picture domain.

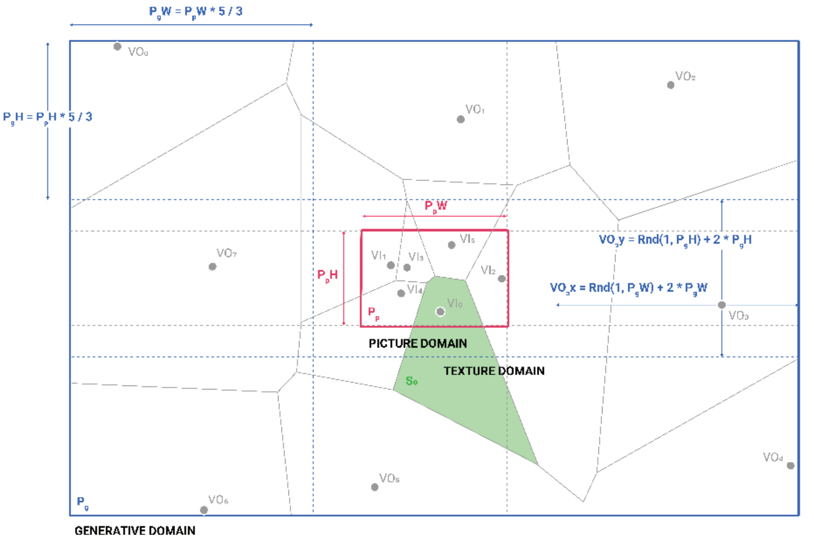
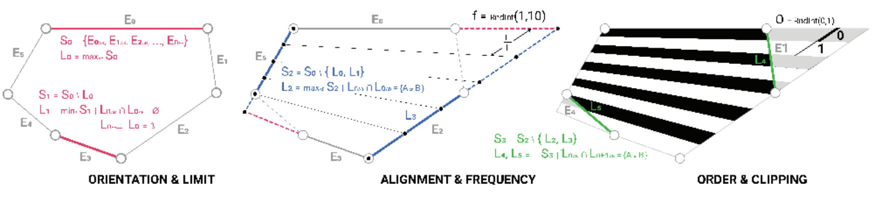


Figure 2. Relative generation of texture domains from a given 2D picture domain (Pp in red). A generative domain is set (Pg in blue) with a ratio to Pp width and height (PpW, PpH). Points are randomly placed within Pp (VIn) together with 8 points outside in their respective subdomains of Pg (VO0-7) to produce a closed Voronoi diagram containing all VIn. The resulting diagram provides for texture domains at each of its cells (S0 in green).



A most minimal texture gradient in the form of alternating black and white bands of equal dimensions (i.e., Square-Wave grating) takes advantage of the convex nature of the produced Voronoi cells and their bounding trapezoids (Figure 3). To generate the frequency and order of the bands, random integer parameters are used, and their orientation is predefined by each region's boundary.

Given the same sets of geometrical elements, the same texture segments must be generated. A generator encodes the procedural generation of these pictures, and each generative information is stored in a dataset for retrieval and further modeling. Eventually, a token may be the seed for generating new texture segments with additional random points updating the initial Voronoi diagram along the iterative sequence. A simple domain union is made between the new and existing texture

Figure 3. Selection of the longest edge (L0) and its opposing most-parallel edge (L1) to define orientation and limit of the texture. Middle: Selection of the longest connected edge (L2) and its opposed (L3) for each of the two previously selected to divide them by a generated frequency parameter (f) and define the alignment of the bands. Right, random order of the bands (O) by starting as 0-1 or 1-0. The remaining edges (L4, L5) form connected sets to clip the generated bands within the texture domain.

Figure 4. Two samples of generated images with texture segments through an iterative sequence. Left: the image generated after 4 iterations contains 4 texture segments. Right: the same image iterated over after 9 iterations. Highlighted in green is the same texture segment being updated over iterations.

segments. That way, previously generated texture segments are preserved, and only their clipping boundaries may be updated regardless of whether the edges initially used for defining the texture have been altered (Figure 4).

To recall the initial problem of inverse modeling from a monocular viewpoint (Palmer, 1999; Zygmunt, 2001): for a given surface projected on a 2D plane, there is an infinite number of possible configurations in the 3D field of view. Texture gradients provide for that missing information to perceive an uncertain but most probable surface orientation and position from a distance (Figure 6). As previously introduced, the geometric information provided by a texture segment on the picture plane, which also serves as perceptual information, is a function of the texture distribution. Segment-wise, local symmetry can be exploited to extract the slant and tilt angles. Two vanishing points may be retrieved from the longest edges of the trapezoid orienting the pattern generation, and their opposing ones generating its distribution (i.e., respectively L0, L1 and L2, L3 in Figure 3). Their bisector intersects at the center of transformations to be applied on the planar surface; the origin of the surface normal vector projected on the picture plane. Following (Ribeiro and Hancock, 1999), a linear approximation of the local distortion due to perspective can be derived. The projected normal vector components (p, q) can be retrieved by measuring the angle between the bisectors and the XY axes of the picture plane. From there, slant and tilt angles can be found. To serve as ground truth, a reference measure is used as the biggest possibly perceived pattern from a picture (half of the diagonal of the presentation screen). That reference is put in the Thales reciprocity formula with the measure of the distorted pattern along the bisector produced by the trapezoid edges generating the pattern distribution. The resulting distance is used for translating the rotated surface along the line of sight relatively to each other texture segments.

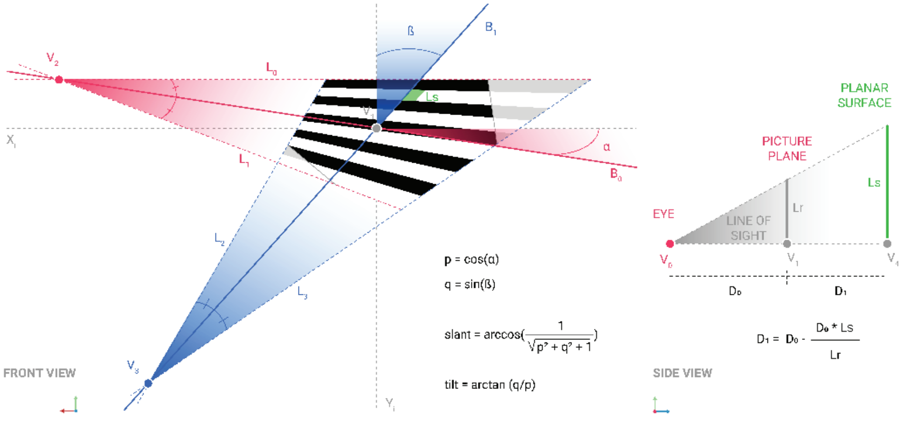


Figure 6. Left, front view, in the picture plane. The point of transformation V1 is retrieved by intersecting the bisectors B0 and B1. B0 and B1 pass respectively by the vanishing points V2 and V3 retrieved from the trapezoid edges generating the texture (L0, L1 and L2, L3). The linear equation that approximates the slant and tilt angles can be retrieved by measuring the angles alpha and beta between the bisectors and the axes of the picture plane. A distance measure of the texture Ls is taken along B1. - Right, side view, in the plane of the line of sight passing from the viewpoint origin V0 and V1. The distance D1 along this line to position the plane at V4 is deduced from the reference measure Lr by reciprocity.

4. Results

For reproducibility of the model under variance, random samples have been taken in the dataset of generated pictures (Figure 7). For every transformed sample, exactly similar transformations have been found for identical texture segments with different boundaries and were found to share the same plane in the inverse 3D space. The transformations were equally successfully applied to every other configuration presented in other generated sequences and matched perceptual evaluation of their orientation and relative distance as per the developed method. As expected, plane intersections vary greatly across generated samples. Since information regarding boundary conditions cannot be retrieved from the pictures, or if so, could be found in conflict with the present information regarding planar transformations, the former is considered to result from the latter. The retrieved boundary conditions become either the edge limiting a planar surface occluding geometries on the background and leaving a gap in-between, or planar intersections along a line or a point. From these resulting conditions, procedural geometric operations have been deduced and systematically applied to generate a CAD model and fabricate a physically tangible mock-up.

5. Conclusion and Outlook

This paper presents a method to design an inverse geometric model of 3D planar geometries from 2D perceptual information under a certain degree of uncertainty and for the rapid presentation of early design scenarios. Its low level of 2D information and presentation pace aims to develop applicative methods in BCI for design and architectural modeling that support the increase of fluency between humans and computers. The core of such an approach lies in two basic assumptions of productivity driven by cumulative empirical findings from cognitive science literature: the generative productivity of computers to output solutions from large dimensional spaces and the discriminative productivity of humans to output responses of higher dimensions from reduced ones. Intractable problems due to a lack of sufficient and timely information may be approached heuristically and iteratively. Several assumptions have been made during the development of this method and kept here for discussion because their relevance mainly concerns the degree of generalization for further research. The basic assumption which allows for correlating a geometric inverse model with perceptual cues examines the hypothesis that regardless of the complexity of visual inputs (i.e., natural images), the principle of inverse geometric modeling operates naturally at the level of human perceptual segmentation. In addition, monocular models are known to leave out a significant amount of uncertainty. For example, a perceived slant angle may vary from two different instances of the same texture segment (Blake, Bülthoff, & Sheinberg, 1993; Cutting & Millard, 1984; Knill, 1998a). The estimation error in perceived angles may also vary by angle thresholds and pattern differences. Additionally, the statistical nature of such information is related to the presentation context (e.g., the overall composition of all texture segments, the modality from person to person, …). Because RSVP integrates by design the repetitive presentation of similar instances on an individual basis, and parameters of uncertainty in the perception of texture gradient are continuously accumulated in the literature, a certain degree of confidence could be further developed for these 3D transformations. Regarding the remaining uncertainty parameters regarding relative distance estimations and surface occlusions, further investigation in providing additional viewpoints could give the method more accuracy. Optimal cue combination methods can support this development, and the deployment of such a process could be thought either stationary (presentation of multiple viewpoints of the same scene on a static screen) or mobile (in the context of augmented reality devices). Finally, generated texture gradients should account for visual discomfort under prolonged presentation modalities such as RSVP. Artificially generated distributions could mitigate this effect by eventually accommodating noise found in natural images. The generation of such noise-based gradients would involve applying generalized methods, including stochastically distributed patterns and non-planar surface geometries.

Acknowledgments

This research is part of the Neuramod project (105213_192500), fully funded by the Swiss National Science Foundation SNSF Project Funding Div. 1-2 for Humanities, Social sciences, Mathematics, Natural, and Engineering sciences.

References

Andersen, G. J., Braunstein, M. L., & Saidpour, A. (1998). *The perception of depth and slant from texture in three-dimensional scenes*. Perception, 27(9), 1087–1106. https://doi.org/10.1068/p271087

Aurenhammer, F., & Klein, R. (2000). *Voronoi Diagrams*. Handbook of Computational Geometry, 5(10), 201–290.

Blake, A., Bülthoff, H. H., & Sheinberg, D. (1993). *Shape from texture: Ideal observers and human psychophysics.* Vision Research, 33(12), 1723–1737. https://doi.org/10.1016/0042-6989(93)90037-w

BuHamdan, S., Alwisy, A., & Bouferguene, A. (2021). *Generative systems in the architecture, engineering, and construction industry: A systematic review and analysis*. International Journal of Architectural Computing, 19(3), 226–249. https://doi.org/10.1177/1478077120934126

Clerc, M., & Mallat, S. (2002). *The texture gradient equation for recovering shape from texture*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(4), 536–549. https://doi.org/10.1109/34.993560

Cutellic, P. (2019). *Towards encoding shape features with visual event-related potential based brain–computer interface for generative design*. International Journal of Architectural Computing, 17(1), 88–102. https://doi.org/10.1177/1478077119832465

Cutellic, P. (2021). *Growing Shapes with a Generalised Model from Neural Correlates of Visual Discrimination*. In P. F. Yuan, J. Yao, C. Yan, X. Wang, & N. Leach (Eds.), Proceedings of the 2020 DigitalFUTURES (pp. 68–78). Springer. https://doi.org/10.1007/978-981-33-4400-6_7

Cutting, J. E., & Millard, R. T. (1984). *Three gradients and the perception of flat and curved surfaces*. Journal of Experimental Psychology: General, 113(2), 198–216. https://doi.org/10.1037/0096-3445.113.2.198

Darko, A., Chan, A. P. C., Adabre, M. A., Edwards, D. J., Hosseini, M. R., & Ameyaw, E. E. (2020). *Artificial intelligence in the AEC industry: Scientometric analysis and visualization of research activities.* Automation in Construction, 112, 103081. https://doi.org/10.1016/j.autcon.2020.103081

Emaminejad, N., North, A. M., & Akhavian, R. (2022). *Trust in AI and Implications for the AEC Research: A Literature Analysis*. ArXiv:2203.03847 [Cs]. http://arxiv.org/abs/2203.03847

Erkelens, C. J. (2017). *Perspective Space as a Model for Distance and Size Perception*. I-Perception, 8(6), 2041669517735541. https://doi.org/10.1177/2041669517735541

Gibson, J. J. (1950). *The perception of the visual world* (pp. xii, 242). Houghton Mifflin.

Gillam, B. J. (1968). *Perception of slant when perspective and stereopsis conflict: Experiments with aniseikonic lenses*. Journal of Experimental Psychology, 78(2), 299–305. https://doi.org/10.1037/h0026271

Gramann, K., McKendrick, R., Baldwin, C., Roy, R. N., Jeunet, C., Mehta, R. K., & Vecchiato, G. (2021). *Grand Field Challenges for Cognitive Neuroergonomics in the Coming Decade*. Frontiers in Neuroergonomics, 2. https://doi.org/10.3389/fnrgo.2021.643969

Grenander, U. (1976). *Pattern Synthesis: Lectures in Pattern Theory*. Springer-Verlag. https://doi.org/10.1007/978-1-4612-6369-2

Grenander, U. (1978). *Pattern Analysis: Lectures in Pattern Theory II* (Softcover reprint of the original 1st ed. 1978 edition). Springer.

Kanatani, K., & Chou, T.-C. (1989). *Shape from texture: General principle*. Artificial Intelligence, 38(1), 1–48. https://doi.org/10.1016/0004-3702(89)90066-0

Klingberg, T. (2000). *Limitations in information processing in the human brain: Neuroimaging of dual task performance and working memory tasks*. In Progress in Brain Research (Vol. 126, pp. 95–102). Elsevier. https://doi.org/10.1016/S0079-6123(00)26009-3

Knill, D. C. (1998). *Discrimination of planar surface slant from texture: Human and ideal observers compared.* Vision Research, 38(11), 1683–1711. https://doi.org/10.1016/S0042-6989(97)00325-8

Koch, E., Baig, F., & Zaidi, Q. (2018). *Picture perception reveals mental geometry of 3D scene inferences*. Proceedings of the National Academy of Sciences, 115(30), 7807–7812. https://doi.org/10.1073/pnas.1804873115

Korteling, J. E. (Hans)., van de Boer-Visschedijk, G. C., Blankendaal, R. A. M., Boonekamp, R. C., & Eikelboom, A. R. (2021). *Human- versus Artificial Intelligence*. Frontiers in Artificial Intelligence, 4. https://doi.org/10.3389/frai.2021.622364

Kulkarni, T. D., Whitney, W., Kohli, P., & Tenenbaum, J. B. (2015). *Deep Convolutional Inverse Graphics Network*. ArXiv:1503.03167 [Cs]. http://arxiv.org/abs/1503.03167

Lees, S., Dayan, N., Cecotti, H., McCullagh, P., Maguire, L., Lotte, F., & Coyle, D. (2018). *A review of rapid serial visual presentation-based brain–computer interfaces*. Journal of Neural Engineering, 15(2), 021001. https://doi.org/10.1088/1741-2552/aa9817

Lotte, F., & Jeunet, C. (2018). *Defining and quantifying users' mental imagery-based BCI skills: A first step*. Journal of Neural Engineering, 15(4), 046030. https://doi.org/10.1088/1741-2552/aac577

Manzoor, B., Othman, I., & Pomares, J. C. (2021). *Digital Technologies in the Architecture, Engineering and Construction (AEC) Industry—A Bibliometric—Qualitative Literature Review of Research Activities*. International Journal of Environmental Research and Public Health, 18(11), Article 11. https://doi.org/10.3390/ijerph18116135

Marois, R., & Ivanoff, J. (2005). *Capacity limits of information processing in the brain*. Trends in Cognitive Sciences, 9(6), 296–305. https://doi.org/10.1016/j.tics.2005.04.010

Marr, D., Poggio, T. A., & Ullman, S. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. The MIT Press.

Maruya, A., & Zaidi, Q. (2020). *Mental geometry of perceiving 3D size in pictures*. Journal of Vision, 20(10), 4. https://doi.org/10.1167/jov.20.10.4

Mohammadpour, A., Karan, E., & Asadi, S. (2019). Artificial Intelligence Techniques to Support Design and Construction. ISARC Proceedings, 1282–1289. https://doi.org/10.22260/ISARC2019/0172

O'Hare, L., & Hibbard, P. B. (2011). *Spatial frequency and visual discomfort*. Vision Research, 51(15), 1767–1777. https://doi.org/10.1016/j.visres.2011.06.002

Oliver, S. (2019). *Communication and trust: Rethinking the way construction industry professionals and software vendors utilise computer communication mediums*. Visualization in Engineering, 7(1), 1. https://doi.org/10.1186/s40327-019-0068-y

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. M. (2003). *Cognitive Load Measurement as a Means to Advance Cognitive Load Theory.* Educational Psychologist, 38(1), 63–71. https://doi.org/10.1207/S15326985EP3801_8

Palmer, S. E. (1999). *Vision science: Photons to phenomenology (3rd printing).* MIT Press.

Pan, Y., & Zhang, L. (2021). *Roles of artificial intelligence in construction engineering and management: A critical review and future trends*. Automation in Construction, 122, 103517. https://doi.org/10.1016/j.autcon.2020.103517

Perdikis, S., & Millan, J. del R. (2020). *Brain-Machine Interfaces: A Tale of Two Learners*. IEEE Systems, Man, and Cybernetics Magazine, 6(3), 12–19. https://doi.org/10.1109/MSMC.2019.2958200

Pizlo, Z. (2001). *Perception is viewed as an inverse problem*. Vision Research, 41(24), 3145–3161. https://doi.org/10.1016/S0042-6989(01)00173-0

Ribeiro, E., & Hancock, E. R. (1999). *Improved orientation estimation for texture planes using multiple vanishing points*.

Riener, A. (2017). *Chapter 19—Subliminal Perception or "Can We Perceive and Be Influenced by Stimuli That Do Not Reach Us on a Conscious Level?"* In M. Jeon (Ed.), Emotions and Affect in Human Factors and Human-Computer Interaction (pp. 503–538). Academic Press. https://doi.org/10.1016/B978-0-12-801851-4.00019-7

Saunders, J. A., & Knill, D. C. (2001). *Perception of 3D surface orientation from skew symmetry.* Vision Research, 41(24), 3163–3183. https://doi.org/10.1016/S0042-6989(01)00187-0

Spence, R., & Witkowski, M. (2013*). Rapid Serial Visual Presentation: Design for Cognition* (2013th edition). Springer.

Stevens, K. A. (1983). *Surface tilt (the direction of slant): A neglected psychophysical variable.* Perception & Psychophysics, 33(3), 241–250. https://doi.org/10.3758/BF03202860

Sun, W., Mollaoglu, S., Miller, V., & Manata, B. (2015). *Communication Behaviors to Implement Innovations: How Do AEC Teams Communicate in IPD Projects?* Project Management Journal, 46(1), 84–96. https://doi.org/10.1002/pmj.21478

Sweller, J., van Merrienboer, J. J. G., & Paas, F. G. W. C. (1998). *Cognitive Architecture and Instructional Design.* Educational Psychology Review, 10(3), 251–296. https://doi.org/10.1023/A:1022193728205

Todd, J. T., Thaler, L., & Dijkstra, T. M. H. (2005). *The effects of field of view on the perception of 3D slant from texture.* Vision Research, 45(12), 1501–1517. https://doi.org/10.1016/j.visres.2005.01.003

Tomasi, C. (2006). *Vision, Early*. In Encyclopedia of Cognitive Science. John Wiley & Sons, Ltd. https://doi.org/10.1002/0470018860.s00071

Venturino, M., & Eggemeier, F. T. (1987). *Capacity Limitations in Human Information Processing: Theory and Applications.* Proceedings of the Human Factors Society Annual Meeting, 31(6), 672–673. https://doi.org/10.1177/154193128703100613

Voronoi, G. (1909). *Nouvelles applications des paramètres continus à théorie des formes quadratiques. Deuxième Mémoire. Recherches sur les paralléloèdres primitifs*. Journal Für Die Reine Und Angewandte Mathematik, 136, 67–182.

Wickens, C. D. (1974). *Temporal limits of human information processing: A developmental study.* Psychological Bulletin, 81(11), 739. https://doi.org/10.1037/h0037250

Xie, B., & Salvendy, G. (2000). *Review and reappraisal of modelling and predicting mental workload in single- and multi-task environments*. Work & Stress, 14(1), 74–99. https://doi.org/10.1080/026783700417249

RE-WEATHERING

A Nature-Inspired Experimental Method for Re-Generating Porous Architectural Systems Based on Environmentally Data-Driven Performance

CHENXIAO LI[1], DING WEN BAO[2,*], XIN YAN[3], RENHU WU[4] and CHENMING HE[5]
[1,4]University of Pennsylvania, [2]RMIT University, [3]Tsinghua University, [5]Cornell University
[1]licx@upenn.edu, 0000-0002-3871-8312
[2]nic.bao@rmit.edu.au, 0000-0003-1395-8747
[3]yanxin2022@tsinghua.edu.cn, 0000-0002-5033-3597
[4]wurenhu@upenn.edu, 0000-0002-3591-9290
[5]hecm216@outlook.com, 0000-0002-3871-8312
[*]corresponding author: Ding Wen Bao, nic.bao@rmit.edu.au

Abstract. Weathering scenery, as one of the most representative time-consuming natural features, is associated with many characteristics through aspects of specialized morphology, porosity, adaptability, and regenerative ability. Because of its unique porous structure and majestic, randomly three-dimensional composition, it has inspired lots of creations in industrial design, art crafts, and graphic design fields. But in architectural generative design, very limited projects are related to weathering. However, in recent decades, the advancements in computer-aided design tools have made it possible to implement rigorous computational methods in complex geometrical systems. This research tries to learn from the weathering evolution and apply the characteristics in the pursuit of enriching nature-inspired architecture. The experiments will focus on one specific weathering behavior, using wind-blown particles as weathering agents, and introduce an experimental method of generating the performance-based porosity architectural system responding to environmental agents' effects. Computational Fluid Dynamics (CFD), Bi-Directional Evolutionary Structural Optimization (BESO), and Swarm Intelligence are employed to negotiate between surface, structure, and environmental space. This study of the porosity architectural system is concerned not only with discovering new possibilities for intricate and complex nature-inspired architectural forms, but also with negotiating and considering humans and nature as equal decision-makers.

Keywords. Porosity, Topology Optimization, Bi-directional Evolutionary Structural Optimization (BESO), Swarm Intelligence, Computational Fluid Dynamics (CFD), Nature-inspired Form-Finding

1. Introduction

Weathering, which describes the breaking down or dissolving of rocks and minerals on the surface of the Earth (Tavsan & Sonmez, 2015), is the chemical and physical change in time of the surface under the influence of the atmosphere, hydrosphere, cryosphere, biosphere, and nuclear radiation (temperature, rain, circulating groundwater, vegetation, etc.) (Hack, 2019). On the microscale, the weathering evolution contains both additive and subtractive processes, but in general, it's a "loss of structure" (Hack, 2019). As rocks weather chemically, their mineral constituents change to new, more stable assemblages, and the contained elements are preserved in resistant minerals, partly redistributed into new minerals, or taken into solution (Mcqueen & Scott, 2008). Within random donations from diverse environmental agents alongside the whole process, the unusual shape, texture, and amazing mechanical balance of each natural sandstone formation has left a strong impression of intelligent design in industrial design, art crafts, and graphic design fields (Ostanin et al., 2017).

However, because of the technical limitations of simulating the dynamic environmental factors and the stereotypes associated with the definitions of "nature-inspired" and "biomimicry," weathering-inspired design is rarely mentioned in architecture. Few well-known projects, such as Aqua by Studio Gang and Mountainous Chaoyang Park Plaza by MAD Architects, were all inspired solely by the natural appearance rather than the outcome of a chaotic process.

The goal of this paper is to find the suitable computational design methods and propose an efficient experimental workflow in this overlapping field so that we can fully understand and use the properties of weathering in nature-inspired architecture.

The general setting leads us to the following questions, which will be addressed in the presented study: (1) Firstly, how is the artificial fluid condition simulated? (2) Second, how can structural optimization based on the previous pressure analysis be performed? (3) Third, is it possible to ensure architectural complexity yet simultaneously maintain structural efficiency? And finally, (4) what potential value does additive manufacturing provide material performance?

2. Literature Review

Computational fluid dynamics (CFD) is a method of discretizing the basic equation describing a flow field and numerically solving it by using a computer. It has been widely applied since the 1960s for simulation in the wind engineering field (Kataoka, Ono, & Enoki, 2020), water currents (Rezakazemi, Shahverdi, Shirazian, Mohammadi, & Pak, 2011; Montazeri, Blocken, & Hensena, 2015). Also, CFD has merited more precise and rational methods in architecture and urban planning in the past ten years, such as using CFD simulations for environmental assessment (Toja-Silva, Kono, Peralta, Lopez-Garcia, & Chen, 2018; Blocken, W.D., & Hooff, 2012; Toja-Silva, Lopez-Garcia, Peralta, Navarro, & Cruz, 2016; Toja, Peralta, Lopez-Garcia, Navarro, & Cruz, 2015), and shaping optimization (Kormaníková, Chronis, Kmeť, & Katunský, 2018; Wang, Tan, & Ji, 2016). With more in-depth research, the shapes of the wind-based model are not limited to the simulation of fixed environmental parameters, but more about how to add dynamic environmental parameters to the iterative generation process in the past five years. For instance, previous studies have reported the possibilities of designing or testing adaptive building envelope systems in a dynamic environment (Kabošová, Katunský, & Kmet, 2020), and applications in innovative, non-standard shapes design (Feng et al., 2021). There are many CFD softwares, but their basic working principles are similar. This experimental method will be based on ANSY/Fluent 2021, the same as used in Project Artificial Taihu Stone proposed by Bao's team at RMIT in 2021 (Feng et al., 2021).

Over the last few decades, structural form-finding optimization applications have gradually emerged in architectural design practice. Topology optimization allows architects to seek creative shapes without giving up structural performance. One widely applied topological optimization method, Evolutionary Structural Optimization (ESO), was first proposed in the 1990s and now has inspired a more advanced method: Bi-Directional Evolutionary Structural Optimization (BESO). With BESO, the optimization is not only able to remove material from the structure but is also able to add the necessary parts. From the perspective of numerical optimization, the benefit of this bidirectional optimization is obvious; that is, the wrongly removed material can be recovered through the growth process (Xie, Zuo, & Lv, 2014), which is the ideal working principle of the virtual weathering process. Many researchers started to try integrating the BESO method with other algorithms to increase the diversity of the outcome. For instance, SwarmBESO, proposed in 2021, combined the BESO and Multi-Agent Systems (MAS) together to generate intricate and complex geometries more efficiently and speculates their potential realisation through additive manufacturing (Bao et al., 2021). But limited systematic examples can be found to know how to make it work in weathering-inspired design.

The scheme of the developed design loop that was concluded by Kabošová, Kmet, & Katunský in 2019 is a classic one for related workflows: wind data collection, parametric definition, shape generation, CFD analysis, and final shape. In wind-based shaping design, the most recent research on the collaboration between BESO and CFD to generate an artificial Taihu Stone in 2021 (Feng et al., 2021), on which our research is also based, has indicated the possibility of a dynamic process with these two digital methods. But these two methods are not enough to realize the simulation expectations of the complete weathering process. Therefore, a specific question is how to systematically combine CFD, BESO and MAS together to generate the rough skin texture and intricate porous structure caused by particle collision and wind erosion.

This paper aims to test the feasibility of using CFD, BESO and MAS together to generate the rough skin texture and intricate porous structure caused by particle collision and wind erosion.

3. Hybrid Workflow

According to a previous study, we took the three-dimensional rock-like components of different shapes as our original model and tested our method on them in three phases, as shown in Figure 1. Firstly, Computational Fluid Dynamics (CFD) is introduced to produce quantitative predictions of fluid-flow phenomena and get the results of pressure analysis on the original model surface (Hu, 2012). Secondly, apply Bi-Directional Evolutionary Structural Optimization (BESO) techniques (Huang & Xie, 2007) to run topology optimization after CFD simulation (Feng et al., 2021). Then, a swarm algorithm based on the self-organizing behavior of birds was employed for the project to negotiate between surface, structure, and environmental space, so that the system could undertake the autonomous construction by converting multiple cantilevers into one architectural system efficiently (Snooks & Stuart-Smith, 2010). In sum, first, find a cost-efficient structural system while also meeting basic building functionality goals and then generate intricate representations.

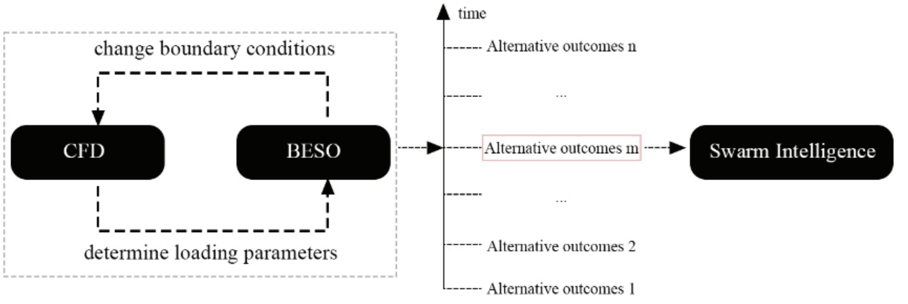


Figure 1. The diagram of the generative process

3.1. PROTOTYPE DESIGN

To put the proposed workflow for generating rock-like shapes to the test, we first create several geometrical rock-like components of various shapes in Rhino3D® (Figure 2). The massing rules follow classic architectural form and spatial order, such as a cave, a cantilever, and interlocking spaces. All components were determined to represent architectural scale and shape. For example, prototype X, depicted in Figure 2, is designed with a cuboid volume of 4 m* 5.6 m* 6 m and is expected to function as a two-story high block.

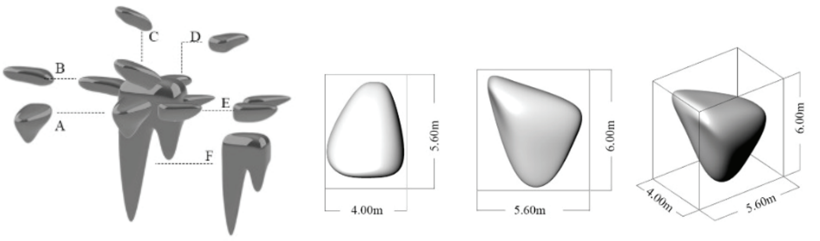


Figure 2. Prototypes: (left) numbering of each component in massing; (right) the scale of prototype A: (a) top view, (b) right view, (c) axonometric view

3.2. FLUID ANALYSIS

The fluid analysis is operated in Ansys® Academic Research Mechanical, Release 22.1. This software has quantitative predictions of fluid-flow phenomena and has good compatibility with the mechanical analysis plug-in in Grasshopper. To be more straightforward about the specific interaction during this "loss of structure" progress, the pressure from flowing wind is the main environmental factor being tested. We only focus on the erosion caused by wind. So, in the following CFD simulation, the flow direction is parallel to the negative Z axis of world coordinates, and the velocity of the environment in Rhino3D® is 11km/h. The reason for these parameters will be discussed in a later section. The fluid analysis of the parametrical component A generated in the first step is shown in Figure 3.

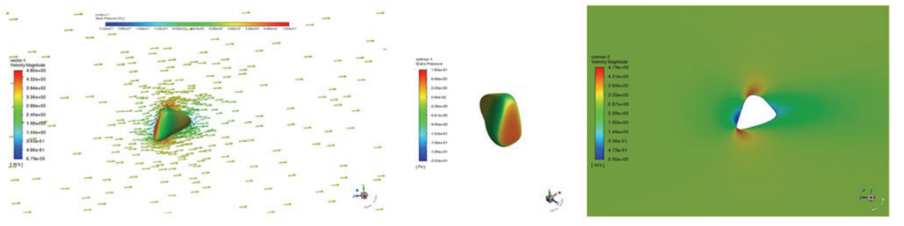


Figure 3. CFD result of prototype A: (a) static pressure shown in vector, (b) pressure in coloured contours, (c) eddy viscosity in coloured contours.

3.3. TOPOLOGY OPTIMIZATION

With the force data collected in CFD, we downloaded the database as an Excel file that is readable in Rhino3D®. This database works as the loading force in Ameba (Zhou et al., 2018), which is a topology optimization grasshopper plug-in tool based on the BESO method. As demonstrated by the existing test of *Artificial Taihu Stone*, the collaboration of CFD and BESO works well on a single component. So, apart from single component analysis, we also tested two adjacent components attached together (Figure 4). In addition to the vertical load, a series of even torques is also added in. In the load setting, the fundamental architectural components, floor and ceiling, are preserved.

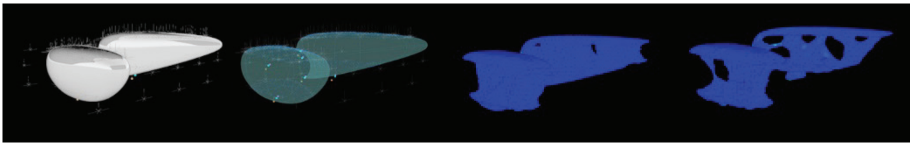


Figure 4. BESO analysis result of prototype X and Y together in Ameba (Zhou et al., 2018): (a) initial design mesh with force analysis from CFD; (b) force analysis with mesh structure curves on; (c) topological optimization in progress, (d) the result of topological optimization.

3.4. SWARMBESO

The logic for single and multiple components is the same in the later process of applying Multi-Agent Systems (MAS). In order to better display the details, we still take prototype A as an example, as shown in Figure 5.

First, we generated the optimised structural forms from Ameba (Zhou et al., 2018), and ran 200 agent points, to simulate the flocking behaviors of birds around the coloured surface. Then, we set the behaviour rules for the agents, with the cohesion and alignment working as the same as the SwarmBESO method proposed by Bao in 2021 (Bao et al., 2021). But the difference is that the swarm algorithm in the paper uses the different type of agents, which is called the continuous fibrous network. So instead of starting with multiple agents and finding their movements, we recorded the trajectories of these agents and generated ribbons from these curves by using Culebra.NET (*Culebra*, 2022). We also add the noise algorithm to improve complexity. The parameters tested are shown in Figure 6 with their corresponding results.

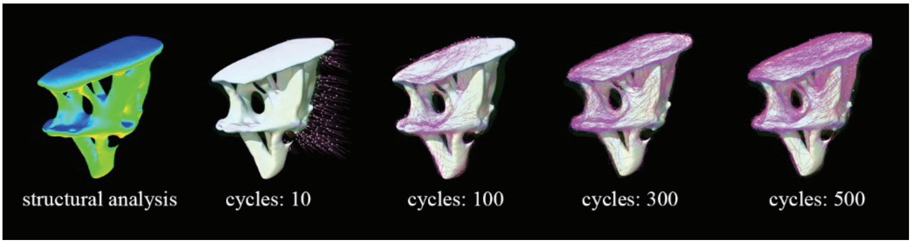


Figure 5. Structural analysis shown in coloured contours with the iterative processes of flocking.

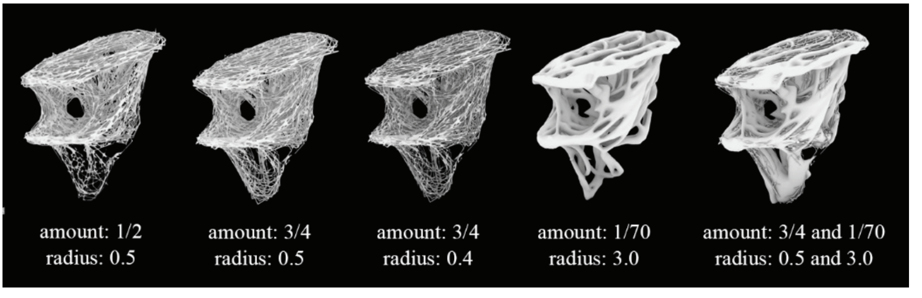


Figure 6. The flocking results of different "amount" and "radius" parameters

4. Result and Discussion

In the final design application, first we designed a series of individual models and combined them together as a reference spatial model. Then we ran, fluid analysis in CFD with a fluid velocity of 11 km/h and a fluid direction from the Z axis from top to down. Later, we input loading data into Ameba (Zhou et al., 2018) for the process of topological optimization one by one and then grouped them together for the structural analysis. The colour information is preserved with the vertex's position and affects the flocking behaviors in the final step (Figure 7).

The collaboration of these digital methods can significantly improve the efficiency of the complex geometry optimization of the weathering-inspired structure. Though this hybrid method is inspired by weathering, what we aimed at was more than just weathering simulation. With the reinforcement from SwarmBESO, we can generate more intricate alternatives more efficiently with better structural performance. The archetype of the weathering process reveals the feasibility of this combination. The complexity and mechanism both remained in the final design application (Figure 8).

This method also provides an efficient methodology for form-finding. The device remained different kinds of architectural elements, and massing rules ensure multiple architectural possibilities, which can exhibit the realities of the untamed ecology (Figure 9). Even after the topology optimization, the new cascading structure enriches the whole form and spatial environment (Figure 10).

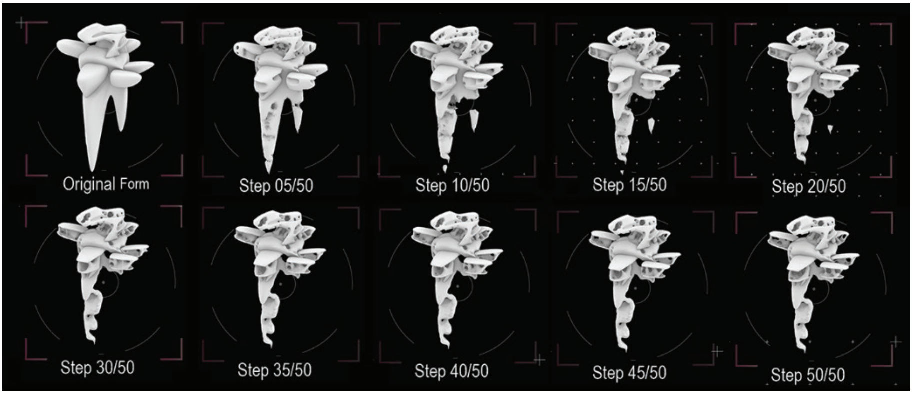


Figure 7. Iteration process of the whole prototypes from step 0 to step 50.

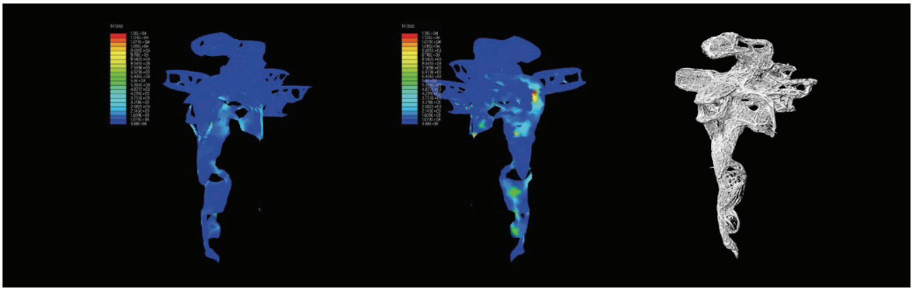


Figure 8. (left) BESO analysis result of the whole prototype shown in view 1; (middle) BESO analysis result of the whole prototype shown in view 2; (right) digital model of SwarmBESO result.

However, there still some limitations remain. To purify the loading environment, we choose the unusual fluid direction of top to bottom, working in the same direction as self-loading. And the results are surprising without making much aesthetic difference.

Besides, the idea of swarmBESO in this paper is an extension of the one that Bao proposed, and it is still at the experimental stage. No obvious difference could be easily observed with the influence of the mises analysis. Also, after we generated the intricate networks using the multi-agent algorithm, no suitable method was found to ensure the continuity of each fibrous structure. At this stage, it can only rely on human observation and manual adjustment.

Figure 9. Rendering 1

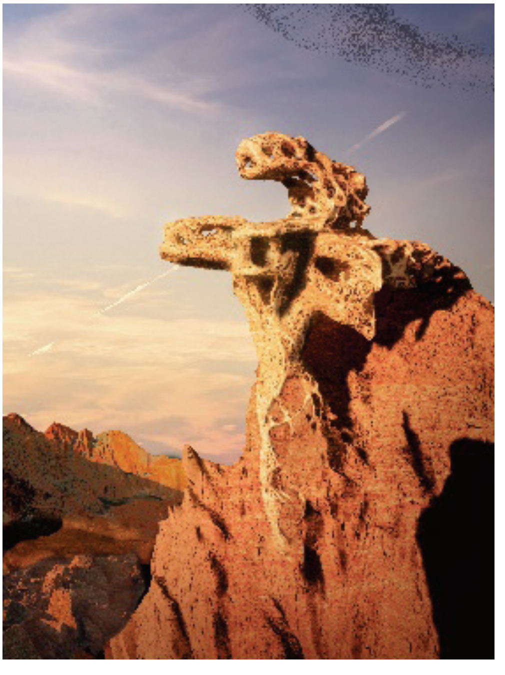


Figure 10. Rendering 2

5. Future Work

The composite fibrous skin registers the ripples of bifurcating and converging strands that blur the distinction between artificial structure and natural weathering ornament. Moreover, the final generated alternatives with features in porous shape and façade can also be applied in other generative designs for architecture, such as porosity city. But for a more comprehensive understanding of this method, more experiments are going to be tested.

In the CFD simulation, the real-time dynamic environmental data will be collected and tested to work in the ANSY software and trigger a real-time generation within the CFD-SWARMBESO workflow.

Also, more complex environmental parameters are going to be tested, including not only vertical fluids but also vertical ones such as rainfall and horizontal blowing wind working together.

Moreover, to improve the working principles of SwarmBESO, we will conduct more experiments to find a better way to ensure both complexity and continuity in the generated structures.

References

Ansys (2023) Academic Research Mechanical, Release 22.1, Help System, Coupled Field Analysis Guide, ANSYS, Inc.

Avdić, I. (2019). *Bio-Inspired Approach to Early-Stage Structural Form Finding*. TU Delft Library. Retrieved September 3, 2022, from http://resolver.tudelft.nl/uuid:ebed5ec7-7951-4139-b80f-eba8052c86c1.

Bao, D. W., Yan, X., & Xie, Y. M. (2022). Encoding topological optimisation logical structure rules into multi-agent system for architectural design and robotic fabrication. *International Journal of Architectural Computing*, 20(1), 7–17. https://doi.org/10.1177/14780771221082257

Bao, D. W., Yan, X., Snooks, R., & Xie, Y. M. (2021). Swarmbeso: Multi-agent and evolutionary computational design based on the principles of structural performance. In *26th International Conference on Computer-Aided Architectural Design Research in Asia: performative envelopes, CAADRIA 2021* (pp. 241-250). The Association for Computer-Aided Architectural Design Research in Asia (CAADRIA).

Blocken, B., W.D., J., & Hooff, T. V. (2012, April). CFD simulation for pedestrian wind comfort and wind safety in urban areas: General decision framework and case study for the Eindhoven University campus. *Environmental Modelling & Software, 30*, 15-34. https://doi.org/10.1016/j.envsoft.2011.11.009

Culebra. (2022, April 3). Food4Rhino. https://www.food4rhino.com/en/app/culebra

Feng, Z., Gu, P., Zheng, M., Yan, X., & Bao, D. W. (2021). Environmental data-driven performance-based topological optimization for morphology evolution of Artificial Taihu Stone. In *Proceedings of the 3rd International Conference on Computational Design and Robotic Fabrication (CDRF 2021)* (pp. 117-128). Springer Nature, https://doi.org/10.1007/978-981-16-5983-6_11

DeSalvo, G.J., & Swanson, J.A. (1985). ANSYS Engineering Analysis System User's Manual. Houston, Pa.: Swanson Analysis Systems.

Hack, H. R. (2019). Weathering, erosion, and susceptibility to weathering. In: M. Kanji, M. He, L. Ribeiro e Sousa (Eds), Soft *Rock Mechanics and Engineering* (pp. 291–333). Springer, Cham. https://doi.org/10.1007/978-3-030-29477-9

Hall, K. (1989). Wind-blown particles as weathering agents? an Antarctic example. *Geomorphology*, 2(4), 405–410. https://doi.org/10.1016/0169-555x(89)90024-x

Hu, H. H. (2012). Computational Fluid Dynamics. *Fluid Mechanics*, 421–472. https://doi.org/10.1016/b978-0-12-382100-3.10010-1

Kabošová, L., Katunský, D., & Kmet, S. (2020). Wind-Based Parametric Design in the Changing Climate. *Applied Sciences*, 1-18.

Kabošová, L., Kmet, S., & Katunský, D. (2019). Digitally Designed Airport Terminal Using Wind Performance Analysis. *Buildings 9*, 59.

Kataoka, H., Ono, Y., & Enoki, K. (2020). Applications and prospects of CFD for wind engineering fields. *Journal of Wind Engineering and Industrial Aerodynamics*, 205. https://doi.org/10.1016/j.jweia.2020.104310.

Kormaníková, L., Chronis, A., Kmeť, S., & Katunský, D. (2018). Wind-formed Architectural Shapes. In *Proceedings of the 36th International Conference on Education and Research* *in Computer Aided Architectural Design in Europe: Computing for a better tomorrow, eCAADe 2018* (pp. 377-384). The Education and research in Computer Aided Architectural Design in Europe (eCAADe).

Montazeri, H., Blocken, B., & Hensena, J. (2015, January). Evaporative cooling by water spray systems: CFD simulation, experimental validation and sensitivity analysis. *Building and Environment, 83*, 129-141. https://doi.org/10.1016/j.buildenv.2014.03.022

Rezakazemi, M., Shahverdi, M., Shirazian, S., Mohammadi, T., & Pak, A. (2011). CFD simulation of water removal from water/ethylene glycol mixtures by pervaporation. *Chemical Engineering Journal, 168*(1), 60-67. https://doi.org/10.1016/j.cej.2010.12.034

Scott, K., & Pain, C. F. (2011). Rock Weathering and Structure of the Regolith. In K.M. Scott and C.F. Pain (Eds.), *Regolith science* (pp. 103–124). CSIRO.

Snooks, R., & Stuart-Smith, R. (2010). Non-linear formation: or how to resist the parametric subversion of computational design. In Xanthropoulos P, Zindrou I (Eds), *Apomechanes: nonlinear computational design strategies* (pp 140–147). Harry Tzannis.

Tavsan, F., & Sonmez, E. (2015). Biomimicry in furniture design. *Procedia - Social and Behavioral Sciences*, 197, 2285–2292. https://doi.org/10.1016/j.sbspro.2015.07.255

Toja, F., Peralta, C., Lopez-Garcia, O., Navarro, J., & Cruz, I. (2015). On Roof Geometry for Urban Wind Energy Exploitation in High-Rise Buildings. *Computation, 3*, 299-325. https://doi.org/10.3390/computation3020299.

Toja, F., Kono, T., Peralta, C., Lopez-Garcia, O., & Chen, J. (2018). A review of computational fluid dynamics (CFD) simulations of the wind flow around buildings for urban wind energy exploitation. *Journal of Wind Engineering and Industrial Aerodynamics,* 180, 66-87. https://doi.org/10.1016/j.jweia.2018.07.010

Toja-Silva, F., Lopez-Garcia, O., Peralta, C., Navarro, J., & Cruz, I. (2016). An empirical–heuristic optimization of the building-roof geometry for urban wind energy exploitation on high-rise buildings. *Applied Energy*, 164, 769-794. https://doi.org/10.1016/j.apenergy.2015.11.095

Wang, L., Tan, Z., & Ji, G. (2016). Toward the wind-related building performative design a wind-related building performance optimization design system integrating fluent and rhinoceros based on iSIGHT. In *Proceedings of the 21st International Conference on Computer-Aided Architectural Design Research in Asia: Living Systems and Micro-Utopias: Towards Continuous Designing, CAADRIA 2016* (pp. 209-218). The Association for Computer-Aided Architectural Design Research in Asia (CAADRIA).

Xie, Y., Zuo, Z., & Lv, J. (2014). Liyong shuangxiang jianjin jiegou youhua suanfa jinxing jianzhu sheji [Application of bi-directional evolutionary structural optimization on architectural design]. *Time+Architecure, 5*, 20-25.

Zhou, Q., Shen, W., Wang, J., Zhou, Y., & Xie, Y. (2019). Ameba: A new topology optimization tool for architectural design. In *Proceedings of International Association for Shell and Spatial Structures Annual Symposia: Scandinavian Exceptionalisms, IASS 2018* (pp. 1-8). The International Association for Shell and Spatial Structures (IASS).

ARTICULATING FACADE MICROBIOMES AT HUMAN SCALE

A Cellular Automata Driven Bioreceptive Facade Design Approach to Communicate a New Dimension of Urban Health

CHRISTIANE M. HERR[1] and CHENXIAO LI[2]
[1,2]Southern University of Science and Technology, Shenzhen, China.
[1]cmherr@sustech.edu.cn, 0000-0002-1381-8288
[2]licx@mail.sustech.edu.cn, 0000-0002-3871-8312

Abstract. Rapid urbanization has led to the deterioration of urban living environments and reduced urban populations' access to green space. This has not only affected urban residents' health but also decimated biodiversity in urban environments. In this study, we respond to both issues by introducing a new approach to façade design focusing on microbial biodiversity on building surfaces. Based on the results of an earlier empirical study, we use a custom Cellular Automata (CA) system as generative design strategy to explore the relationships between microbial community diversity and several design factors relevant to creating favourable building surface conditions, in particular surface texture, substrate material characteristics and sunlight exposure. By translating these factors into localized CA design parameters, we create micro-habitats supporting microbial biodiversity in bio-receptive façade elements. Beyond generating desired physical shapes, we employ CA to generate expressive patterns as visually comprehensible articulations of invisible scales of microbial growth.

Keywords. Cellular Automata, Generative Design, Bio-receptive Facade Design

1. Introduction

Rapidly increasing urbanization across the world, in particular in Asia, has led to much of the world's population now living in environments characterized by artificial materials, limited access to green spaces as well as high levels of pollutants and noise. Consequently, urban residents living in such environments increasingly experience health issues, many of them related to autoimmune deficiencies (Flies et al., 2017) linked to lack of exposure to environments featuring diverse microbial communities. In this context, access to natural environments has been identified as a key resource for supporting public health within the broader goal of protecting and enhancing urban biodiversity across all scales.

In response, urban and architectural design are called upon to create environments that respect, protect and (re)integrate natural biodiversity into urban environments, as

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 281-290.

in various recent initiatives aiming to design “multi-species cities” or to “re-wild” cities (Mills et al., 2017; Roudavski, 2020). The benefits of introducing biodiversity into artificial environments extend beyond the visible scale of organisms and relate to a significant extent to microbial life and its key role in human health. This paper presents a particular aspect of a broader study that examines microbial diversity on building facades, with the aim of creating bio-receptive façade elements hosting microbial communities beneficial to human health. Initial results of the earlier empirical part of the study indicate a strong correlation between diversity in microbial communities and several design factors, most importantly substrate material characteristics and sunlight exposure, as shown in Herr and Duan (2020). Translating these factors into design parameters for façade elements offers a broad range of possibilities for creating micro-habitats supporting microbial biodiversity.

In this paper, we focus on a particular aspect of the project, the design of façade patterns to not only support diversity in microbial growth, but also to communicate the nature of this new type of façade element to both inhabitants and the public – which we consider relevant since microbial diversity is invisible to the human eye. In the following, we present a digital design approach that integrates physical design requirements derived from empirical research with cellular automata-based generative design of biofilm-like growth patterns, translating the design intention of promoting microbial growth and biodiversity into visually legible scales of façade elements.

2. Background

Distinct from a range of recent studies discussing “bioreceptive” architecture with a strong focus on lichen and bryophyte growth (Cruz and Beckett, 2016) we adopt the term of bioreceptivity in Guillite’s original definition as *“the aptitude of a material to be colonized by one or several groups of living organisms without necessarily undergoing any biodeterioration”* (Guillite, 1995, p. 216). A bioreceptive façade element thus denotes a façade element that can support a biofilm and potentially also larger organisms for an extended period of time, with its material providing a suitable substrate for biofilm growth (Gorbushina, 2007; Pentecost and Whitton, 2012). Previous empirical studies point out that key physical and chemical material characteristics improving bioreceptivity include increased surface roughness, porosity, capillary water content, an abrasion pH below 10, surface geometry and substrates supporting various metabolic functions (Miller et al., 2012; Herr and Duan, 2020).

Cellular automata (CA) are discrete massively parallel digital simulation tools that have been employed for various architectural design purposes (Kim, 2013; Herr, 2015; Zawidzky, 2015). They have also found wide applications to model biofilm growth, as demonstrated by Wimpenny and Colasanti (1997), but these two areas of cellular automata application have not been connected previously. Developing a generative method for bio-receptive façade design at the pattern level, we employ CA predominantly for exploration and speculation, inspired by Cedric Price's *Generator* project.

Despite their inherent simplicity, CA are capable of generating intricate dynamics and properties difficult to predict for external observers and thus often described as “emergent”. Architects have adopted CA-based patterns for various uses, for example in “open” adaptive façade design strategies (Kim, 2013) for increased visual interest.

Other studies focusing on daylighting criteria have tested CA-based façade design proposals under static and dynamic conditions (Zawidzki, 2015; Ayoub, 2018).

In the following sections, we employ CA-based patterns in a directed manner to generate physical design proposals in a custom developed CA model, aiming to support biofilm growth while also featuring some form of visual legibility. The paper first outlines relevant façade design parameters identified based on empirical results examining microbial communities on building facades (located in South China). The second part of the paper introduces cellular automata-based modelling of biofilm growth patterns. The third part of the paper applies selected design parameters to CA patterns and presents first results of three-dimensional façade patterns translating invisible microbial scales of growth to the scale of conventional building facades.

3. Methodology

We employ a custom-designed CA tool, implemented in Rhino Grasshopper IronPython 2.7, that offers a layered approach to physical pattern generation to afford designers a certain level of control over the inherent tendency of CA dynamics. In addition to the conventional approach of developing rule sets to control outcomes per se, we also employ rule sets that interact with a variable “substrate” layer, offering a more indirect level of design control. Figure 1 illustrates the generative sequence.

According to previous empirical studies, surface recesses of substrates have a positive correlation with biofilm growth. Thus, to generate a bio-receptive facade surface, we focus on aspects of geometry first, in particular to ensure an overall proportion of surface recesses. This approach is carried out at two levels of the facade design, macro-level and micro-level: (1) the macro level defines the topological expression of the CA system with respect to the overall façade surface using the analogy to nutrient availability in natural biofilm growth, and (2) at the micro level, invisible behaviours are rendered visible as colour gradients and transformed into surface recesses. The CA models developed for this study deviate somewhat from classic CA models in that they combine a variety of project specific mappings and translations of CA dynamics to offer some degree of design control. This aspect is of relevance beyond this project as most previous applications of CA for design purposes do not explicitly discuss the issue of design control.

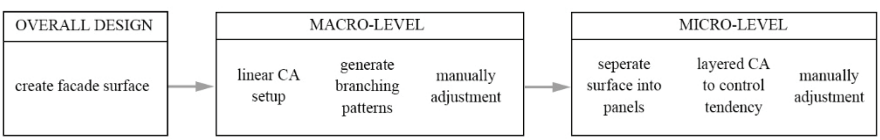


Figure 1. CA-based bioreceptive façade generation process

4. CA Design

The CA model presented here forms the core generative method for the following “layers” of CA application. The first step of the ruleset is to generate a branching pattern in analogy to microbial growth patterns in nature. In the second step, we vary the CA parameters to generate a series of alternative patterns.

4.1. MACRO-LEVEL: LINEAR CA FOR BRANCHING PATTERNS

The CA system operates as a vertically growing linear CA system, which is then mapped to a surface in Rhinoceros 3D, with UV values used as grid coordinates. As shown in Figure 2c, each cell has three possible values: “0”, “1”, or “2”, corresponding to white, grey and black colour. To generate the desired diversity in branching patterns, we modified the CA rules developed for an earlier project called Smart Branching (Li at al., 2022), adding a two-step logic. Figure 2a illustrates the vertical growth dynamics of the CA system, distinguishing two rule sets, one for vertical growth in “empty” space and a second rule (Figure 2b) determining growth at intersection points.

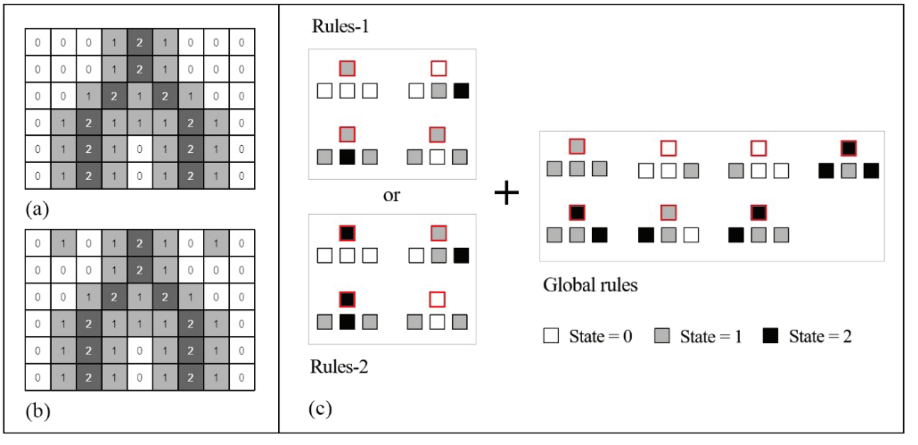


Figure 2. Patterns generated by the linear CA system: (a) the first version of customized CA rules; (b) adjusted CA rules; (c) rule sets for the macro-level model: rule 1: for vertical growth: [rule 1 exclusive] + global rules; rule 2: for intersection and variety: [rules 2 exclusives] + global rules.

The designer directly controls the initial (bottom) two rows of cells to determine the starting points of vertical branches. Cell dynamics follow the conventions of classic linear CA, with local transition rules (TR) for each cell taking into account their own current state as well as the states of the three nearest cells below. For instance, TR (0,1,2) indicates that in one step of CA development, if the three cells below the self-cell value 0, 0, 0 from the left to right, and if the self-cell is not within a pre-defined “increased nutrient availability” area, then the state transition follows rule-1: (0,1,2) → 0. But if the self-cell is within a pre-defined “increased nutrient availability” area as further detailed below, the state transition follows rule-2: (0,1,2) → 1.

To offer a first additional layer of design control, the linear CA system is mapped to a Rhino3D surface with variable properties influencing the linear CA system’s growth, in analogy to a physical substrate providing varying conditions for microbial growth. To allow the linear CA system to “read” the properties of the (façade) surface on which it develops, the overall surface is visualized in Rhino3D's ‘Arctic’ display style. The grey-scale output image is later transferred into a pixelated mesh, so that Rhino3D's IronPython 2.7 GHPython components can read the mesh vertex and its corresponding “lightness” value. This value determines an “attraction force” that influences the growth of the linear CA system. Further details describing the layered

CA growth model are explained in the following sections.

4.2. MICRO-LEVEL: A DIFFUSION MODEL RESPONDING TO CONTEXT

At the micro level, we aim to further control the growth tendency of the linear CA system in a gradual and non-obtrusive manner, which we implement by using gradients in analogy to nutrition availability in natural substrates for microbial growth. To this end, the CA system is set up as an evenly distributed grid of 2D points grid, corresponding to the plan of a 500*400mm (width * height) 3D panel (Figure 3), a constituent element of the larger façade surface generated in the first step of the generative process. At this stage, the CA development follows a new and more flexible set of transition rules different from those at the macro-level, where the cell locations are allowed to shift on the surface grid similar to an agent-based model.

Figure 3. (left) 500*400mm (width * height) panel area; (middle) gray-scale sample image; (right) original point grid.

In each single step of CA development, the self-cell cycles through the neighbours within a unit to determine cell density to decide the next state of the self-cell according to a new transition rule. The micro-level transition rules take into account the state of the self-cell, neighbour cells and a pixelated "substrate" surface grid. The states of the surface grid are evaluated in addition to direct neighbours within a pre-defined radius.

The cell states generated in this micro-level evaluation will translate into relative movements of "aggregation" or "detachment" depending on the "attraction force" calculated based on the states of cells within a wider radius. To understand the effects of this process, we tested two different kinds of transition rules: a static one (where rules stay the same) and a dynamic one (where rules can change between iterations).

5. Exploration of the layered generative CA design approach

The overall design approach features two steps integrating the macro- and micro-level models. In the first step, the CA system generates an 8*12m (W*H) facade base with feasible parameters derived from previous experiments, with subsequent manual modification of the resulting surfaces. The surface is then split into 50 equal panels. The overall surface is expressed in Rhino3D's 'Arctic' display style and rendering results are imported for testing at the micro-level for further design.

At the macro-level CA "layer", the design parameters tested are: (1) input state value, and (2) one evolution rule (0,0,0) → 0. A 30*50 matrix was generated on the 8*12m (width*height) façade. In the first comparison, the original state of all cells is "0", so we evenly distribute "state=1" and "state=2": "if i%3 == 0 and j%3 == 0: self_state = 1; if i%5 == 0 and j%3 == 0: self_state = 2"; and additive step to renew parts of the cells: "if i%3 == 0 and j%3 == 0: self_state = 0". The rules are then run and cells are updated with distance measurements to three input attractor points, denoting "nutrient availability". Cells within a given attractor radius follow global rules and rule 2, others follow global rules and rule 1. To make the outcome more readable in terms of shape, the point grid is transferred into a 3D mesh, and the state of each cell point is given a depth value indicating a layer of recess from the original surface.

To systematically assess the generated surface characteristics, we apply the Visual Character Analysis (VCA) method involving viewport capture and image processing in Rhino3D™ proposed by Stuart-Smith and Danahy (2022). We simplify this method to focus on our primary concern of creating and managing 3D surface recesses. Outcomes are expressed as grey-scale images and evaluated through a custom script in Rhino3D's IronPython 2.7 that reads mesh vertices relative to the lightness value of local mesh colour, resulting in a quantitative evaluation of surface recesses on a given surface (Figure 4).

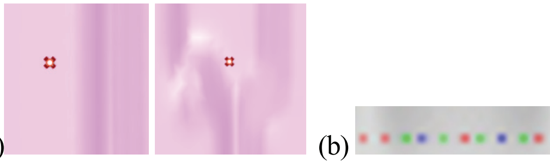
(a)  (b)

Figure 4. (a) The comparison of surfaces generated from initial and revised CA rules, showing a more effective response to the "nutrition" points. (b) pre-defined initiating cell sequence

At the micro-level, the CA development mostly depends on the surface grid values denoting nutrient availability. Each CA cell is seeking a local balance between the geometric characteristics of the substrate material (denoted by lightness values) and density of the community (denoted by cell density). On this base condition, we tested a further layer of complexity, consisting of static or dynamically evolving rule sets.

We first tested a static model to determine the characteristics of the generated outcomes. In this model, the neighbours are defined by their index and remain static in later iterations, along with a simplified attractor model recognizing three simplified categories of lightness on the base grid. The most representative results are shown in the below "results" section. To understand and visualise the cell development or "growth" of the CA system over time, we also record the trajectory of the cells.

A second model examined a more dynamic development mode, where density and attractor points within a pre-defined neighbourhood radius are the key determinants: if the distance between self-cell and the target cell is less than a given value, the target cell is considered a neighbour. If the population of neighbours per unit is less than five cells, the self-cell will implement "aggregation" behaviour, and check all corresponding lightness values of the pixels to look for the most "nutritious" location. If local density exceeds five cells per unit, "detachment" behaviour is prioritized.

6. Results

Results generated from the above outlined explorations indicate that the layered CA-based generative system is able to offer designers a degree of control while "growing" surface patterns emulating the growth of microbial communities on surfaces. The below sections detail the results of this ongoing experimental process, which we will implement in the production of physical bioreceptive façade panels.

6.1. MACRO-LEVEL: OVERALL FAÇADE MODEL

The macro-level model generated a series of overall surface patterns (table 1). Considering the overall architectural façade composition, we selected the second, fifth and fifteenth generations as most suitable for further development. In the selected results, the A test series contains more variations in surface recesses but only responds to 2/3 of the substrate related rules, with only 1/3 responding in test series C. We selected series B as most appropriate for a more evenly distributed façade application.

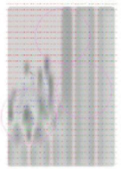
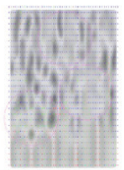
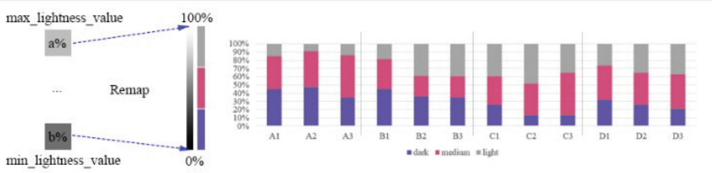
input state	[1, 2]	[1, 2]	[0, 1, 2]	[0, 1, 2]
rule 0	0, 0, 0 → 0	0, 0, 0 → 1	0, 0, 0 → 0	0, 0, 0 → 1
gen = 2	 A1	 B1	 C1	 D1
gen = 5	 A2	 B2	 C2	 D2
gen = 15	 A3	 B3	 C3	 D3
VCA result				

Table 1. CA generated results and evaluation of the first series of tests at macro-level

The Visual Character Analysis (VCA) method (table 1, bottom) offers insights into the generative parameter interrelationships. As the number of CA generations increases, the proportion of dark areas decreases, the proportion of light areas increases, and the proportion of grey remains constant. While parameters are varied, the overall VCA results do not show a distinctive corresponding trend in recess distribution.

6.2. MICRO-LEVEL: SURFACE VARIABILITY

“Static” CA rule models offer regularity in generated outcomes compared to “dynamic” CA rule models, which generate visually “noisier” results. Increasing iterations in the static model produce fewer visible points in the panel area, and the overall cells show regular convergence. The dynamic model, in contrast, does not show an obvious density change aside from some convergence behaviour.

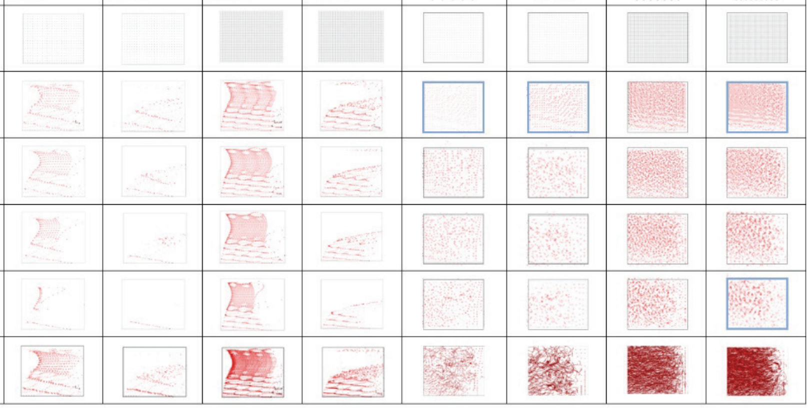
input image	self.state = attract_force = k * Lightness				attract_force = k * Lightness			
CA grid	U*V=20*25 H*W=400*500mm	U*V=20*25 H*W=400*500mm	U*V=40*50 H*W=400*500mm	U*V=40*50 H*W=400*500mm	U*V=20*25 H*W=400*500mm	U*V=20*25 H*W=400*500mm	U*V=40*50 H*W=400*500mm	U*V=40*50 H*W=400*500mm
parameters	k = 0.05	k = 0.09	k = 0.05	k = 0.09	thredshold = 20	thredshold = 30	thredshold = 20	thredshold = 30
	neigh_by_index	neigh_by_index	neigh_by_index	neigh_by_index	orig_neigh = 5/unit	orig_neigh = 9/unit	orig_neigh = 13/unit	orig_neigh = 25/unit
gen = 0								
gen = 2								
gen = 5								
gen = 10								
gen = 20								
trajectories over time								

Table 2. Collection of the results and assessment in the first series of tests at macro-level

Table 2 illustrates how recording the “cell motion” points in different generations renders the development trends over time visible and allows assessment for further CA modifications.

6.3. EXPERIMENTAL DESIGN APPLICATIONS

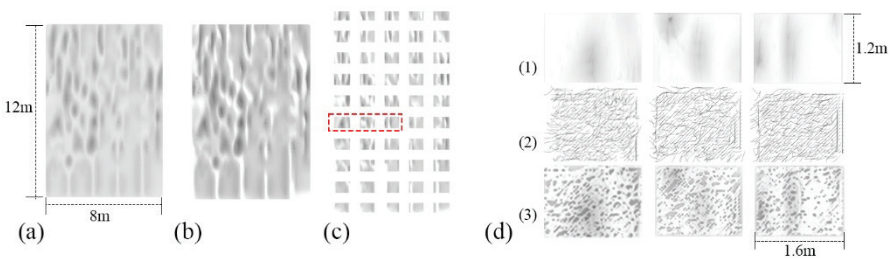


Figure 5. The generative process of the experimental design application.

The results generated at each stage in the experimental design are shown in Figure 5. The macro-level model first generated the refined façade surface. Then with manually modified and split, the whole surface was separated into 50 1.2m*1.6m panels of same width and height. Then project intricate curves generated at micro-level onto the panel surface to form 3D geometry. The three adjacent sample panels of three different texture types are selected and displayed in Figure 5(d). Figure 6 illustrates a first tentative application of the generated geometry to a physical facade panel, consisting of a prefabricated composite panel comprising several 10mm deep material sheets, comprising metal, growth substrate and organic materials. Details will be determined at the next stage of physical prototype tests.

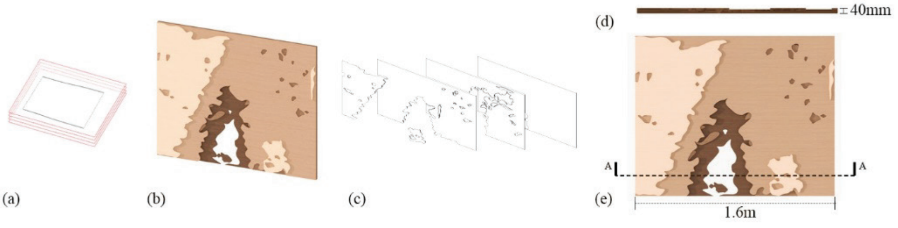


Figure 6. First façade panel prototypes: (a) Schematic diagram of panel layers; (b) and (c) panel composed of four 10mm deep material sheets; (d) proposed panel dimensions.

7. Summary and Outlook

The experiments presented here outline interim results from the ongoing development of a CA-based method to generate façade assemblies and elements with two aims: The generated geometry follows a decentralized and locally determined logic in analogy to microbial growth patterns and can be employed to not only generate varied surface geometries supportive of bio-colonization, but also to communicate these growth dynamics visually to a broader public. To achieve this goal, we develop a novel, "layered" way to design with the dynamic, locally-generated expression of the CA system that can support architectural design. Employing this approach, we control average variation of curved surface geometries on a basic level and create superimposed intricate surface patterns at the same time.

Contributions offered by this CA-based design method include a CA model employing variations of rules in relation to the states of a "base surface," which offers increased control for human designers and extends previous research by Li et al. (2022) as well as Herr (2015). At this current stage, we focus primarily on the interaction of the "layers" of the CA system, whereas the next stage of research will include a more systematic investigation of the CA rule sets. A further contribution to previous CA studies is the integration of CA-generated form with VCA evaluation as proposed by Stuart-Smith and Danahy (2022). Yet other evaluation methods currently under development include the production of panel prototypes for empirical testing of various surface geometry and material configurations to map their effectiveness in supporting microbial biocolonization.

References

Ayoub, M. (2018). Integrating illuminance and energy evaluations of cellular automata controlled dynamic shading system using new hourly-based metrics. *Solar Energy*, 170, 336–351. https://doi.org/10.1016/j.solener.2018.05.041

Cruz, M. and Beckett, R. (2016) Bioreceptive design: A novel approach to biodigital materiality, *Architectural Research Quarterly*, 20(01), 51-64, https://doi.org/10.1017/S1359135516000130.

Flies, E.J., Skelly, C., Negi, S.S., Prabhakaran, P., Liu, Q., Liu, K., Goldizen, F.C., Lease, C. and Weinstein, P. (2017). Biodiverse green spaces: a prescription for global urban health, Frontiers in Ecology and the Environment, 15(9), 510-516, https://doi.org/10.1002/fee.1630.

Gorbushina, A. A. (2007). Life on the rocks. *Environmental Microbiology*, 9(7), 1613–1631. https://doi.org/10.1111/j.1462-2920.2007.01301.x

Guillitte, O. (1995). Bioreceptivity: a new concept for building ecology studies. *Science of the Total Environment*, 167(1–3), 215–220. https://doi.org/10.1016/0048-9697(95)04582-l

Herr, C.M. (2015) Second-order cellular automata to support designing. *Kybernetes* 44(8/9), pp. 1251-1261. https://doi.org/10.1108/K-11-2014-0268.

Herr, C.M. and Duan, Y. (2020). Designing Facade Microbiomes. In *Imaginable Futures: Design Thinking and the Scientific Method, Proceedings of the 54th Intl Conference of the Architectural Science Association* (pp. 365-374). The Architectural Science Association.

Kim, J. (2013). Adaptive façade design for the daylighting performance in an office building: the investigation of an opening design strategy with cellular automata. *International Journal of Low-Carbon Technologies*, 10(3), 313–320. https://doi.org/10.1093/ijlct/ctt015

Li, C., Yuan, M., Han, Z., Faircloth, B., Anderson J. A., King, N., & Stuart-Smith, R. (2022). Smart Branching: An Experimental Method for Heterogeneous Branching Networks using Non-Planar 3D Printed Clay Deposition. In *42nd Annual Conference of the Association for Computer Aided Design in Architecture: Hybrids & Haecceities, ACADIA 2022*. The Association for Computer Aided Design in Architecture (ACADIA).

Mills, J.G., Weinstein, P., Gellie, N.J.C., Weyrich, L.S., Lowe, A.J. & Breed, M.F. (2017). Urban habitat restoration provides a human health benefit through microbiome rewilding: the Microbiome Rewilding Hypothesis. Restoration Ecology, 25(6), pp. 866-872. https://doi.org/10.1111/rec.12610

Miller, A., Sanmartín, P., Pereira-Pardo, L., Dionísio, A., Saiz-Jimenez, C., Macedo, M., & Prieto, B. (2012). Bioreceptivity of building stones: A review. *Science of the Total Environment,* 426, 1–12. https://doi.org/10.1016/j.scitotenv.2012.03.026

Pentecost, A., Whitton, B.A. (2012). Subaerial Cyanobacteria. In: Whitton, B. (eds) *Ecology of Cyanobacteria II* (pp. 291–316). Dordrecht, DOI:10.1007/978-94-007-3855-3_10.

Roudavski, S. (2020) Multispecies Cohabitation and Future Design. In *Proceedings of Design Research Society (DRS) 2020 International Conference: Synergy* (pp. 731-750). Design Research Society, https://doi.org/10.21606/drs.2020.402.

Stuart-Smith, R., & Danahy, P. (2022). Visual Character Analysis within Algorithmic Design: Quantifying Aesthetics Relative to Structural and Geometric Design Criteria. In *24th International Conference on Computer-Aided Architectural Design Research in Asia: Intelligent and Informed, CAADRIA 2019* (pp. 131-140). The Association for Computer-Aided Architectural Design Research in Asia (CAADRIA),.

Wimpenny, J. & Colasanti, R. (1997) A unifying hypothesis for the structure of microbial biofilms based on cellular automaton models. FEMS Microbiology Ecology, 22(1), pp. 1–16, https://doi.org/10.1111/j.1574-6941.1997.tb00351.x

Zawidzki, M. (2015). Dynamic shading of a building envelope based on rotating polarized film system controlled by one-dimensional cellular automata in regular tessellations (triangular, square and hexagonal). *Advanced Engineering Informatics*, 29(1), 87–100. https://doi.org/10.1016/j.aei.2014.09.008

URBANISING MARS

Establishing An Urban Framework For Martian Settlement Through Sequential Evolutionary Simulations

LE ANH GIANG[1], MINH TRIET TRAN[2], STEFANO BALOSSINO[3] and MOHAMMED MAKKI[4]
[1,2,4] *University of Technology Sydney*
[3] *Politecnico di Milano*
[1]*leanh.giang@student.uts.edu.au, 0000-0001-6424-7094*
[2]*minhtriet.tran-1@student.uts.edu.au, 0000-0002-2636-4043*
[3]*stefano.balossino@mail.polimi.it, 0000-0002-6121-7110*
[4]*mohammed.makki@uts.edu.au, 0000-0002-8338-6134*

Abstract. Current initiatives for urbanising space are focused on three key locations; Orbital settlements; Lunar settlements and Martian settlements. Of the three, Martian settlements present themselves as the most hospitable location for interplanetary colonisation due to the various similarities between Earth and Mars. Despite significant scientific research for the settlement of Mars, there lacks an equally significant Urban analysis and impact of such settlements. The presented study explores the habitation of Mars through an urban lens of efficiency that responds to geographic and climatic conditions, as well as addressing behavioural challenges of a human settlement on Mars. The study is presented through two key stages, the first establishes life support modules, while the second develops an urban solution for habitation. In both stages, a multi-objective evolutionary algorithm is implemented, with a focus on the selection mechanisms that choose the optimal solution from the population generated by the algorithm. The results showcase an adaptable model that allows for multiple stakeholders with varying expertise the ability to influence the design of a Martian settlement.

Keywords. Martian Urban Settlement, Extra-terrestrial Urbanization, Valles Marineris, Interplanetary Colonization, Multi-Objective Evolutionary Algorithm (MOEA).

1. Introduction

For decades, there has been increased interest in space colonisation for a variety of reasons, such as reducing human impact on earth, finding off-planet resources, setting bases for further exploration into space or discovering potential avenues for improving future urban conditions on Earth. Interplanetary colonisation, seemingly new, has been

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 291-300.

targeted by human for centuries and since then, a significant amount of research-based urban proposals have been made with an attempt to turn this idea into real matter (Reid, 2019). Current initiatives for urbanising space are focused on three key locations; orbital settlements, primarily between earth and the moon; Lunar settlements and Martian settlements (National Space Society, 2017). Where an orbital settlement is dependent on the nearest planet for resources (the most immediate example being the international space station), a habitat located on the Moon or Mars can be sustained by locally sourced materials. When comparing the two, there is a stronger argument for a martian settlement than a lunar one (Zubrin & Wagner, 2011), primarily due to the similarities between Mars and Earth. Despite differences in size (Mars is half the size), year duration (twice as long as Earth's year), distance from the sun, temperature and temperature variations between Mars and Earth, the potential to inhabit and urbanise Mars outweighs its drawbacks (Zubrin, 2014). This is driven by the red planet's capacity to provide the basic elements for human survival and growth through the availability of raw and life-essential materials such as treatable air, permafrost layer (for water extraction), and solar energy (Petranek, 2015). Moreover, a 24-hour day/night cycle, four seasons and the potential for plant growth and food generation categorises Mars as one of the most suitable extra-terrestrial locations for a self-sustaining settlement (Moses & Bushnell, 2016).

This paper examines the 2040 scheduled timeline for Mars exploration (Nasa, 2021), which as per the projected timeline, assumes that the essential life support mechanisms have been established. By this date, mankind is technically, technologically, financially, and socially prepared to build the first colony on Mars. between the years 2040 and 2050, the first settlement on Mars would be expected to be reaching a target population of roughly 3331 inhabitants to form the first self-efficient and sustainable module (Szocik et al., 2020). Through the use of sequential multi-objective evolutionary algorithms (MOEA) with a thorough selection process applied to the generated results, the presented experiment investigates the potential to integrate urban sustainability, efficiency and self-sufficiency to the establishment and growth of the first Martian colony. The results present how the use of multi-objective decision making processes play a key role in the integration of urban design goals to inhabit an extreme climate.

2. Background and Context

2.1. MARTIAN ENVIRONMENT

Mars is a cold desert world of which the surface has many geological features that have recognizable counterparts on Earth, some of which are relatively larger compared to those on our planet including volcanoes, canyons and river channels. Due to its distance from the sun, a year on Mars is equal to 687 Earth days. Approximately half the size of our planet, the red planet has one-third the gravity of earth and a much thinner atmosphere, which contains more than 95% carbon dioxide and less than 1% oxygen. Water is present only in very low abundance on Mars in its atmosphere, surface ice and regolith (Haberle et al., 2017). The temperature of the Martian surface produces permafrost, due to the mean average temperature near the Martian equator reaching -60 degrees Celsius (Leighton & Murray, 1966). Consequently, there are elevated levels of solar radiation due to no stratosphere (Catling & Kasting, 2017). These conditions significantly affect design strategies to promote a balance between segregation and interaction between the urban habitat and its context.

2.2. SITE - VALLES MARINERIS

Identifying the location on Mars for establishing the habitat is driven by several key factors; these include temperature, geography and water availability. While Mars is generally too cold for humans to inhabit, selecting warmer and low altitude areas would benefit future urban settlements in terms of energy efficiency and protection from natural extreme phenomena such as sand storms. These areas are categorised as D (Tropical), E (Low albedo tropical) and G (Tropical Lowland) based on temperature, modified by topography, albedo, and actual solar radiation (Hargitai, 2010). Missions to Mars will need to be as self-sufficient as possible, which means using local resources to meet the needs of the mission and astronauts – a process known as in-situ resource utilisation (ISRU) (Moses & Bushnell, 2016). Water is another determining factor for site selection of the first urban colony, of which a significant amount has been found at the heart of Mars' dramatic canyon system, Valles Marineris, located in Martian climate zone E (European Space Agency, 2021). This water-rich area (similar in size to the Netherlands) overlaps with the deep valleys of Candour Chaos, part of the canyon system considered promising in the formation of the human settlement (Figure 1).

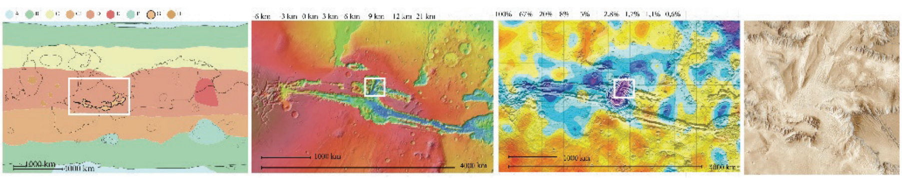


Figure 1. Analysis of site based on (from left to right) Martian climate zone, altitude and percentage of solid water under surface.

2.3. URBAN CHALLENGES OF A MARTIAN COLONY

Human space settlement may be regarded as the next step in human expansion; this part of the paper discusses a broad set of challenges for creating a human settlement on Mars, which should be taken into consideration by mission planners. In the possible scenarios of a human society being formed on Mars, there are three aspects among many urban challenges that impact the urban growth and configuration of the settlement.

2.3.1. Life-Support System

Unlike Earth, where breathable air, drinkable water, energy, and material of different forms naturally exist, inhabiting Mars must address the technical challenges of creating and maintaining life-support systems such as air, water and energy. This changes the urban structure of the settlement and affects site location, module hierarchy and design decisions. This issue is addressed in the first design stage where population-based efficient location and area of essential facilities such as water, oxygen and energy are generated based on a set of constraints and relative connections to provide future inhabitants with a self-efficient and self-sustaining life-support framing system.

2.3.2. Protection and Interaction

Martian extreme climate conditions, as described above, demand complete segregation between internal and external living environments for human survival, yet it is important for the urban settlement to visually respond to the existing background of Mars. Considering these challenges, proposals of active and/or passive shielding from extreme solar radiation and Martian harsh climatic conditions are among the prioritised solutions, which influences morphologies, façade and material use of different facilities depending on functions. Enclosed, covered and underground spaces are part of the strategies.

2.3.3. Efficiency and Sustainability

If humans were to make their lives multi-planetary, sustainability must be core to the habitation framework. Efficiency in transportation networks, facility density and distribution, resource production, and sustainability in how each module functions and supports each other provide the Martian module with the capacity to be self-efficient and make the best use and reuse of the limited resources available.

3. Methods and Experiment Setup

The first phase of this study addresses the technicality of survival on Mars, where the location and overall size of life support systems are generated based on population needs and relative relationships between each module. As a result of this framework, the second phase examines habitation in response to multiple urban objectives, focusing on system efficiency and quality of life. Both phases are examined through a sequential automated and semi-automated algorithmic process that utlise a multi-objective evolutionary algorithm, through the software Wallacei (Makki et al., 2018) within the Rhino 3d and Grasshopper 3D ecosystems. The formulation of both sequential simulations is presented below.

3.1. SEQUENCE 1

3.1.1. Constructing the Phenotype

The phenotype (the geometry that the algorithm evolves) of stage one focuses on defining the boundary of six main modules that form the foundation of the Martian habitat: Energy, Water, Oxygen, Dwelling, Agriculture, and Launching Pad. The goal of this stage is to construct a basic framework that identifies the most suitable locations for each module based on their relationship to one another. Therefore, the phenotype's construction is highly sequential, in which modules are established based on their significance to ensure viability of the colony. As such, the simulation is divided into 5 major parts. As Energy, Water, and Oxygen are the most essential for human survival, these structures' boundaries are defined first. The Dwelling space is established next, driven primarily by the efficiency of resource delivery to the various habitats. This is followed closely with the identification of recreational spaces. Although critical, the location of agricultural land is identified in response to solar gain and the impact of the agricultural land (which is the largest in terms of area) on the surrounding facilities. Finally, the Landing Pad location is identified, ensuring it is located 300 metres from the nearest modules.

The modules are located within a one square kilometre site that utilises a 50-metre grid as a guiding system, primarily driven by the objective of achieving an urban tissue that allows for a 15-minute walking distance throughout the settlement. To achieve stage one's goal and allow the simulation to run effectively, four key variables (the chromosomes that the algorithm mutates) define the construction of the phenotype: the starting location of each module, the direction of growth for each module, the size of habitable spaces, and finally, the size and location of the agricultural area (Figure 2).

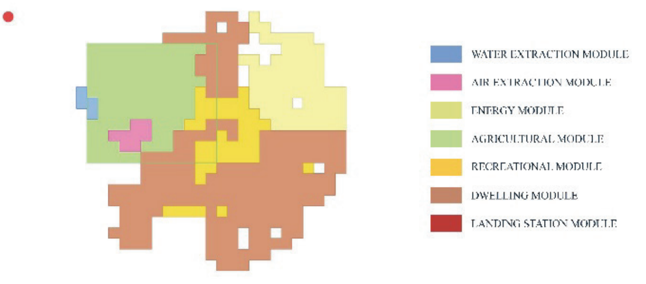


Figure 2. The base phenotype for the first evolutionary sequence.

3.1.2. Fitness Objectives

The design goals (fitness objectives) that the algorithm will optimise for have been defined for the purpose of achieving operational efficiency and safety of the selected modules. In total, four fitness objectives have been identified; first, to maximise the distance between the Water and Energy modules so as to avoid any overlap between the two which may put either (or both) modules at risk. The second objective maximises food production through location and solar exposure on the agricultural module. As the colony is an inhabited settlement, the third objective ensures there is sufficient recreational space for the inhabitants, which contributes to the mental health of those living in the colony. Finally, the fourth objective optimises the location of the landing pad and its relationship to the other modules in the colony.

3.2. SEQUENCE 1

3.2.1. Constructing the Phenotype

Where the first simulation examins the efficient distribution and size of the life support modules, the second simulation brings attention to the inner workings of each module and the relationship between them, primarily from the perspective of transport. This is established through the integration of a hierarchical system of circulation that is driven by the clustering of various units that form each module. The variables (chromosomes) for this second simulation are primarily focused on the number and location of units within each module, which in turn defines how units are clustered and which hierarchical pathway they connect to. Consequently, the length of connection lines establishes the mode of transportation attributed to it (for example, light rail, high speed transit or pedestrian) (Figure 3).

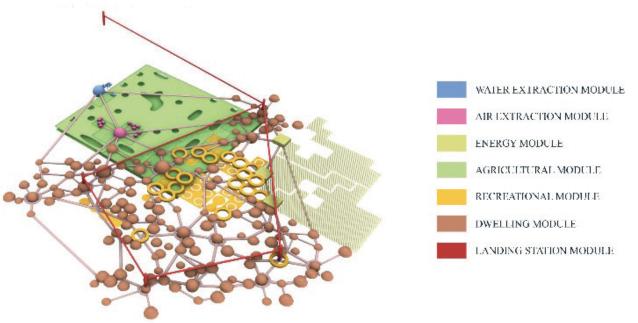


Figure 3. The base phenotype for the second evolutionary sequence

3.2.2. Fitness Objectives

Four main fitness objectives have been identified for the second simulation, focusing on efficiency as well as human comfort. Due to the significance of ensuring efficient connections are established between the modules towards the overall productivity of the colony, the first objective optimises for connections between the dwelling units to resource modules within the colony; using space syntax, density is optimised along connections that achieved a high integration value. The second objective optimises for solar gain on the energy modules to maximise solar energy generation in the colony. Lastly, and considering the potential for events that require immediate closure and evacuation of various modules, the final two fitness objectives optimise for views to key focal points, which in this case were identified as evacuation bunkers and landmarks.

4. Results

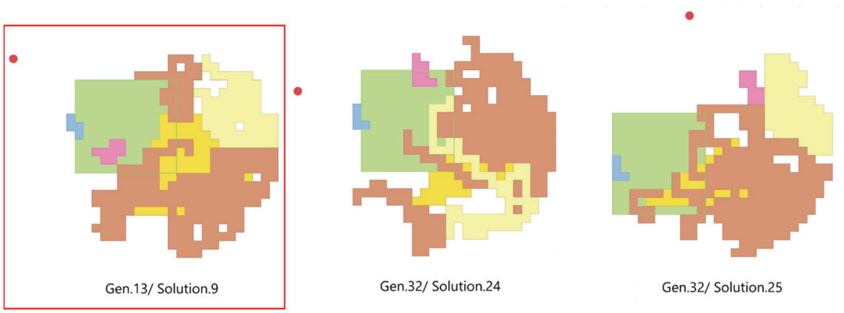
4.1. SEQUENCE 1

The algorithm for the first sequence generated a population size of 5000 solutions, from which 400 solutions formed part of the Pareto front. The Pareto Front solutions are the solutions that are not dominated by any other solution and are considered the best performing solutions to the different fitness objectives driving the algorithm. As the objective is to select a single solution for progression to simulation 2, a multi-tier selection process (Showkatbakhsh & Makki, 2022) is implemented to filter through the 400 Pareto Front solutions and select one. Three selection mechanisms informed the filtration process, the first was to identify solutions that contained an optimal and efficient distribution of dwelling modules, ensuring the 3331-target population is achieved and distributed throughout the 1km site. This first step reduced the Pareto Front to 12 solutions that met the population threshold. The second step in the filtration process evaluated the 12 selected solutions against a set of phenotypic urban indicators (Table 1) that aimed towards achieving a polycentric distribution of dwelling modules and the preference for cross pollination between modules. Each solution was attributed a performance score and the top three performing solutions were selected for further analysis (Figure 4). In the final step, the top three solutions were examined further, through a second round of phenotypic analysis that focused on efficient relationships between adjacent modules (Table 2), leading to the selection of the solution that has performed the best against the defined metrics. This solution formed the phenotype for the second simulation.

	Design goals	Selection criteria	Quantification method	Vector direction	Weighting	Sol 1 (G11.S18)	Sol 2 (G13.S09)	Sol 3 (G16.S35)	Sol 4 (G18.S05)	Sol 5 (G24.S32)	Sol 6 (G32.S24)	Sol 7 (G32.S25)	Sol 8 (G32.S44)	Sol. 9 (G39.S30)	Sol. 10 (G56.S41)	Sol. 11 (G61.S42)	Sol. 12 (G88.S06)
1	Reponses to matter of survival, self-efficiency and constant transportation between Earth and Mars	Distance between Water and Energy blocks	Distance between 2 central points of W and E	+	0.25	0.10	0.21	0.24	0.10	0.12	0.17	0.25	0.13	0.25	0.13	0.15	0.11
2		Proximity distance from water to argicultural block	Percentage of Agricultural block covered by W area	+	0.50	0.17	0.42	0.17	0.00	0.50	0.25	0.17	0.00	0.17	0.17	0.00	0.17
3		Surface coverage of Dwelling units over Agricultural block	Percentage of Agricultural block covered by Dwelling blocks	-	0.25	0.06	0.09	0.04	0.05	0.19	0.10	0.06	0.08	0.06	0.08	0.09	0.25
4		Distance between landing pad and dwelling blocks	Distance between central points of landing pad and dwelling blocks	+	0.25	0.21	0.24	0.22	0.14	0.25	0.23	0.17	0.24	0.21	0.18	0.21	0.21
5		Total distance between landing pad and (W+E)	Distance between central points of landing pad and water + energy	-	0.25	0.18	0.15	0.09	0.16	0.09	0.17	0.17	0.16	0.10	0.08	0.25	0.09
6		Adequate distribution and sufficient area for Public blocks	The number of public points ÷ The number of public clusters	+	1.00	0.47	0.67	0.40	0.53	0.47	0.73	1.00	0.67	0.67	0.60	0.60	0.67
7		Exposure of Agricutural block to sunlight	Percentage of Agricultural block's area receiving sunlight	+	1.00	0.72	0.60	0.78	0.89	0.36	0.56	0.63	0.40	0.50	0.40	0.43	0.56
8	Reduction of urban stress/ radiation island effect and concentration of dwelling clusters	Partial distribution of dwelling blocks	The number of dwelling clusters divided	+	0.50	0.33	0.50	0.33	0.33	0.33	0.42	0.50	0.33	0.42	0.42	0.33	0.33
9	Encouragement of social interaction and mixed land use	Equal integration of public blocks into dwelling	How equally public blocks are spreaded among dwelling	+	1.00	0.24	0.28	0.12	0.20	0.30	0.61	1.00	0.24	0.35	0.28	0.40	0.42
	Total points				**5.00**	2.48	**3.16**	2.39	2.40	2.62	**3.24**	**3.96**	2.26	2.72	2.35	2.47	2.81

Table 1. The analysis and results of the second filtration process, highlighting the phenotypic indicators used in the analysis.

Figure 4. The top 3 selected solutions from the first simulation. The solution selected to progress to



Simulation 2 is highlighted.

		Phenotypic indicators		Quantification method	Vector direction	Weighting	Sol 2 (G13.509)	Sol 6 (G32.524)	Sol 7 (G32.525)
Selection criteria	1	Reduce radiation-island effect	Concentration of energy block	Futhest distance between energy points	-	0.75	0.45	0.075	0.15
	2		Isolation of energy block	Distance of exposing edges of energy blocks to other block types	-	0.25	0.125	0	0.1875
	3	Promote mixed land use	Variation of industial blocks over-crossing agricultural block	Number of industrial blocks covering agricultural	Target: 3	0.50	0.33	0.5	0.165
	4	Obtain balance between efficient partition and compactness of dwelling blocks	Adequate distribution of dwelling blocks	The number of dwelling clusters divided	Target: 2-3	0.50	0.5	0.2	0.5
	5		Equal area distribution among dwelling clusters	Relative subtraction between clusters'area	Target: 0	1.00	0.7	0.4	0.2
		Total points				3.00	2.105	1.175	1.2025

Table 2. The analysis and results of the third filtration process on the top three solutions.

4.2. SEQUENCE 2

Similar to the first simulation, the second simulation generated a population size of 5000 solutions, however only 50 Pareto front solutions were generated, likely a result of the more complex search space and fitness landscape of the design problem. As in sequence 1, simulation 2 implemented a multi-step filtration process to identify the best performing solution from the Pareto Front. Due to the small size of the Pareto Front, a two-step filtration process was selected instead of three. In the first step, the Pareto Front was clustered into 14 clusters, where the centre of each cluster was selected for further analysis (as a representative of that cluster). In addition to the cluster centres, the fittest solution for each fitness objective as well as the solution with the average fitness value were also added to the solution set, resulting in 19 solutions. The second step evaluated the 19 solutions against a varied set of phenotypic indicators that examined each solution's performance against urban metrics such as circulation efficiency, solar gain, and access to views (Table 3). Each solution was awarded a score based on their cumulative performance to the identified indicators, which resulted in a clear standout solution that has performed better than all solutions in the selected pool against the identified metrics.

	Design goals	Selection criteria	Quantification method	Vector direction	Weighting	Sol 1 - F04 (G00.510)	Sol 2 - C1 (G35.544)	Sol 3 - C2 (G44.512)	Sol 4 - C3 (G49.547)	Sol 5 - C4 (G50.520)	Sol 6 - F02 (G58.503)	Sol 7 - C5 (G65.524)	Sol 8 - C6 (G71.537)	Sol 9 - C7 (G73.548)	Sol 10 - C8 (G74.502)	Sol 11 - C9 (G82.546)	Sol 12 - C10 (G94.541)	Sol 13 - C11 (G95.524)	Sol 14 - F01 (G100.503)	Sol 15 - C12 (G107.538)	Sol 16 - C13 (G110.527)	Sol 17 - Ave (G111.543)	Sol 18 - F03 (G118.504)	Sol 19 - C14 (G118.545)
1	Density and efficient distribution of qualified place-making to achieve context-driven objectives	Density of dwelling units in the vinicity of public units	Average number of dwelling units per public unit	+	1.00	0.89	0.92	0.75	0.90	0.97	0.79	0.81	0.84	1.00	0.79	0.84	0.76	0.77	0.88	0.85	0.96	0.88	0.82	0.88
2		Average public/ private ratio within a dwelling cluster	Ratio between public space area over private area within a dwelling block (average)	Target: 35%	0.50	0.36	0.46	0.50	0.47	0.43	0.41	0.43	0.43	0.41	0.41	0.40	0.46	0.39	0.47	0.36	0.47	0.41	0.43	0.36
3	Solar gain	Solar gain on energy, public and residential blocks	Total maximum solar gain area on public, residential and energy blocks	+	0.25	0.20	0.23	0.24	0.22	0.23	0.22	0.25	0.24	0.21	0.22	0.22	0.24	0.22	0.25	0.23	0.22	0.21	0.20	0.20
4	View	Views from dwelling units to public units (emergency bunker)	Total view rays from dwelling units to public units	+	0.75	0.34	0.55	0.39	0.75	0.36	0.53	0.43	0.62	0.68	0.53	0.52	0.54	0.55	0.60	0.66	0.56	0.60	0.71	0.35
5		Views from dwelling units to surrounding landmarks	Avarage views to at least 2 landmarks from dwelling units	+	0.25	0.22	0.24	0.24	0.24	0.24	0.25	0.24	0.25	0.24	0.25	0.23	0.24	0.24	0.24	0.25	0.23	0.25	0.24	0.22
6	Pedestrian distance	Walking distance between 2 furthest points	Distance of the longest path in the module	-	0.5	0.48	0.49	0.49	0.50	0.48	0.48	0.49	0.46	0.50	0.48	0.49	0.48	0.49	0.48	0.46	0.49	0.50	0.50	0.48
7	Effective circulation	Transportation convenience	Number of dwelling units not being covered by 500m radius of secondary circulation	-	0.50	0.25	0.03	0.10	0.17	0.03	0.02	0.02	0.17	0.25	0.02	0.13	0.25	0.08	0.01	0.04	0.50	0.02	0.50	0.25
8		Efficient travelling distance of secondary connectors	Total length of secondary connectors	+	0.25	0.24	0.19	0.22	0.21	0.22	0.11	0.19	0.22	0.24	0.11	0.20	0.22	0.22	0.14	0.13	0.25	0.12	0.22	0.24
				Total point	4.00	2.99	3.12	2.94	3.46	2.96	2.81	2.85	3.23	3.53	2.81	3.03	3.18	2.96	3.07	2.98	3.69	2.97	3.61	2.98

Table 3. Phenotypic analysis and filtration of the simulation's results with the selected solution.

The selected solution demonstrates five key clusters that serve as distribution points, generated through circulation pathways that favoured connections between dwelling units to public amenities. The distribution points ensure that every dwelling unit is within a serving radius of 400 metres, enabling functional efficiency and maximum capacity of each cluster. Additionally, a secondary circulation is identified that allows pedestrian movement between any point of the colony within a 15-minute timeframe (Figure 5). Most importantly, the final selected solution demonstrates a decentralised and contained layout of the original 5 modules, distributed through critical relationships derived through processes of efficiency and urban sustainability (Figure 6). The proposed Master plan of Martian settlement is as shown in Figure 7.

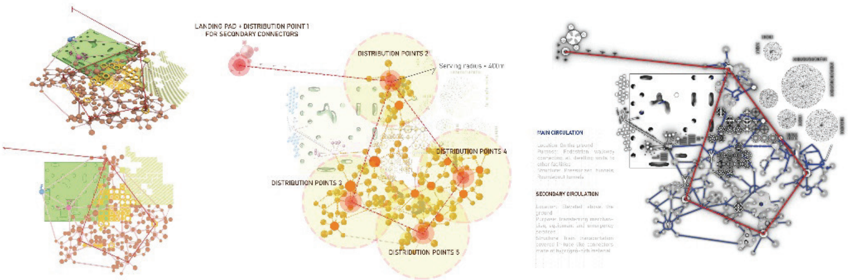


Figure 5, 6: Analysis of selected solution: Zoning diagram, Distribution clusters & Circulation

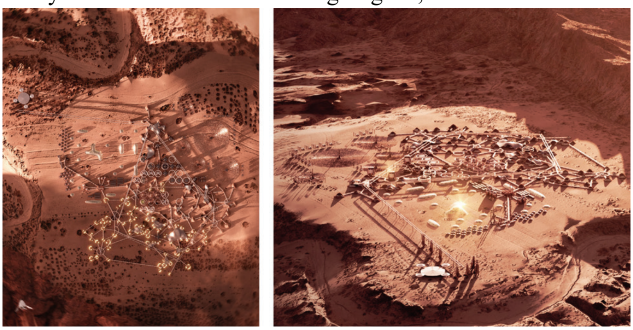


Figure 7: Master plan of the Martian settlement

5. Conclusions

Through the use of evolutionary computation methods, the research examines the potential for, and challenges to, the urban colonisation of Mars as well as emphasises the efficiency, internal relationships and sustainability of the proposed urban framework that addresses existing contextual conditions and distinctive relationships between various life-support and foundational living blocks within the Martian habitat. The research also demonstrates the advantages of generating the urban framework for the colony being dependent on ensuring life supported for a 3331-person habitat; while the second simulation brings the focus on the urban qualities that hold the greatest impact on the inhabitants' quality of life. This two-tiered approach allows for different teams with different expertise the opportunity to hold greater weight as stakeholders for the project as a whole. The presented model, although specific to Valles Marineris, can be adapted to alternative locations with similar characteristics on Mars by generating a distinct relationship between variables (chromosomes), design goals (fitness objectives) and morphological characteristics (phenotypes).

Interplanetary colonisation is an extremely new and young field of study, one that is evolving from exploratory missions to the establishment of sustainable and long-term habitations; a field that requires significant urban direction. Future work is needed to continue to address challenges of interplanetary colonisation, such as the need to address the impact of a lower gravitational pull on long-term human settlement, material studies, geological conditions and climate-resilient structure experiments. Although the data used to drive the presented study may be revised and could vary depending on future studies, the general framework and process would play a critical role in establishing urban settlements that are primarily dependent on efficiency and optimisation of modular components.

References

Catling, D. C., & Kasting, J. F. (2017). *Atmospheric evolution on inhabited and lifeless worlds.* Cambridge University Press.

European Space Agency. (2021). *ExoMars discovers hidden water in Mars' Grand Canyon*. European Space Agency. https://www.esa.int/Science_Exploration/Human_and_Robotic_Exploration/Exploration/ExoMars/ExoMars_discovers_hidden_water_in_Mars_Grand_Canyon

Haberle, R. M., Clancy, R. T., Forget, F., Smith, M. D., & Zurek, R. W. (2017). *The Atmosphere and Climate of Mars.* Cambridge University Press.

Hargitai, H. (2010). *Mars Climate Zone Map Based On TES Data*. 41st Annual Lunar and Planetary Science Conference, 1533, 1199.

Leighton, R. B., & Murray, B. C. (1966). *Behavior of Carbon Dioxide and Other Volatiles on Mars: A thermal model of the Martian surface suggests that Mars's polar caps are solid carbon dioxide.* Science, 153(3732), 136–144, DOI: 10.1126/science.153.3732.136.

Makki, M., Showkatbakhsh, M., & Song, Y. (2018). *Wallacei: An evolutionary and Analytic Engine for Grasshopper 3D*. Wallacei. https://www.wallacei.com/

Moses, R. W., & Bushnell, D. M. (2016). *Frontier in-situ resource utilization for enabling sustained human presence on mars*.

Nasa. (2021). *Timeline of Mars Exploration*. NASA Mars Exploration. https://mars.nasa.gov/mars-exploration/timeline

National Space Society. (2017, August 14). *NSS Space Settlement Nexus*. NSS Space Settlement Nexus - National Space Society. https://space.nss.org/space-settlement-national-space-society/

Petranek, S. (2015). *How We'll Live on Mars*. Simon and Schuster.

Reid, C. (2019). *A Hypothetical Exploration of Survival, Colonisation and Interplanetary Relations Around the planet Mars* [Master Thesis, Faculty of Humanities]. https://open.uct.ac.za/handle/11427/30530

Showkatbakhsh, M., & Makki, M. (2022). *Multi-Objective Optimisation of Urban Form: A Framework for Selecting the Optimal Solution*. Buildings, 12(9), Article 9. https://doi.org/10.3390/buildings12091473

Szocik, K., Wójtowicz, T., & Braddock, M. (2020). *The martian: Possible scenarios for a future human society on mars*. Space Policy, 54, 101388, https://doi.org/10.1016/j.spacepol.2020.101388.

Zubrin, R. (2014). *Colonising the Red Planet: Humans to Mars in Our Time*. Architectural Design, 84(6), 46–53. https://doi.org/10.1002/ad.1832

Zubrin, R., & Wagner, R. (2011). *The Case for Mars*. Simon and Schuster.

A SURROGATE-ASSISTED OPTIMIZATION APPROACH TO IMPROVE THERMAL COMFORT AND ENERGY EFFICIENCY OF SPORTS HALLS IN SUBTROPICAL CLIMATES

YU LI1, LINGLING LI[2] and NAIHUA YUE[2]
[1,2]School of Architecture, Harbin Institute of Technology; Key Laboratory of Cold Region Urban and Rural Human Settlement Environment Science and Technology, Ministry of Industry and Information Technology
[3]School of Architecture and Urban Planning, Qingdao University of Technology
[1]19b334003@hit.edu.cn, 0000-0002-7670-3998

Abstract. Balancing the thermal comfort and energy efficiency has been recognized as a critical issue in sports hall design, which is yet to be properly implemented in early design stages due to the huge computational cost and delayed simulation feedback. This paper develops an accelerated optimization approach for thermal comfort and energy efficiency of sports halls by combining surrogate modelling with evolutionary algorithms. An integrated computational workflow designated for early-stage application was established that consists of Design of experiments, Surrogate modelling, Surrogate-assisted multi-objective optimization, and Multi-criteria decision making. Specifically, a parametric sports hall model was set up for batch physics-based simulations to generate abundant training samples, which was then utilized to develop surrogate models for the rapid prediction of thermal comfort and energy efficiency. The validated surrogate models were eventually linked with evolutionary algorithms to quickly identify the optimal design solution(s). The performance of the developed approach was evaluated against the traditional simulation-based optimization (SBMOO) method. Results indicated that the proposed approach could save 70.91% of total computational time for this case study, whilst achieving better optimized thermal comfort and energy efficiency with a reduction of mean PMV and site EUI by 0.001 and 1.60 kWh/m2/yr versus the SBMOO method.

Keywords. Thermal comfort, Energy efficiency, Multi-objective Optimization, Surrogate Model, Sports Hall.

1. Introduction

Thermal comfort is regarded as an essential environmental factor of designing sports halls in subtropical climates. For a long time, mechanical ventilation has become a

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 301-310.

dominant way to adjust the indoor thermal environment of sports halls in subtropical climates, which accounts for a significant proportion of building energy consumption (Yue, Li, Morandi, & Zhao, 2021). Trade-off between thermal comfort and energy efficiency remains a challenging task at early-stage sports hall design because of the contradictory nature of the two. Generally, to improve occupant thermal comfort, more energy consumption is expected from air conditioning (AC) systems. Thus, how to take advantage of passive measures to improve building energy efficiency without scarifying occupant thermal comfort need to be explored in the design of sports halls in subtropical climates.

Vast studies have been conducted to facilitate the energy saving of large sports halls applying active and passive strategies, the former one focuses on improving the AC systems, smart control systems, and so on, while the later one concentrates more on optimizing building form, envelopes, apertures and materials, which has been verified as an effective way at early-design stages (Yue, Li, Morandi, & Zhao, 2021). As a common passive cooling strategy in subtropical climates (e.g. Guan Zhou, Shen Zhen, China), natural ventilation has been extensively utilized in sports hall design practices, which shows great climatic potentials from January to May and September to December, accounting for nearly 75% of annual time span (Guo et al., 2022). Besides, natural ventilation can effectively remove thermal stress from indoor spaces thus saving the cooling energy use which is a large building energy consumer in subtropical climates. Nevertheless, satisfactory natural ventilation is difficult to realize in large-scale sports halls due to the thermal stratification phenomenon and low ventilation rate (Guo et al., 2022; Wang et al., 2021).

Simulation-based multi-objective optimization (SBMOO) method has been utilized to identify optimal building forms, materials, ventilation strategies, etc. for achieving the best thermal comfort and energy efficiency of large sports halls, unfortunately, most of which have not been fully integrated to early design stages due to the high computational cost and delayed simulation feedback (Yue, Li, Morandi, & Zhao, 2021; Guo et al., 2022). Another shortcoming comes from the fact that the simulated optimal solutions cannot be adjusted to suit practical implementation and meanwhile, fail to explain why a design alternative is good or not. Therefore, there still requires a flexible and efficient optimization method designated for early-stage decision-making to support the design of thermal comfortable and energy-efficient sports halls in subtropical climates, the new method should not only address the expensive computational load, but also available to provide designers with the direction to advance the design process and help them find out the most influential design factors.

2. Methodology

This study develops a surrogate-assisted multi-objective optimization (SA-MOO) method that can quickly optimize the thermal comfort and energy efficiency of sports halls at early design stages. An integrated computational workflow coupling Design of experiments (DoE), Surrogate modelling (SM), Surrogate-assisted multi-objective optimization (SA-MOO) and Multi-criteria decision making (MCDM) was established using Grasshopper-modeFRONTIER (GH-MF) platform, as depicted in Figure 1. The proposed approach was applied to a pilot study using a parametric sports hall model in Shen Zhen, China, the performance of which was compared with traditional SBMOO method in terms of optimization process, optimal solution(s) quality and computational efficiency.

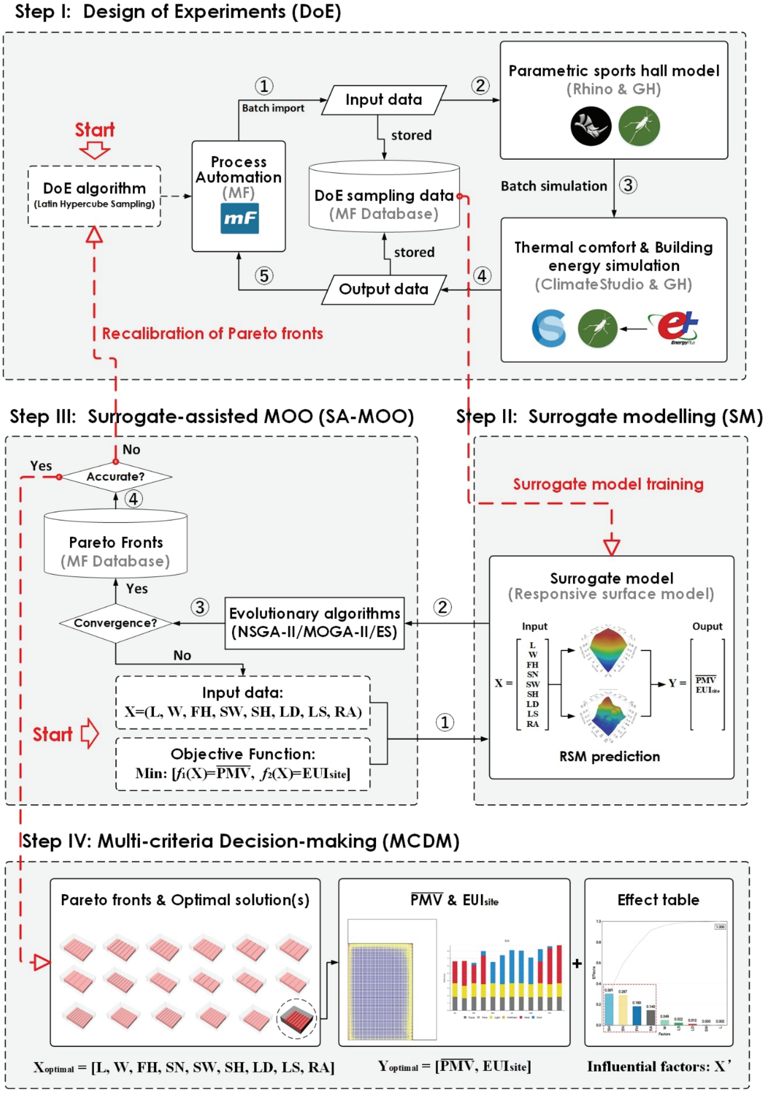


Figure 1. Methodological framework of the SA-MOO method.

2.1. DESIGN OF EXPERIMENTS

A parametric sport hall model was set up to assess the thermal comfort and energy efficiency of sports halls, which consists of 9 design parameters (inputs) controlling the building massing, skylights and shadings (Table 1), the second storey was selected as the target area in this study (Figure 2(a)). Then, a physical simulation model (Figure 2(b)) was established in Rhino-Grasshopper-ClimateStudio (RH-GH-CS) platform for calculating the thermal comfort and building energy efficiency, i.e., predicted mean votes (PMV), site energy use intensity (EUI_{site}). Afterwards, Design of experiments (DoE) was conducted to collect abundant and representative design samples (Input/output combinations) for training surrogate models, a widely-used sampling algorithm, Uniform Latin Hypercube (ULH) was utilized to generate totally 600 design variable combinations given the computational cost and surrogate model quality, based on which, batch physical simulations were performed using the RH-GH-CS platform called by modeFRONTIER (MF), referred to (Yang, Di Stefano, Turrin, Sariyildiz, & Sun, 2020; Giouri, Tenpierik, & Turrin, 2020), the main difference is that the proposed workflow utilizes parallel simulation of thermal comfort and energy efficiency which can quickly generate the required design samples.

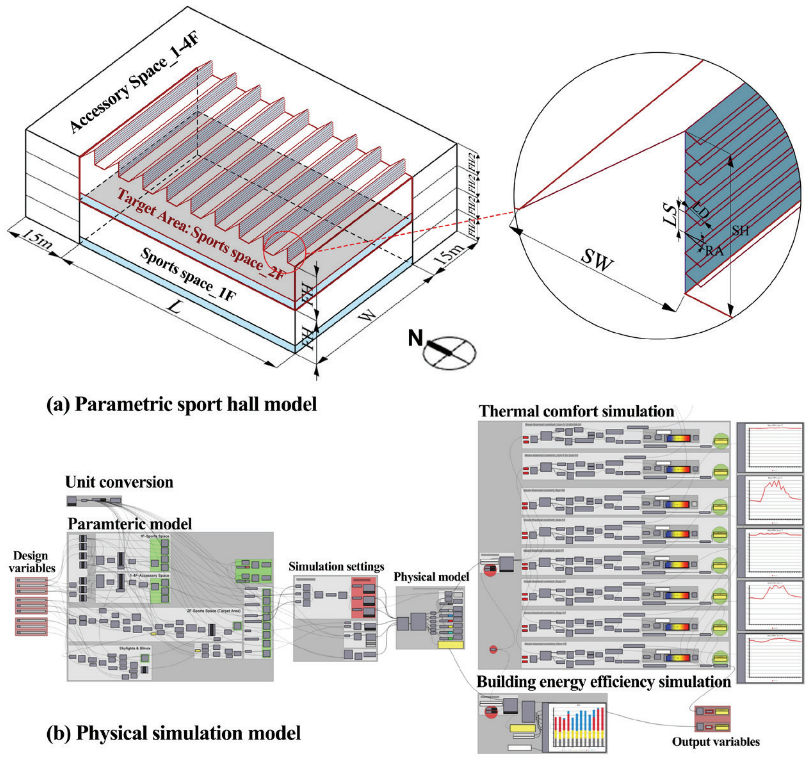


Figure 2. (a) Parametric sports hall model; (b) Physical simulation model.

Table 1. Description of design variables of the parametric sports hall.

	Design variable	Range	Step	Unit
Building Massing	Sports Space_Length (L)	[50, 70]	1	(m)
	Sports Space_Width (W)	[30, 50]	1	(m)

Table 1(continued)

	Sports Space_Floor Height (FH)	[9.0, 15.0]	0.1	(m)
Skylights	Skylights_Number (SN)	[3, 12]	1	-
	Skylights_Width (SW)	[1.0, 4.0]	0.1	(m)
	Skylights Height (SH)	[1.0, 4.0]	0.1	(m)
Shadings	Shades_Louver Depth (LD)	[0.20, 0.50]	0.01	(m)
	Shades_Louver Spacing (LS)	[0.20, 0.50]	0.01	(m)
	Shades_Rotation Angle (RA)	[-45, 45]	5	(°)

Note: West and south windows of the sports spaces are 2.5m high.

In thermal comfort simulation, we selected 6 seasonal representative days (Spring: Apr-16; Summer: Jun-12, July-17, Aug-17; Autumn: Sep-17 and Nov-15) (Guo et al., 2022) to conduct spatial thermal comfort calculations with an analysis grid spacing of 2m in a height of 1.5m, the hourly PMV values (0:00-24:00) in the whole target area were obtained and finally the mean PMV of those analysed days across the target area ($\overline{PMV}$) was defined to evaluate the overall thermal comfort, calculated as:

$$\overline{PMV} = \frac{1}{N}\sum_{n=1}^{N}\left(\frac{1}{6}\sum_{d=1}^{6}\left(\frac{1}{24}\sum_{h=0}^{23} PMV_{n,d,h}\right)\right)$$

where: n is the specific sensor point in the target area, $n \in [1, \mathrm{N}]$; d is the d^{th} representative day, $d \in [1, 6]$; and h is the particular hour of a day, $n \in [0, 23]$.

In building energy simulation, 6 thermal zones were set up, including four-story accessory spaces and two-story sports spaces. The target area (sports space_2F) was selected to calculate EUI_{site} in this study with no mechanical ventilation.

2.2. SURROGATE MODELLING

Response surface methodology (RSM) model is one of the widely used surrogate models in building performance prediction due to its fast approximation speed and validated accuracy, which can use limited training samples to map the complex function relationship between design variables and response variables (Myers, Montgomery, & Anderson-Cook, 2016). Thus, RSM model was adopted to quickly approximate the thermal comfort and energy efficiency in this study instead of physics-based simulations, the input variables are aforementioned design variables, $X = (L, W, FH, SN, SW, SH, LD, LS, RA)$ and the output variables are $f_1(\mathrm{X}) = \overline{PMV}$ and $f_2(\mathrm{X}) = EUI_{site}$. To achieve the best fitted RSM, 17 different training algorithms were adopted to train the independent RSM for $\overline{\mathrm{PMV}}$ and EUI_{site} using the DoE sampling data, with a proportion of training and validation data of 80% and 20%, from which the RSM models with the best prediction accuracy were identified and used in later SA-MOO.

2.3. SURROGATE-ASSISTED MULTI-OBJECTIVE OPTIMIZATION

The surrogate-assisted multi-objective optimization combined with evolutionary algorithms (i.e., Non-dominated Sorting Genetic Algorithm (NSGA-II), Multi-

Objective Genetic Algorithm II (MOGA-II) and Evolution strategy (ES)) were applied for the fast optimization of thermal comfort and energy efficiency using modeFRONTIER software (RSM node), the MF workflow is shown in Figure 3, where parallel prediction of $\overline{PMV}$ and EUI_{site} was enabled by Synchronizer node in MF. The objective function is to improve the thermal comfort and energy efficiency of the target area of the sport hall, i.e. to minimize the $\overline{PMV}$ and EUI_{site} values, defined as:

$$Minimize: \quad F(X) = [f_1(X), f_2(X)],$$

$$X = (L, W, FH, SN, SW, SH, LD, LS, RA) \; and \; X \in S,$$

where: $f_1(X), f_2(X)$ denote the optimization function of $min\_\overline{PMV}$, min_EUI_{site}, $X = (L, W, FH, SN, SW, SH, LD, LS, RA)$ corresponds to 9 independent design variables; S is the entire search space.

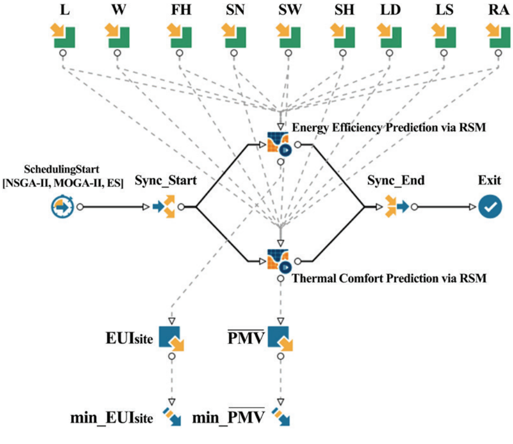


Figure 3. Surrogate-assisted MOO workflow in MF.

2.4. MULTI-CRITERIA DECISION MAKING

To identify the optimal solution(s) form the Pareto front, a genetic-algorithm based multi-criteria decision making (GA-MCDM) method was adopted, which calculates the utility function with a genetic algorithm taking into account the weight and alpha values of two optimization objectives in this study. The GA-MCDM can rank all available Pareto solutions depending on the calculated utility function and thus identifying the best ones as the optimal solution(s).

3. Results and discussion

3.1. PERFORMANCE OF SURROGATE MODELS AND RE-CALIBRATION RESULTS

Table 2 depicts the performance of the best fitted RSM model for $\overline{PMV}$ and EUI_{site}, respectively, which indicates good prediction performance in terms of low MAE, MRE and MNE. By combining the best fitted surrogate models with three evolutionary algorithms (NSGA-II, MOGA-II and ES), the SA-MOOs were conducted and 2420, 1815 and 19740 alternative design solutions were generated, among which 41, 35 and 128 Pareto solutions were identified (Figure 4(a)). To recalibrate the RSM prediction accuracy, those generated Pareto solutions were re-calculated using physics-based simulation and were compared with the RSM prediction data (Figure 4(b-d)). Results suggested that the optimization process using ES algorithm achieved the Pareto solutions with the best prediction accuracy and were chosen in this study.

Table 2. Performance of the best RSM model of $\overline{PMV}$ and EUI_{site}.

RSM model & Training algorithm	Mean absolute error (MAE)	Mean relative error (MRE)	Mean normalized error (MNE)
$EUI_{site}\_H2O\_MLP$	2.1450	0.0238	0.0175
$\overline{PMV}\_H2O\_GBM$	0.0088	0.0052	0.0587

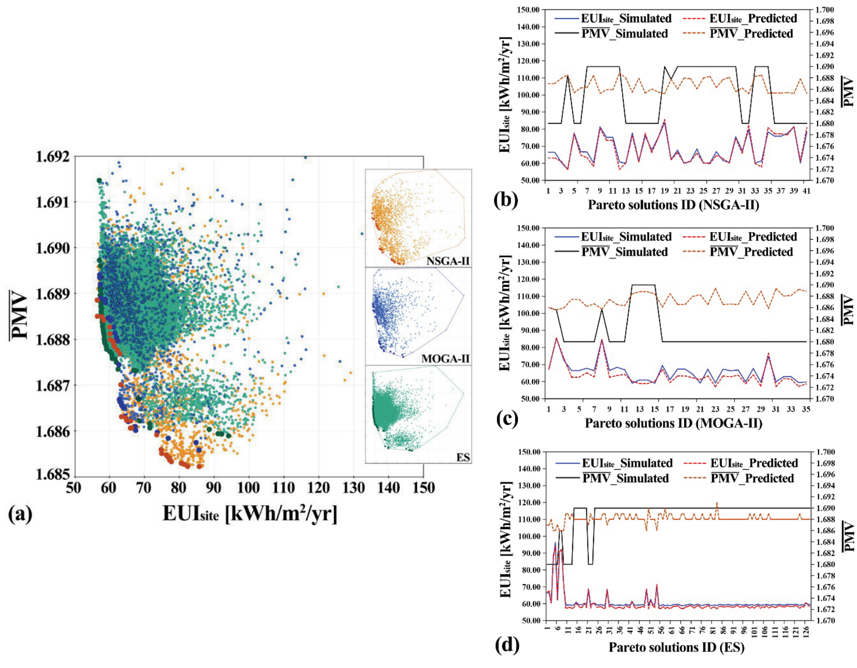


Figure 4. (a) Pareto solutions and all alternative solutions from SA-MOO using NSGA-II, MOGA-II and ES optimization algorithm; (b-d) Re-calibration of the generated Pareto solutions.

3.2. FINAL OPTIMAL SOLUTION(S)

By determining those Pareto solutions achieved by ES optimization algorithm as the final ones, we further conducted the GA-MCDM on them using the simulated data, rankings of all simulated Pareto solutions are shown in Figure 5(a-b), where it could be noted that ID = {16738, 15750, 15630, 15294, 15694, 16829, 8793, 13056, 4945, 3937} valued the highest and were identified as the final optimal solutions. Fig 5(c) depicts the influential design parameters for $\overline{PMV}$ and EUI_{site}, it can be seen that LS and W has a significant effect on $\overline{PMV}$, while for EUI_{site}, the top four influential design parameters are SN, SH, W and FH, the findings can support designers to better understand where and how to revise the design alternatives and thus achieving better optimal design solutions.

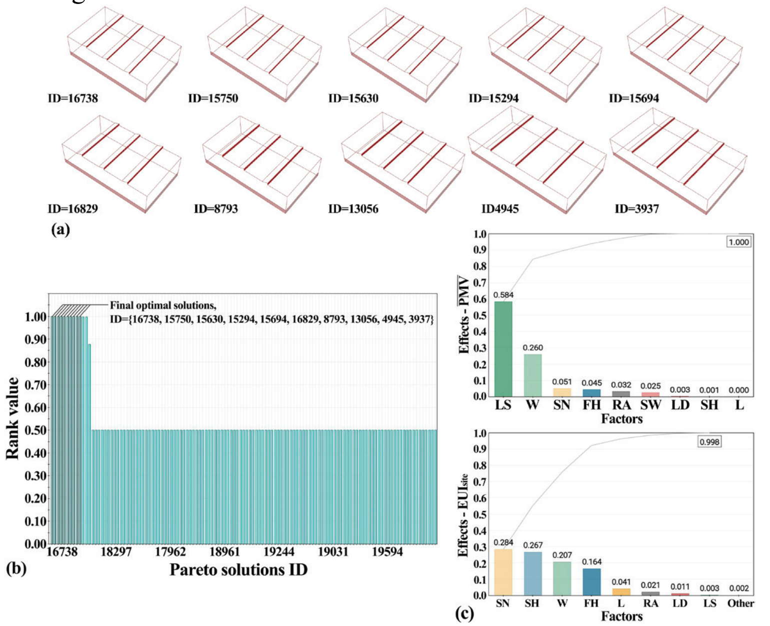
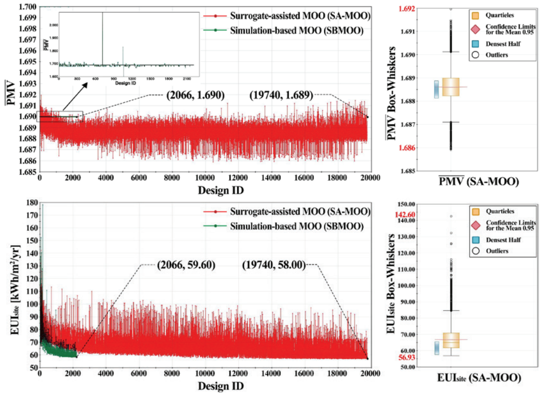


Figure 5. (a) Ranking of simulated Pareto solutions; (b) Optimal solutions; (c) Influential design variables.

3.3. DISCUSSION ON THE PROPOSED APPROACH

Optimization process, optimal solution quality and computational efficiency of the proposed approach were compared against the traditional SBMOO method. Figure 6 indicates that both SA-MOO and SBMOO achieved reasonable convergence towards better thermal comfort and energy efficiency but at different generations. For $\overline{PMV}$, the SBMOO method converged at the ID=2066 ($\overline{PMV} = 1.690$), while SA-MOO at the ID=19740 ($\overline{PMV} = 1.689$); For EUI_{site}, the SBMOO converged with a EUI_{site} of 59.60 $kWh/m^2/yr$, while SA-MOO achieved a lower EUI_{site} of 58.00 $kWh/m^2/yr$. In general, SA-MOO outperformed SBMOO in achieving better-performing design solutions, and meanwhile, the wider solution space, allowing designers to freely make decisions according to actual needs at early design stages.

Figure 7 shows the computational time of SA-MOO and SBMOO method, the calculation time for thermal comfort and building energy simulation of a single sample was measured to be 86.18s, totally 600 samples were run to obtain the DoE sampling dataset, taking about 51708s. For SA-MOO, the surrogate model training time was 3600s, the surrogate-assisted MOO time was about 5s, Thus, the total computational time was about 55313s. The total computational time of SBMOO can be calculated as: $y = 86.18 * x$ (x represennts the number of optimized samples). It can be indicated from Figure 7 that when the number of optimized samples is less than 642, the SBMOO method has a time advantage versus SA-MOO because there is no need for training surrogate models, while when it exceeds 642, the gap of total computational time between the SA-MOO and SBMOO becomes larger with the increase of optimized samples because the trained surrogate modes can gradually save total computational time. For this case study (the number of optimized samples was 2206), the SA-MOO outperformed the SBMOO in computational efficiency, suggesting a 70.91% time saving which is quite considerable at early design stages. In this study, all simulations were ran using a common computer with Intel (R) Core (TM) i7-6700 CPU @ 3.40 GHz, 8GB RAM and NVIDIA (R) GeForce GTX 1060, 6GB graphics accelerator.



. Figure 6. Optimization process of $\overline{PMV}$ and EUI_{site} by the SBMOO and SA-MOO approach.

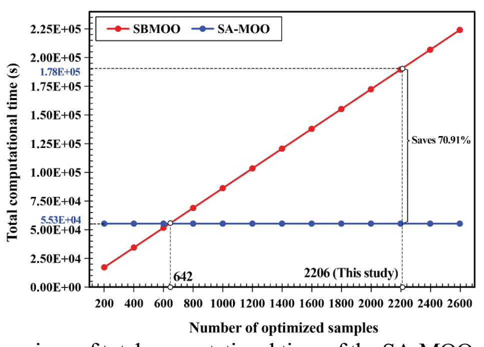


Figure 7. Comparison of total computational time of the SA-MOO and SBMOO approach.

4. Conclusions

This paper presents an accelerated optimization approach for thermal comfort and energy efficiency of sports halls by combining surrogate modelling with evolutionary algorithms. The novelty of this study resides in the adoption of surrogate models for rapid prediction of building performance, which is combined with evolutionary algorithms to identify the optimal design solutions. A pilot study using the proposed SA-MOO approach and the traditional SBMOO method was conducted and compared. Case study results verified the superiorities of the SA-MOO and its potential application at early design stages.

There exist some limitations of the SA-MOO approach. First, the training dataset is non-reusable, meaning that the surrogate models need to be trained every time in different cases and scenarios, though some studies have tried to automatically develop the surrogate models for different application scenarios (climate zones, building types, simulation periods, etc.), the computational cost still remains high, which is far away from acceptable in practical projects. Second, differentiated optimization efficacies have been observed between thermal comfort and energy efficiency, indicating that for cases with many optimization objectives, there may exist trade-offs of optimization algorithms for different optimization objectives that should be carefully addressed; Last, this study utilized mean PMV value of representative calendar days to evaluate the thermal comfort of different seasons, which may exist risks in assessing the overall thermal comfort, especially when the analysis time-span is across a whole year. Besides, the predicted value from surrogate models fails to reflect the spatial distribution of thermal comfort, hence, a more comprehensive metric of thermal comfort in sports spaces should be utilized in future studies.

References

Giouri, E. D., Tenpierik, M., & Turrin, M. (2020). Zero energy potential of a high-rise office building in a Mediterranean climate: Using multi-objective optimization to understand the impact of design decisions towards zero-energy high-rise buildings. Energy and Buildings, 209, [109666]. https://doi.org/10.1016/j.enbuild.2019.109666

Guo, W., Liang, S., He, Y., Li, W., Xiong, B., & Wen, H. (2022). Combining EnergyPlus and CFD to predict and optimize the passive ventilation mode of medium-sized gymnasium in subtropical regions. *Building and Environment*, 207, 108420, https://doi.org/10.1016/j.buildenv.2022.109951.

Myers, R. H., Montgomery, D. C., & Anderson-Cook, C. M. (2016). *Response surface methodology: process and product optimization using designed experiments*: John Wiley & Sons.

Wang, X., Yang, Y., Xu, Y., Wang, F., Zhang, Q., Huang, C., & Shi, C. (2021). Prediction of vertical thermal stratification of large space buildings based on Block-Gebhart model: Case studies of three typical hybrid ventilation scenarios. *Journal of Building Engineering*, 41, 102452. DOI:10.1016/j.jobe.2021.102452

Yang, D., Di Stefano, D., Turrin, M., Sariyildiz, S., & Sun, Y. (2020). Dynamic and interactive re-formulation of multi-objective optimization problems for conceptual architectural design exploration. Automation in Construction, 118, [103251]. https://doi.org/10.1016/j.autcon.2020.103251

Yue, N., Li, L., Morandi, A., & Zhao, Y. (2021). A metamodel-based multi-objective optimization method to balance thermal comfort and energy efficiency in a campus gymnasium. *Energy and Buildings*, 253, 111513. DOI:10.1016/j.enbuild.2021.111513

STYLIZED SPACE SYNTHESIS

Exploring The Stylized Generative Design Method Of Architecture Based On Wave Function Collapse Algorithm

WANZHU JIANG[1]and JIAQI WANG[2]
[1,2] *South China University of Technology.*
[1] *ucbqwj0@ucl.ac.uk, 0000-0002-3450-670X*
[2]*ucbq121@ucl.ac.uk, 0000-0002-3242-7709*

Abstract. It has been a frequent task and challenge for architects to translate and transfer a specific style so that the design works can fit into a particular built environment or a unique period. The wave function collapse algorithm is an image generator inspired by constraint solving, which generates numerous images with similar styles by analyzing the potential connections of discrete segments in instances. This paper explores the application of the wave function collapse algorithm in the generation of stylized architecture. By deconstructing architectural style templates into predefined spatial tiles and connection rules, this research models style expression as a constraint-solving process, establishing a stylized spatial synthesis algorithm with discrete design logic to generate self-similar aggregations, shaping architecture as a unique semantic system. Based on the generative experiments of cultural architecture in the traditional Chinese style, this method was tested in two stages. While demonstrating a complete workflow, it has been fully verified for the feasibility, creativity and adaptability in stylized synthesis problems.

Keywords. Stylized Synthesis, Discrete Aggregation, Wave Function Collapse, Spatial Module, Constraint Solving, Generative Design

1. Introduction

Style is connotated with particular, identifiable features, positioning buildings in a specific time-frame and culture (del Campo and Manninger, 2020). Architectural style includes not only explicit external forms and detailed patterns, but also implicit functional relationships and spatial configurations, reflecting complex artistic precepts, cultural sympathy, and social concepts. For a long time, architects often have to face the challenge of preserving, protecting, restoring, tracing, translating, transferring or expressing a specific style to integrate projects into a particular built environment or commemorate a unique background of the times. This kind of work relies on sensibility and experience without systematic workflow. With the update of digital design methods, how to achieve the stylized generation of architecture in an efficient, convenient, clear and complete way has become an essential research topic.

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 311-320.

The earliest exploration of stylized generation can be traced back to the 1970s. Shape grammar was used in the space generation of Palladian (Stiny and Mitchel, 1978) and Wright's styles (Koning and Eizenberg, 1981), where the style was converted as procedural shape growth. Its structural rules can make the outcomes overfit the target style, resulting in the limitation of scales and forms. In addition, deep learning algorithms such as GAN have recently been used for the style transfer of 2D images and stylized generation of architectural plans (Huang and Zheng, 2018). However, architectural spaces differ from blendable pixels; the limited data dimension, non-vectorized output and the uninterpretable computing model are still barriers preventing it from establishing a mature design workflow.

As a new data-driven generative design method, the wave function collapse algorithm originated from the research on procedural content generation (PCG) in the computer field, using the atypical approach of constraint solving to complete automatic constructions of game scenes. Its characteristics of local similarity, incremental creation, minimal entropy heuristic, and preserving semantics (Karth and Smith, 2017) provide innovative ideas for the stylized architectural generation.

This paper explores the application of the wave function collapse algorithm in the generative design of stylized architecture. This research models style expression as a constraint-solving process, deconstructing architectural style templates into predefined spatial tiles and their local connection rules to establish a bottom-up discrete design logic. Based on the experiments of cultural buildings in the traditional Chinese style, this paper demonstrates a complete workflow of the method and fully verifies it for the feasibility, creativity and adaptability in stylized space synthesis.

2. Literature review

Constraint Satisfaction Problems (CSP) (Karth and Smith, 2017), defined in terms of decision variables and values, are widely applied in modeling and solving computational logic problems such as space allocating, decision making, puzzle reasoning, etc. The architectural style has long been understood as a system as a whole rather than a generating system (Christopher, 1968). The traditional design methods require the architect to focus on the holistic property in a top-down way, ignoring the interactive relationship between the underlying parts that produce style features. In fact, the generation of featured architectural systems often involves a complex arrangement and combination of various functional and form elements according to specific rules, depending on a large number of trial-and-error searches, which is difficult to achieve only by experience or inspiration. Digital technology can encode architectural design as an objective-oriented CSP problem, interpreting it as a system of parts and rules to construct a spatial synthesis logic of "deconstructing-encoding-aggregating".

Wave function collapse is derived from quantum mechanical theory and refers to the process by which a microscopic system with uncertainty is reduced from a superposition of several eigenstates to a single eigenstate due to interaction with the external world. "Entropy" measures the unpredictability of the information in this process, which reduces as the remaining probability decreases. The Wave Function Collapse (WFC) algorithm uses the above principles to solve the constraint satisfaction problem by progressively checking the possibilities through the mechanisms of local search, constraint propagation, and backtracking optimization.

The WFC algorithm was developed by game developer Maxim Gumin and inspired by Merrell's Model Synthesis (Gumin, 2017). It was initially an example-driven image generation method, which can capture local fragments and relationships from source materials and recreate similar distribution of features. Recently, The algorithm has been increasingly used as a crucial avenue for discrete architecture research. These studies focus on spatial generation in three dimensions, where the subjects vary in scale from building components to modular units, and the connection constraints are set to geometric or topological relationships. Different settings yield rich results, demonstrating the powerful spatial creativity of the algorithm. Specifically, the current research can be broadly classified into two categories according to the different operation modes of the WFC algorithm.

The Tiling Wave Function Collapse Model (TWFC): It is the model used in most studies, with no target case input, requiring the manual setup of the tiles library and constraints. Based on this model, Athiana Athina defined a discrete modular architecture system consisting of building components in multiple scales that can be constructed and reconfigured adaptively (Hosmer et al., 2020). Youyu Lu sets tiles as functional units (represented by standard cubes) in kindergarten and translates constraints such as light requirements to generate spatial layouts (Lu and Tong, 2020). Its advantage lies in controllable operation steps and generation outcomes, but the lack of target guidance weakens the significance of tile semantics. The workflow of tile design and constraint formulation is relatively abstract and complicated, placing high demands on the designers.

The Overlapping Wave Function Collapse Model (OWFC): Instead of manually setting tiles, Immanuel Koh's proposed voxel synthesis algorithm explores automatic sampling directly from the input examples to synthesize a variety of locally similar outputs (Koh, 2022). This model alleviates the tediousness of manually defining parts and marking rules, but excessive subdivision may lead to scattered spaces. Different from 2D images, after the algorithm disassembles 3D models, the change of voxel positions may lead to the deviation of their architectural semantics (the same voxel configuration may represent completely different meanings in different positions).

According to the analysis of the above two types of models, the focus of this study is not just to enrich the existing tests but to formulate an efficient tile encoding mode suitable for architectural space and construct a spatial synthesis method for specific style samples combined with the TWFC model, establishing a controllable, universal and logical workflow from the target template to the new derivatives.

3. Methodology

3.1. STYLIZED SPACE SYNTHESIS WORKFLOW

Based on the WFC algorithm, the stylized space synthesis process includes explicitly three main steps of "deconstructing-encoding-aggregating":

First, the overall architectural style template is deconstructed into featured modular tiles. According to the positions of essential components such as walls and floors, the axis network of plans and sections is divided and adjusted appropriately to unify the tile scales. Tiles with mirror and rotation relationships are classified or deleted. Thus, an original spatial semantic library with style features is formed.

Then, digital information and constraints are encoded in these tiles. In addition to grading and naming, functional topology and geometric connectivity relationships, as well as subsequent evaluation indicators, are recorded in tiles.

Finally, the WFC algorithm arranges and combines the tiles according to the constraints and resynthesizes them into expected stylized architectural spaces.

3.2. STYLIZED SPACE SYNTHESIS ALGORITHM

This study develops WFC as a 3D stylized space synthesis algorithm, which mainly consists of five parts (Figure 1):

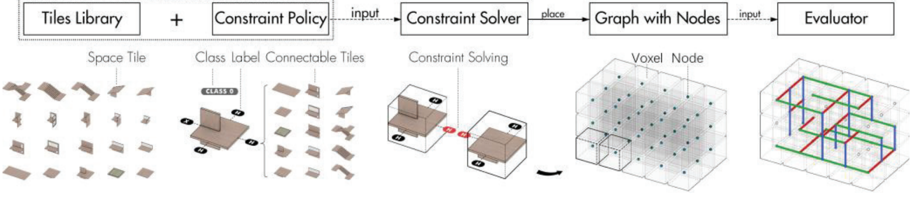


Figure 1. The components and organizations of the stylized space synthesis algorithm

1) The 3D Grid is a virtual generation environment with a topological structure, including nodes in space and voxels indicating the node scale. Each voxel can accommodate a predefined tile. The array mode and voxel shape can be modified according to design requirements, such as tetrahedron, cube, and octahedron.

2) The Tile Library contains all the parts comprised of a prototype and the constraints involved in space generation. The prototype design follows the voxel size and form, integrated with the target style, to form a featured architectural component. The constraints provide necessary data for the implementation of subsequent connections and combinations. For example, the class information is responsible for marking the tiles' placing zones, while the label information marks the adjacency relationship on each face.

3) The Policy defines the connection and exclusion rules among the classes and labels of all tiles in the library. Tiles can be connected to each other if their classes and labels are attachable.

4) The Solver is the decision-making core of this algorithm. By triggering the process of wave function collapse, spatial tiles are gradually selected and placed in the grid environment to aggregate a complete solution. The computing process can be divided into two steps: searching the following planning location and choosing the exact tile. More concretely, the solver will assess all the nodes adjacent to the planned area and take the one with the fewest options as the next planning location (with the largest limitations and the smallest entropy). The evaluator precisely determines the embedded tile, minimizing the probability of making wrong choices. Besides, to prevent the algorithm from getting stuck due to insufficient selection, the solver is given the function of backtracking optimization, where the collapse order and constraint propagation are saved as historical data, and the tile can be reselected by

revoking some actions when the constraint error is larger than the threshold value.

5) The Evaluator provides real-time spatial assessment for the generated proposals to affect the collapse process and the result filtering. By integrating the tile information in the aggregated clusters, multiple data such as building density, greening rate, functional proportions, structural stability, and spatial connectivity can be obtained and visualized.

4. Experiment

Cultural buildings represented by museums and galleries are generally characterized by free spatial organizations and rich spatial types and often need to incorporate local traditional architectural styles, which are suitable as the experimental subjects of this algorithm. A two-stage generative design experiment was conducted for a 1000 m2 cultural building, taking traditional Chinese residences in southern Zhejiang as a stylistic template to verify the application potential of the method.

4.1. ALGORITHM MODEL CONFIGURATION

Before the experiment, the algorithmic components are configured to form a dedicated computational model for cultural architecture.

1) The 3D grid node is a cube with a side length of 4m, close to the geometric logic and spatial scale of typical architectural design. The range is 9*6*3 (36*24*12m) to meet the area requirements.

2) The tile library takes a series of Chinese residences in southern Zhejiang as the style case model and deconstructs them to generate tiles with features including sloping roofs, wooden walls, stone foundations, etc. The tiles are divided into four categories according to utility: Function Class, Greening Class, Connection Class and Vacancy Class (Figure 2).

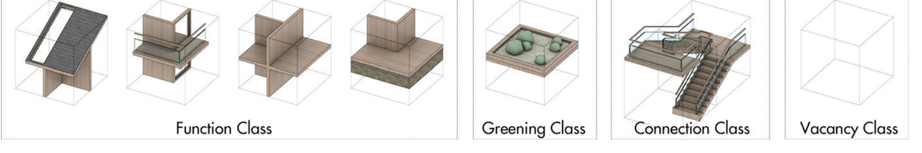


Figure 2. 4 classes of the tiles

As an example, the prototypes of the function class are gradually extracted and hierarchically encoded from the style paradigm. The first-level encoding is related to the global position of the tiles: they are deconstructed vertically into three zones: roof tiles (Roof, R), wall tiles (Wall, W), and base tiles (Base, B); and further subdivided horizontally into convex corner tiles (Corner, Ce), concave corner tiles (Corner, Ca), side tiles (Side, S) and middle tiles (Middle, M). In the second-level encoding, each tile differentiates into details to increase the richness of the library. Four types of balconies can be connected outside the L-shaped wall in corner tiles, deriving 8 subtypes (Wca-Lb1-4, Wce-Lb1-4). In the third-level encoding, each subtype is multiplied by 4 to obtain its rotational variant in each direction (Figure 3). The roof and base tiles also undergo the same process, constituting a total of 271 function tiles.

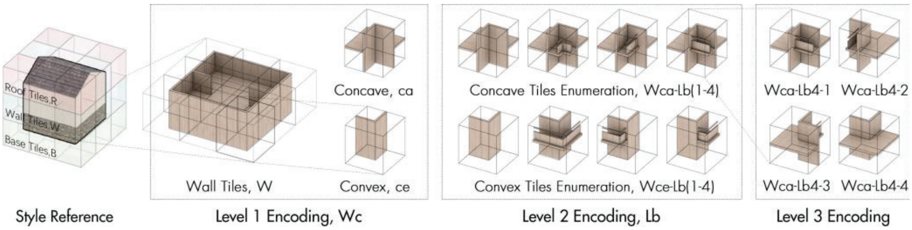


Figure 3. Differentiation from the style template to multiple subtypes

The library also contains 4 types of garden tiles, 8 types of circulation tiles and 1 type of empty tile. There are 284 types of tiles in total. The proportions of garden and circulation tiles are increased to maintain an appropriate ratio: Function Class: Greening Class: Connection Class: Vacancy Class = 10:2:1:1 (Figure 4).

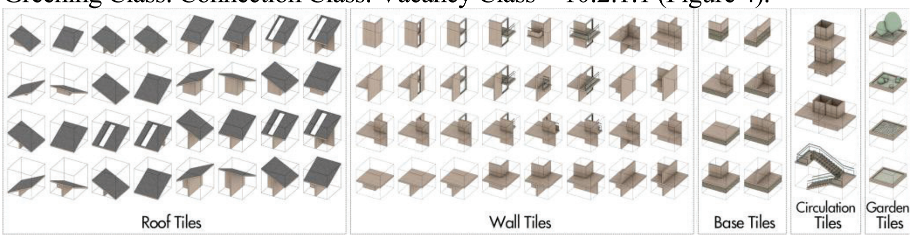


Figure 4. Part of the tile library

After the prototypes are extracted and designed, the tiles need input with constraint information according to the style template. Apart from the 4 class information mentioned above, 6 faces of each tile will be tagged with 32 horizontal and 10 vertical labels (Figure 5). A variety of attributes are also integrated into the tiles to participate in the subsequent evaluation process.

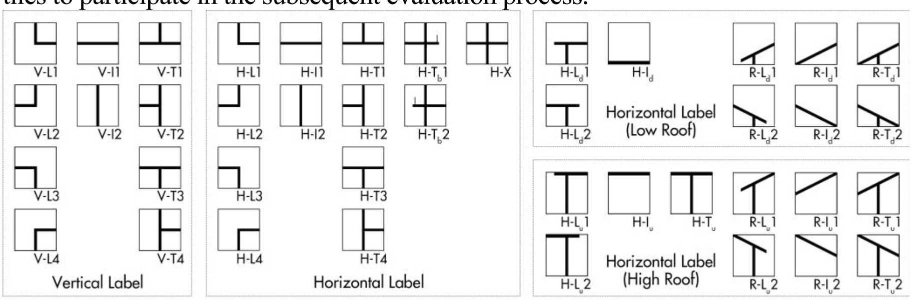


Figure 5. Label graphics on neighboring surfaces

3) The policy setting closely follows the inherent spatial configuration logic of the style template. The continuous alignment of the roof and the non-crossing form of the wall can be incorporated into the rules. The class principles are shown in Table 1, in which the function class can be connected to any other classes with the greatest degree of freedom. The label principles are shown in Table 2 and Table 3.

Table 1: Connection rules for class information

	Function Class	Greening Class	Connection Class	Vacancy Class
Function Class	√	√	√	√
Greening Class	√	√		√
Connection Class	√		√	
Vacancy Class	√	√		√

Table 2. Connection rules for horizontal and vertical labels on wall/base/circulation/garden tiles

	H(V)-L1	H(V)-L2	H(V)-L3	H(V)-L4	H(V)-I1	H(V)-I2	H(V)-T1	H(V)-T2	H(V)-T3	H(V)-T4	H-Tb1	H-Tb2	H-X
H(V)-L1		√					√			√			
H(V)-L2	√						√	√					
H(V)-L3				√				√	√				
H(V)-L4			√						√	√			
H(V)-I1					√		√		√				
H(V)-I2						√		√		√			
H(V)-T1	√	√			√		√						
H(V)-T2		√	√			√				√			
H(V)-T3			√	√	√				√				
H(V)-T4	√			√		√		√					
H-Tb1												√	
H-Tb2											√		
H-X													√

Table 3: Connection rules for horizontal and vertical labels on roof tiles

	H-Lu(d)1	H-Lu(d)2	H-Iu(d)	H-Tu	R-Lu(d)1	R-Lu(d)2	R-Iu(d)1	R-Iu(d)2	R-Tu(d)1	R-Tu(d)2
H-Lu(d)1		√								
H-Lu(d)2	√									
H-Iu(d)			√							
H-Tu				√						
R-Lu(d)1						√				
R-Lu(d)2					√					
R-Iu(d)1								√		√
R-Iu(d)2							√		√	
R-Tu(d)1								√		√
R-Tu(d)2							√		√	

4) The solver still adopts the method of minimum entropy to select the planning position and randomly extracts the connectable tiles.

5) The evaluator uses building area, volume, spatial connectivity, and structural stability as indicators, adjusting the selection preference of different tiles to generate diverse outcomes. At the same time, the results are also visually evaluated to facilitate subsequent optimization.

4.2. SINGLE-OBJECTIVE DIVERSITY SYNTHESIS EXPERIMENT

The experiment at this stage aims to verify the versatility and rationality of the algorithm's outputs, and 60 repeated tests were carried out with the model described above. Figure 6 illustrates the specific process of generating a spatial cluster in a bottom-up manner, while Figure 7 records 18 output results.

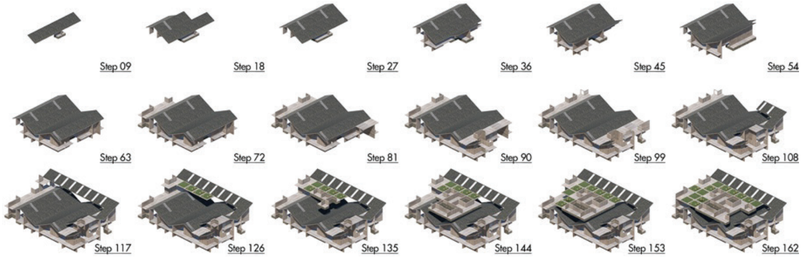


Figure 6. The specific process of generating a spatial cluster

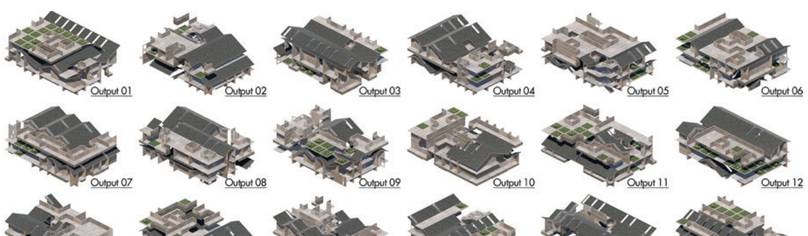


Figure 7. 18 typical outputs of single-objective diversity synthesis experiment

Due to the randomness of the initial tile, the algorithm does not run along a unified path when it collapses, showing intense uncertainty. Therefore, numerous results are produced to satisfy the horizontal exploration and comparison.

Manual observation reveals that almost all the solutions can meet the needs of subsequent deepening, but there are also local defects, the proportion of which is roughly 23%. For instance, Output 02 has suspended tiles; Output 15 gets a cluttered top floor. In contrast, some creative spaces appear in satisfactory results: Output 04 and 12 depict neat small spaces collocated with free common areas; Output 07 and 17 present attractive fold-plate roofs.

Figure 8 demonstrates the visual analysis of some results by the evaluator. When there are no preference settings, the satisfaction degree of the four indicators is relatively average. The first result has a feasible spatial layout, but many slopes on the second-floor lead to poor structural performance; the fourth result has a clear vertical organization with an appropriate area and good structural stability. Therefore, in view of the above analysis, the reliability of this model is roughly 77%, which shows creativity and controllability for design exploration.

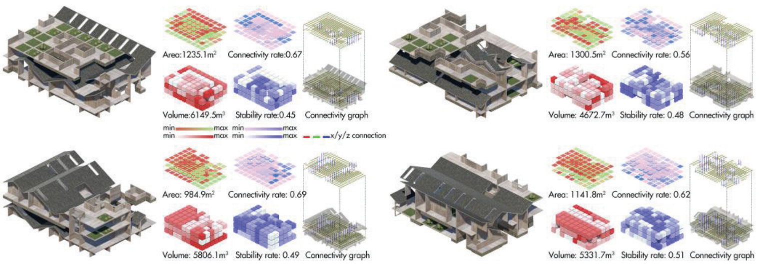


Figure 8. Visual evaluation for 4 outputs

4.3. MULTI-OBJECTIVE ADAPTIVITY SYNTHESIS EXPERIMENT

On the basis of the previous tests, the experiment at this stage verifies the adaptability and expansibility of this method in the face of different design objectives such as site scale, general form and functional ratio. After setting three new generation scenarios, including the courtyard, high-rise tower and country alleys, the algorithm can still generate numerous results with the same style. Figure 9 records 18 typical results in the three scenarios.

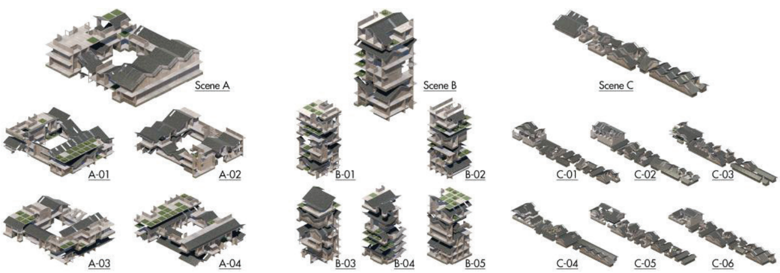


Figure 9. 18 typical outputs of multi-objective adaptivity synthesis experiment

In terms of space style, the three groups of results keep a high degree of unity. After evaluating the rationality, the percentage of eligible outcomes can also be maintained at about 75%. In particular, the completion and innovation of the algorithm increase significantly when dealing with building types that extend in a single direction, such as high-rises and alleys.

5. Conclusions

By modeling the architectural style expression as a constraint-solving process, this paper explores the application of the WFC algorithm to the stylized spaces generation and develops a synthesis workflow that conforms target templates. In two stages of experiments, the algorithm can generate abundant self-similar solutions reflecting the style and complete evaluation and optimization according to the demand tendencies. Meanwhile, it can also adapt to a wide range of design objectives and efficiently perform the task of style synthesis in complex scenarios.

Architecture has more complex and practical constraints than the original application scene of the WFC algorithm to generate 2D images. The solver currently involves high randomness in the tile selections, resulting in homogeneous spaces at the macro level. In further work, this algorithm can be combined with spatial planning and other research to introduce more top-down architectural constraints (such as site environmental conditions, spatial configuration relations, etc.) to improve the comprehension and definition of style and promote the integrity and reliability of the algorithm framework. Intelligent decision-making mechanisms based on artificial intelligence techniques (e.g., deep reinforcement learning) can be embedded in the solver to achieve controllability of the policy to optimize the outputs.

This research questions the idea of considering the architectural system as a whole, shaping it into a generative semantic system with stylized features. By means of deconstructing and encoding, the stylized space synthesis method can retain the stylistic characteristics while producing excellent spatial effects.

References

Christopher, A. J. (1968) Systems Generating Systems. *Architectural Design*, 38, no. December 1968: 605-610.

del Campo, M., Manninger, S., & Carlson, A. (2020). A Question of Style: Style, Artificial Intelligence, and Architecture. In *Architectural Intelligence: Selected Papers from the 1st International Conference on Computational Design and Robotic Fabrication (CDRF 2019)* (pp. 171-188). Springer Singapore.

Gumin, M. (2017, May). WaveFunctionCollapse Readme.md. Retrieved May 20, 2017 from https://github.com/mxgmn/WaveFunctionCollapse/blob/master/README.md . 2017.

Hosmer, T., Tigas, P., Reeves, D., & He, Z. (2020). Spatial Assembly with Self-Play Reinforcement Learning. In *ACADIA 2020: Distributed Proximities. Proceedings of the 40th Annual Conference of the Association for Computer Aided Design in Architecture (ACADIA)*. 2020: 382-393.

Huang, W., & Zheng, H. (2018, October). Architectural drawings recognition and generation through machine learning. In *Proceedings of the 38th Annual Conference of the Association for Computer Aided Design in Architecture, Mexico City, Mexico* (pp. 18-20).

Karth, I., & Smith, A. M. (2017, August). WaveFunctionCollapse is constraint solving in the wild. In *Proceedings of the 12th International Conference on the Foundations of Digital Games* (pp. 1-10).

Koh, I. (2022). Voxel synthesis for architectural design. In *Design Computing and Cognition'20* (pp. 297-316). Cham: Springer International Publishing.

Koning, H., & Eizenberg, J. (1981). The language of the prairie: Frank Lloyd Wright's prairie houses. *Environment and planning B: planning and design*, 8(3), 295-323.

Lu Youyu, & Tong Ziyu. (2020). Research on Discretized Unit Layout Generation Based on Wave Function Collapse Algorithm. *Proceedings of the 2020 National Symposium on Architectural Digital Technology Teaching and Research in Schools of Architecture*.

Stiny, G., & Mitchell, W. J. (1978). Counting palladian plans. Environment and planning B: Planning and design, 5(2), 189-198.

VR/AR/XR & Interactive Environments

AUGMENTED ROBOTIC BRICKLAYING

An Experiment in Remote Programming Robotic Assembly Using Augmented Reality for Brick-Based Structures

YANG SONG[1] and SOOMEEN HAHM[2]
[1] *Department of Architecture, University of Liverpool*
[2] *the Southern California Institute of Architecture (SCI-Arc)*
[1]*yang.song@liverpool.ac.uk, 0000-0003-0340-8629*
[2]*soomeen_hahm@sciarc.edu, 0000-0003-2868-9375*

Abstract. After experiencing the *Covid-19* pandemic, remote communication became one of the key issues in almost every field and discourse. Digital fabrication is no exception, and architects hope to seek a user-friendly way for human-machine interactions. This paper presents experimental research using Augmented Reality (AR) for robotic remote programming. The research tries to develop a unique pipeline and workflow which allows users from different locations to program robots and communicate with machines through AR. A sample workflow has been tested as a series of simple brick assemblies in an online workshop with remote participants. The pipeline allows all users to be able to remotely program and control a robot in AR. For this workshop, we transform the robotic coding method from the traditional computer science way to the plugin-oriented AR visual programming way in *Grasshopper*. As for the physical outcomes, participants all assembled brick-based structures successfully by programming and operating the robotic arm in AR remotely at the end. Associating the interaction in AR with the robotic arm and programming it with interactive visual input methods will make it easier for architectural practitioners to simulate and control industrial robots for complex structure assembly.

Keywords. Augmented Reality (AR), Remote Programming, Robotic Assembly, Brick-based Structure, Online Workshop.

1. Introduction

Due to the *Covid-19* breakout and the social distancing requirements, students were restricted from using on-campus facilities, which disrupted experiential learning that relies heavily on face-to-face interaction (Tatiana et al., 2021). Under the influence of this pandemic, education and practice have to move remotely online. Both instructors and students are experiencing dramatic changes in their modes of teaching and practising (Duaa et al., 2021). In architectural education fields, with few relevant

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 323-332. © 2023 and published by the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA), Hong Kong.

resources and experience, the above challenges have inspired many researchers to explore and develop teaching and practising methods in remote communication modes (Antonio and Lucas, 2021). Moreover, automated architectural fabrication practices, which proceed by machines instead of humans, have become an efficient way to produce physical outcomes from digital design to deal with the Covid-19 situation (Regiane and Luiz, 2020). However, current research gaps exist because the remote software and interfaces are not clear and straightforward enough to complete complex teaching and practising tasks. According to the above changes brought about by the pandemic, there is increasing interest in integrating easier and more intuitive ways for remote education and fabrication in the architectural areas.

The mechanization and industrialization of architecture were a dream of the modernists and gained momentum due to increased access to robotics (Gilles, 2016). Moreover, the coupling of parametric computer-aided design (CAD) with robotic fabrication equipment has enabled the materialization of designs with previously unfathomable levels of complexity and variation (Ryan and Jeffrey, 2018). Current research gaps also exist in that conventional robotic operation requires not only the corresponding computer science knowledge but also demands code debugging and simulation on-site. Even so, the entire operation process is tedious and full of unpredictable errors, requiring people with solid background skills to solve or assist during the process, which makes it challenging and unsafe to introduce to architectural practitioners in a short period. The attempt to introduce robots into the construction industry is not a new research topic (Ines and Merav, 2015). However, exploring a remote and user-friendly robotic operation method convenient for architectural participants is the research in line with the current pandemic trends in the architectural digital fabrication field.

With the development of mixed reality (MR) technology, AR has gradually entered the field of digital fabrication research. The AR system has the following characteristics: combining real and virtual objects in a natural environment, running interactively and in real-time, and registering real and virtual objects contextually in 3D (Bhaskar and Eliot, 2019). Moreover, based on the above unique features, AR is an interfacing technology that has recently become popular in remote architectural practice and industrial robotic fabrication sectors (Chu et al., 2020). Intertwining AR technology with architectural practice opens up possibilities for a remote environment where humans can intuitively communicate with digital content, objects, machines, and spatial context (Song et al., 2021).

This paper presents experimental research using AR technology as an interface for robotic remote programming in architectural practice. The research contributes to developing a unique remote mode and workflow that allows users from different locations to program and operate industrial robotic arms through AR for architectural proposal fabrication. This sample workflow will be tested as a series of simple robotic brick-based structure assemblies through the online workshop with remote participants. Additionally, this online workshop is used as an example to reflect on this new mode of online teaching and remote digital fabrication in architecture education and fabrication reacting to the *Covid-19* pandemic.

2. Methodology

The Augmented Robotic Bricklaying workshop proposes an experiment in remote robotic programming consisting of two phases: A) robotic operation in AR, in which interactive inputs will communicate with the robotic programming process in AR through the screen-based user interface (UI) (e.g., smartphones or tablets) to achieve the trajectory planning and end-effector commanding; B) remote robotic assembly, in which user can edit and preview on the AR virtual robot anywhere for grabbing commands and assembly sequences, and map the operations remotely to the robotic arm in the lab for the physical assembly (Figure 1). This research aims to provide architects with a convenient and easy operation method for robotic fabrication in line with their corresponding knowledge reserves and develop a remote robotic control method that conforms to the pandemic needs. This research adopts the method of preliminary online workshop experimental verification from the participants to validate the feasibility of the above aims and summarise the findings and limitations.

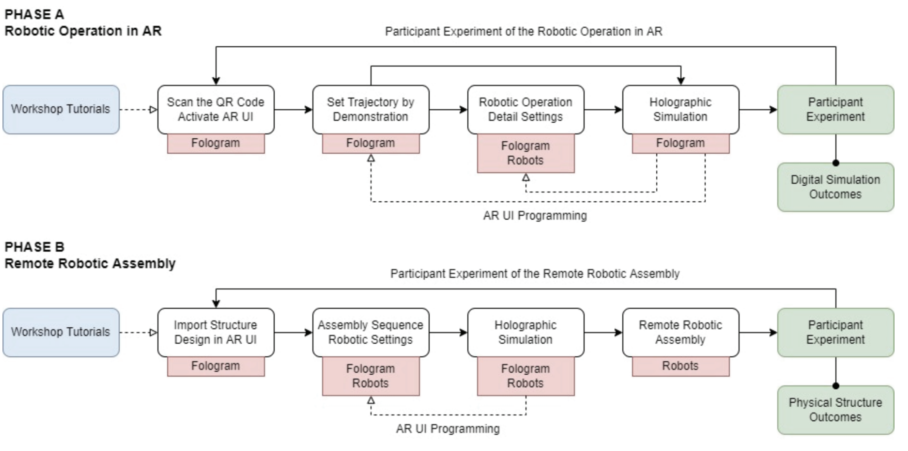


Figure 1. This is the flow chart of the Augmented Robotic Bricklaying workshop, including Phase A (Robotic Operation in AR) and Phase B (Remote Robotic Assembly) (in blue). The related plugin for each critical step (in red) and the outcomes of each phase (in green) are also illustrated.

For this workshop, we transform the robotic coding method from the traditional computer science way to the plugin-oriented visual programming method in the Grasshopper environment, which is familiar for architectural practitioners to understand and manipulate. The employed programming method is driven by an instant connection between the development environment (Grasshopper), AR interactive immersion plugin (Fologram), and robotic operation plugin (Robots). Fologram is an AR plugin developed by architects, which can translate interactive inputs, such as hand gestures, screen taps, device location and QR codes, into digital data in the *Grasshopper*. It bridges humans and machines between physical and virtual through AR. *Robots* is a robotic plugin developed by architects, which enables the user to program the robotic arm in *Grasshopper*, such as trajectory planning, gripper commands, and other operational details. Implementing these plugins will provide architectural practitioners with an easy robotic programming method in AR as a medium. The above plugins work with their integrated graphical algorithm editor, *Grasshopper*, a standard tool in architectural fields, and can easily be integrated into the established immersive robotic programming workflow.

The hardware in this workshop includes mobile devices (smartphones or tablets) for AR-assisted robotic programming, and a *Universal Robot 10e* robotic arm with *RobotiQ 2F-140* gripper in the lab for the remote robotic assembly. Additionally, workshop participants need their personal computers for the AR-assisted robotic programming development in *Grasshopper* and the back-end programs running; as well as the lab-based operator needing a laptop for activating and debugging the robotic operations. Except for the robotic arm, we have to ensure that any software and hardware used in this workshop should be ubiquitous, so that all participants can assess and manipulate the corresponding operations easily and remotely.

This online workshop was taken by 5 bachelor students majoring in architectural design. They all have 3D-modeling experience, but none had prior knowledge and skill in robotic fabrication. Either Fologram or Robots plugin was new to them. Moreover, due to the robotic lab safety restrictions, foam bricks are chosen instead of standard bricks for the physical experiments.

The original contribution of our workshop is to integrate and combine the different functions from various plugin applications into the remote robotic programming method through AR for the brick-based structure assembly and verify the feasibility of this remote method on the participant's experience and outcomes.

3. Experiments and Findings

The online workshop was limited to three days, and its schedule was split between introduction lectures, plugin tutorials, digital workflow development, discussions, remote robotic assembly experiments, and final reviews.

On the first day, participants received introductory lectures about AR and its interactions in architectural and robotic fields, as well as information about the topic and schedule of this workshop. After that, we brought the *Fologram* tutorials to the participants on their mobile devices, as well as computers, to demonstrate the essential interactive AR functions, including interactive input methods, holographic preview, QR code recognition, and AR user interface (UI) developments. These example files are based on the *Rhino/Grasshopper* environment.

On the second day, we delivered the *Robots* tutorial to the participants, so they could understand and program robotic operations in *Rhino/Grasshopper*, such as trajectory planning, grabbing commands, and other robotic operation settings. By combining the AR skills from the first day, participants practiced programming the robot in AR through UI and previewing the robotic operation simulation as 3D holograms to check the details. Furthermore, time slots were set for discussing the robotic operation in AR (Phase A experiment).

On the last day, the participants were asked to play with the prepared brick-based structure AR design scripts developed by Song (Song et al., 2022) for parametric brick-based structure design. After that, participants were asked to program and operate the robotic arm for pick and place assembly remotely. The physical robotic operation of brick-based structure assembly (using foam bricks for our experiments) was live-streamed from the lab. The end of the day was dedicated to the final review and extensive discussion about the findings and user experience of the remote robotic program and operation (Phase B experiment).

In this section, the two phases of workshop outcomes and research findings, robotic operation in AR and remote robotic assembly process, are illustrated and concluded.

3.1. PHASE A - ROBOTIC OPERATION IN AR

Phase A proposes a practical and user-friendly method for architects to control industrial robots in AR, instead of the conventional programming process. Compared with mastering cumbersome computer programming languages, users only need to be familiar with the fundamental common sense of AR and the essential interactive function of AR UI, which will be easy and controllable for beginners. The pre-set background scripts are developed and run in *Grasshopper*. Participants only need to understand the script's logic, but do not need to edit them unless they need other customized robotic functions.

To start the robotic operation in AR, first, participants should scan the *Fologram* QR code generated from *Rhino* with their mobile app to activate and connect their AR environment with the laptops. Second, participants can select the serial robotic arm (*UR 10e* for this workshop) and preview the virtual machine as holograms through the AR UI. Next, after understanding the safety information, such as the working radius of the selected machine, the user can start programming the trajectory through screen-based AR UI inputs. By tapping the mobile screen at the correct position and angle, the waypoints with corresponding coordinate information will be added to the trajectory for participants to preview and modify in AR. Moreover, from the AR UI, users can edit more specific operation settings, including robotic speed, liner or joint movement mode, end-effector commands, and node delay time. Last, participants can preview the virtual robotic movements as 3D AR animations to simulate trajectory and end-effector operation in AR (Figure 2).

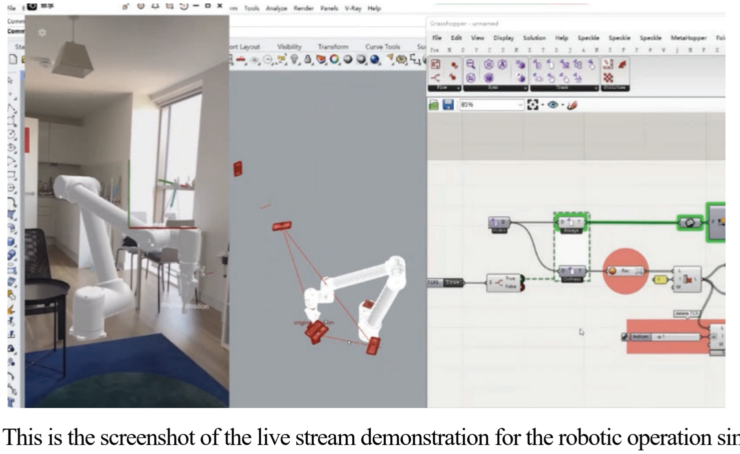


Figure 2. This is the screenshot of the live stream demonstration for the robotic operation simulation in AR during the workshop tutorials. The pre-set script works in the AR UI from Fologram App (left), and Grasshopper (right).

After the demonstration during the workshop, participants started to get familiar with the AR UI operating environment and method and use the pre-set scripts to program and simulate the robotic operation through AR, including trajectory planning, end-effector commands, movement mode, operation speed, and operation node delay time (Figure 3). The participants were asked to document their operation process as AR videos, share and discuss the experiments at the end of the day. Furthermore, it was assessed whether the interactive AR programming was well engaged, e.g., regarding the operation of AR UI and the simulation of customized commands.

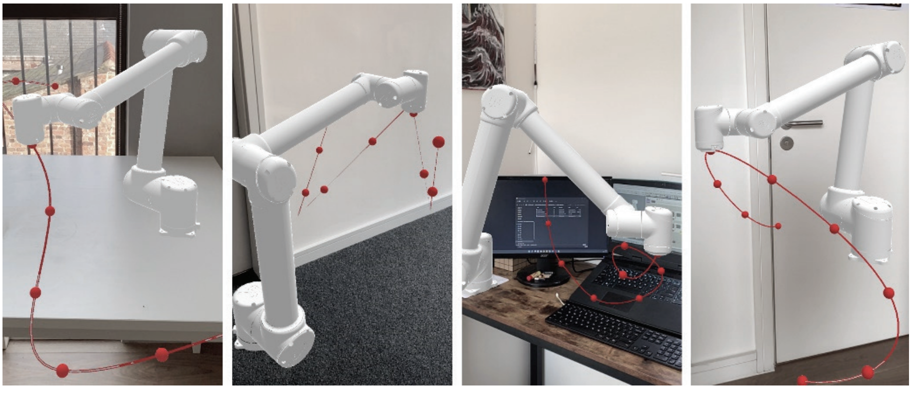


Figure 3. This is the AR UI (Fologram app) screenshot of participants programming robotic trajectory and related detail settings in the AR environment.

The findings of phase A suggest that compared with the conventional coding-based robotic programming method, this screen-based AR programming method can allow the user to demonstrate robotic operation, set commands, and preview the simulation efficiently. After the basic AR interactive function introduction, without any robotic skills, participants can easily program and simulate the robot within ten minutes on their mobile devices. The simulation results are all consistent with the participants' settings. Moreover, the users' feedback was good during the AR programming process. All participants, without a computer science background knowledge, indicated that they could use AR UI easily and freely to program the robotic arm and preview the simulation after the simple pieces of training. However, the commands for the end-effector are currently limited to 'open' and 'close' due to the function of our robotic gripper. But this pre-set script is based on the *Grasshopper*, which is believed to be friendly to architectural practitioners and conducive to developing other subsequent robotic functions depending on different end-effectors.

3.2. PHASE B - REMOTE ROBOTIC ASSEMBLY

Phase B proposes a remote robotic operation method through AR for brick-based structures. Compared with the inflexible traditional on-site operation method, participants can set, edit, and preview the robotic assembly sequence anywhere through mobile devices (smartphones or tablets) in AR, and transmit commands to the lab-based or on-site industrial robots to accomplish the physical assembly, which accommodate the requirements of the pandemic situation. This pre-set script is developed and runs in *Grasshopper*. The remote robotic operation is transmitted by storing and reading command data on QR codes through AR UI. Participants should understand the AR remote operation process logic, but do not need to edit the script unless they need other customized operations. Besides that, participants also need to master the method of storing and reading QR codes to realize the remote control.

To start the remote robotic assembly, the user should first design the brick-based structures as the assembly targets using the pre-set or customized design scripts provided by the workshop instructor. Second, participants need to scan the *Fologram* QR code generated from *Rhino* to pair and activate their mobile app with *Grasshopper* and import the brick-based structures into the AR environment. After selecting the serial robotic arm (*UR 10e*) and gripper (*RobotiQ 2F-140*), participants can preview the brick-based design and the virtual machine as holograms through the AR UI. Third, participants can modify the assembly sequence by touching the holographic bricks one by one from the designed structure on the screen through AR and preview the assembly simulation. Moreover, the simulation can be edited in operation speed, liner or joint movement mode, and node delay time through mobile devices. The determined robotic assembly simulation will be shown as holograms once the AR programming modifications are done. Last, the robotic commands with assembly sequences will be generated as QR codes for remote operations (Figure 4).

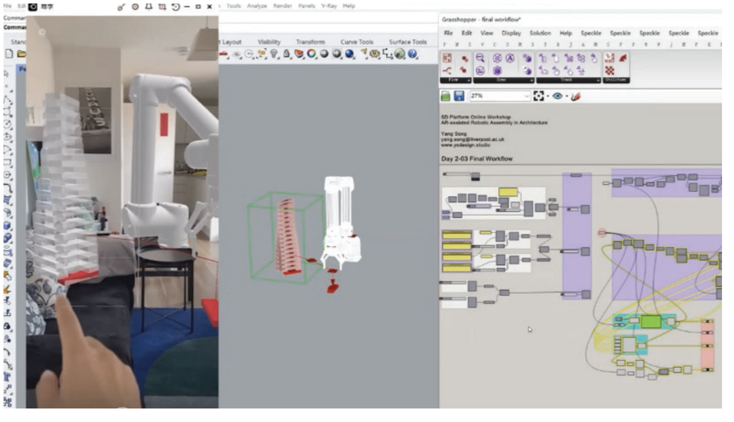


Figure 4. The screenshot of the live stream demonstration for the remote robotic assembly simulation in AR during the workshop tutorials. The pre-set script works in the AR UI from Fologram App (left), and Grasshopper (right).

After the demonstration during the workshop, participants started to design the brick-based structures, and use the pre-set scripts to program the remote robotic assembly through AR UI, including the sequence, movement mode, operation speed, and node delay time. After confirming that the simulations were correct, participants generated the corresponding QR codes for the remote operations. After that, the lab-based operator used the AR UI to scan and run the related commands from the QR codes to help the remote robotic assembly process (Figure 5). Participants were asked to document their remote assembly process as AR videos and the physical outcomes for the final review, share and discuss the experiments at the end of the workshop. Furthermore, it was assessed whether the remote robotic operation was well engaged in AR, e.g., regarding the remote operation of AR UI and the outcomes of physical assembly.

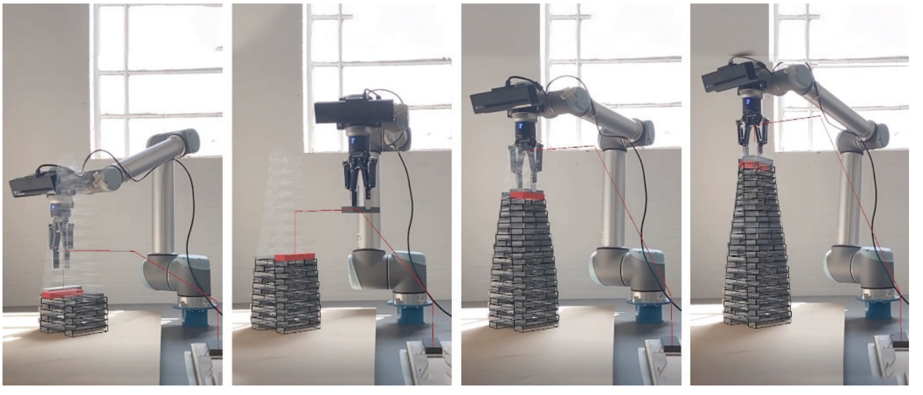


Figure 5. This is the remote robotic assembly of the brick-based structures. The lab-based operator scans the QR codes generated by the participants and runs the robotic assembly process in the sequence set by the participants through AR UI.

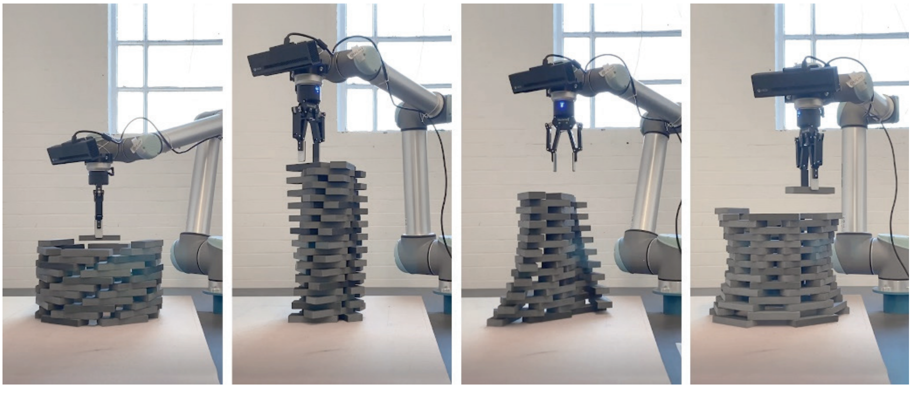


Figure 6. The physical brick-based structure outcomes of participants' designs were assembled by the remote robotic operation method during the Augmented Robotic Bricklaying workshop.

The findings of phase B show that the AR UI programming method fulfilled the remote operation requirements reacting to the pandemic situation. After the basic AR introduction and demonstration, participants can easily program and simulate the robotic assembly operation anywhere and beam the commands remotely to the lab-based or on-site robots in a second for the physical structures assembly. The brick-based structure outcomes are all consistent with the participants' design (Figure 6). Moreover, the users' feedback was good during the program and remote assembly process. All participants indicated that they could use AR UI easily and intuitively to simulate and operate the remote robotic assembly after the simple pieces of training.

However, the lab-based operator is still required to assist with on-site remote robotic assembly. It is because the *Fologram* and *Robots* plugins are temporarily unable to synchronize for remote operation due to network IP address issues. The operator needs to assist in identifying the QR codes, ensuring the assembly operation of the robotic arm on-site, as well as dealing with malfunctions or other special errors. With the update of software and technology, this shortcoming will be solved in the future.

4. Conclusion

Associating the interaction in AR with the robotic arm and programming it with intuitive input methods will make it easier for architectural practitioners to simulate and control industrial robots for complex structure assembly. Using AR as a media between physical and virtual, it simplifies the conventional robotic coding method and replaces it with a user-friendly AR-assisted interface for the safe human-robot collaborative process. This AR programming method also gives the possibility of remote control for robotic assembly, which meets the pandemic requirements for remote communications and makes up for the disadvantage that participants cannot visit the lab in person due to *Covid-19* restrictions.

Closely witnessing the participants' experiments, as well as analysing the *Augmented Robotic Bricklaying* workshop outcomes, it can be concluded that architectural practitioners can quickly employ this augmented programming workflow after an essential AR skill tutorial. Moreover, learning the corresponding interactive input method intuitively and visually through AR and smartphones is much simpler, safer, and more convenient than mastering coding languages. The workshop outcome shows that by programming and remote operating the robot in AR, participants can try digital fabrication tools, such as industrial robots, without any computer science background knowledge. The participants eliminated their fear of industrial machines due to the lack of relevant skills and stimulated their interest in digital fabrication. Through mobile device screen sharing and group discussions during the workshop, participants can learn from each other and exchange their experiences to gain a deeper understanding and reflection.

In conclusion, AR programming and remote robotic assembly is an innovative approach, and this *Augmented Robotic Bricklaying* workshop has also implemented remote digital fabrication methods as online teaching experiments in architectural education and fabrication fields. Future investigations suggest developing more robotic end-effector functions to fit different customized designs. The *Grasshopper* backstage scripts for the AR programming and remote control should be more straightforward and editable that can be quickly developed by architectural practitioners. In the future, with the updates and developments of plugins and devices, the lab-based operator is expected to be replaced to achieve complete remote robotic control and human-machine interaction through the AR environment.

Acknowledgements

This paper presents the process and outcomes of the *AR-assisted Robotic Assembly in Architecture* online workshop from *SDPlatform* in August 2022, instructed by Yang Song (https://soomeenhahm.com/shop/workshop/sdp-2022-aug/). The above selected

experiments in this paper were developed by the following workshop participants: Soomeen Hahm, Park Jongwook, Yejun Yoon, Solah Yoo, and Sowoon Lee.

References

Antonio, R.M.O & Lucas, B.A. (2021). Design and Digital Manufacturing: changes and challenges in product development in the context of remote learning. In *the XXV International Conference of the Ibero-American Society of Digital Graphics* (SiGraDi) 2021, (pp.1345-1357), Brazil.

Bhaskar, B. & Eliot, H.W. (2019). Augmented Reality via Expert Demonstration Authoring (AREDA). In *the Computers in Industry*, 105 (2019) 61-79. https://doi.org/10.1016/j.compind.2018.04.021

Chu, C.H., Liao, C.J. & Lin, S.C. (2020). Comparing Augmented Reality-assisted Assembly Functions: a case study on Dougong structure. In *the Applied Sciences*, 2020, 10, 3383. https://doi.org/10.3390/app10103383

Duaa, A.M., Saba, A. & Amer, A. (2021). Transforming Learning for Archi-tecture: online design studio as new norm for crises adaptation under Covid-19. In *the 9th International Conference of the Arab Society for Computer Aided Architectural Design* (ASCAAD) 2021, (pp. 129-141), Cairo.

Gilles, R. (2016). Discrete Assembly and Digital Materials in Architecture. In *the 34rd International Conference on Education and Research in Computer Aided Architectural Design in Europe* (eCAADe) 2016, Volume 1, (pp. 143-151), Oulu.

Ines, A. & Merav, G. (2015). On-site Robotic Assembly of Double-curved Sefl-Supporting Structures. *In the 19th International Conference of the Ibero-American Society of Digital Graphics* (SiGraDi) 2015, (pp.746-753), Florianópolis

Regiane, T.P. & Luiz, S.R.G. (2020). The importance of collaborative design process and fabrication during Covid-19 emergency - case in Brazil. In *the XXVI International Conference of the Ibero-American Society of Digital Graphics* (SiGraDi) 2020, (pp.357-362), Medellin.

Ryan, L.J. & Jeffrey, A. (2018). Interfaces for Adaptive Assembly. In *the international conference of the Association for Computer Aided Design in Architecture* (ACADIA) 2018, (pp. 126-135), Mexico City.

Song, Y., Koeck, R. & Luo, S. (2021). Review and analysis of augmented reality (AR) literature for digital fabrication in architecture. In *the Automation in Construction*, Volume 128, 2021, 103762. https://doi.org/10.1016/j.autcon.2021.103762

Song, Y., Agkathidis, A. & Koeck, R. (2022). Augmented Masonry Design: a design method using Augmented Reality (AR) for customized bricklaying design algorithms. In *the 10th International Conference of the Arab Society for Computer Aided Architectural Design* (ASCAAD) 2022, (pp. 703-712), Debbieh.

Tatiana, E., Vincent, H. & Lena, M. (2021). The Digital Design Build: modes of experiential learning in the pandemic era. In *the 26th International Confer-ence of the Association for Computer-Aided Architectural Design Research in Asia* (CAADRIA) 2021, Volume 2, (pp. 41-50), Hong Kong.

HOLOGRAPHIC CONSTRUCTION IN CIRCULAR DESIGN

A Literature Review

MIA TEDJOSAPUTRO[1]
[1]*Xi'an Jiaotong - Liverpool University*
[1]*Mia.Tedjosaputro@xjtlu.edu.cn, ORCID 0000-0003-1328-3671*

Abstract. This study looks at a review of literature in regard to holographic construction within the period of 2012-2022, specifically analyses of its potential to aid circular design adaptation. The rationale behind this study is the lack of circularity context in holographic construction and augmented reality studies. The outcome of this literature review reveals aspects of holographic construction that lead to support for circular design opportunities. This paper provides a snapshot and is summarised with identified research gaps related to four key principles of circular buildings.

Keywords. Holographic Construction, Circular Design, Augmented Reality, Paperless Architecture

1. Introduction

According to the 2022 Global Status Report for Buildings and Construction, the buildings and construction sector, which is the highest contributor of carbon emissions in comparison to other industries, is not on track to achieve decarbonisation by 2050 (UNEP, 2022). The report highlights that CO2 emissions in 2021 rebounded back to above pre-pandemic levels in major economies. At the Sharm el-Sheikh COP (Climate Change Conference 27) which was held in November 2022, the imminent need to move away from linear economy with its 'take-make-dispose' sequence to circular economy which includes the restorative use of resources has become a matter of urgency. The second point of departure is the notion of paperless architecture, particularly pushing the notion of paperless fabrication on generating and testing new ideas using holographic construction (utilising Augmented Reality headsets). The proposed study looks at leveraging technology to facilitate a swift transition to circular design and circular architecture.

The aim of this literature review paper is to investigate how technology can be leveraged to facilitate circular design adaptation. The research question is, "How can holographic construction aid the practice of circular design?" This literature review provides an understanding of circular economy, circular design in the building industry; and subsequently four circular design principles are mapped to previous studies in holographic construction and augmented reality to draw future potentials in facilitating the transition.

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 333-342.

2. Search Strategy

Building on the previous systematic literature review on the state of XR research in architecture conducted in 2015-2019 by Stals and Caldas (2022), the search is conducted based on the same combination of two sources of publications. The first source is Cumincad (Cumulative Index of Computer Aided Architectural Design) which is a database for publications related to computer aided architectural design. The second source is based on the same journal publications used in Stals and Caldas (2022)'s study. There are six journals database: Design Studies, Co Design, Architectural Science Review, Frontiers of Architectural Research, Archnet-IJAR and IJAC. The same set of search terms is used in both sources, they are: "holographic construction", "hologram", "holographic instruction", "holographic assembly" and "augmented reality".

In the initial phase, the author identified and processed 4715 publications from the first source, which were then screened by title, abstract, keywords and year of publication. In total n=77 publications from the first source are identified for their suitability for data-extraction and full texts were processed and analysed. From the second source, journal publications, using the same data-extraction technique, n=9 publications are included in this study. The last stage includes coding the 80 publications' full text using a qualitative data analysis software and analysing 45 relevant papers which illustrate AR potentials in regards to four circular building principles as outlined by Rahla et al. (2021).

3. Related Literature

3.1. CIRCULAR ECONOMY

The origins of the term circular economy are not novel and date back at least 60 years, when the concept was discussed by Boulding (1966). With growing awareness of environmental issues, Pearce et al. (1990) posited an economic model based on natural resources and the environment. The term has continued to be refined and the Ellen MacArthur Foundation (2022) suggests that the most generally accepted term is *"...one that is restorative and regenerative by design and aims to keep products, components and materials at their highest utility and value at all times, distinguishing between technical and biological cycles."* Wysokińska (2016) further refines the concept as a 'closed-loop economy' that does not generate excessive waste and whereby any waste becomes a resource. Kirchherr et al. (2017) analysed 114 definitions of circular economy and the authors ultimately defined it as: *"An economic system that replaces the 'end-of-life' concept with reducing, alternatively reusing, recycling and recovering materials in production/ distribution and consumption processes.... Operates in micro-level (products, companies, ...), meso-level (eco-industrial parks) and macro-level (city, region, nation and beyond)..."* Fundamentally, it can be seen from the above that in contrast to a linear economy, in which a product begins its life as raw materials which are processed, manufactured, used, and ultimately disposed of, the circular economy is cyclical and emphasises reuse and recycling of end-of-life products.

3.2. CIRCULAR DESIGN

With regard to the application of this concept to the building industry, Mazur (2021) notes that the traditional linear model bifurcates from the circular model at the end of life of a building. The stages in a building's life up to that point are the same. Raw material is gathered, process, and used for construction. At end of life, the building can be demolished and the resulting waste becomes landfill, or disassembled and the materials either reused or recycled. Kanters (2020) notes that this potential disassembly must be provided for at the design stage and requires cooperation between architect and client- this design-for-deconstruction accords with the cradle-to-cradle principles of Mcdonough and Braungart (2002). The benefits of this approach include increased flexibility in the design, optimised operation and maintenance, and reduced environmental impact. There are disadvantages however as noted by Mazur (2021) including regulations (which tend to focus more on efficiency of operational energy than embodied energy) and issues with materials supply- since this is a relatively novel concept in architecture, new materials are needed in the early iterations of these building as there is a shortage of reused materials.

Rahla et al. (2021) identified four key principles. These are design for disassembly, materials selection, design for adaptability, and designing out waste.

3.2.1. Designing out waste

According to Keys et al (2000, cited in Spreafico (2022), design for reducing waste means reducing exhaust components and consumption of auxiliary materials or consumables. Den Hollander et al (2017, cited in Spreafico (2022) suggest that at end-of-life, the energy in a material can be recuperated during its decomposition. This includes for example combustion of biomaterials for electricity generation, since if the material were left to biodegrade naturally, the carbon released as carbon dioxide during the process would be equivalent to that generated during combustion. Hebel et al. (2014) mention moreover that building materials can be generated from waste through several processes. The first, densification, involves compacting solid waste into building blocks. The second, reconfigured waste, involves the processing of surplus materials into useful building materials, such as OSF (oriented-strand fibreboard). The third method, transformed waste, involves altering the molecular structure of waste. Typically, this involves grinding the waste into powder then melting it and moulding it into new shapes without the need for binders. This process is possible with concrete and other aggregates, and also with plastics and polymers. Designed waste materials, the fourth method, are products designed to never be altered from their original form or function. Organic waste design, the fifth method, incorporates organic waste into building materials, for example coffee grinds into roofing tiles. Finally, cultivated waste materials are organic elements such as fungi and bacteria that can be used not only to grow building elements, but also repair them, for example by filling cracks.

3.2.2. Design for adaptability

Hebel et al. (2014) mention three dimensions of transformation: spatial, structural, and element/material. These three dimensions combined suggest that building elements should be modular. Again, this concept is very much interrelated with the other three concepts. It further fits into the overarching category of MLC, or multiple life cycle approaches (Sassanelli et al., 2020) which is also concerned with standardisation, compatibility, modularity and upgradability. The modularity of building elements should facilitate transfer between structures during renovations and builds, prolonging the life of the elements to potentially a longer duration than that of the buildings themselves, hence the term multiple life cycles. Two related concepts are design for reuse and design for remanufacturing. Hooton and Bickley (2014, cited in Spreafico (2022) suggest that elements be designed with the objective of being reusable for either the same or other purposes, while Nasr and Thurston (2006) suggest that the remanufacturing processes should be selected to minimise energy consumption. This would reduce the addition of embodied energy in the material.

3.2.3. Materials selection

The concept of the materials passport is discussed by Munaro and Tavares (2021). They cite BAMB (2016) who state that it is a tool to *"document and track the circular potential of materials, products and systems"* with a view to recovery and reuse. Munaro and Tavares (2021) findings indicated that within Europe, greater government support and standardisation is required in terms of tax incentives, regulation and cooperation with industry. Rahla et al. (2021) provide guidance on the selection of materials, as mentioned previously. The distinguish between recycling and upcycling, and the related concept of reuse, but also on ease of deconstruction. These four concepts are interrelated as this aspect is equally applicable to DfD. Two other aspects they suggest are energy recoverability and biodegradability. This suggests that materials taken from nature should be able to be returned to nature as part of the earth's natural carbon cycle, while energy recoverability means that embodied energy should be able to be upcycled.

3.2.4. Design for disassembly (DfD)

Crowther (2005) cited in Akinade et al. (2017) suggests that in addition to component reuse, recycling and remanufacture are elements of design for disassembly. However, Akinade et al. (2017) argue that recycling- now common practice in the building industry- should not constitute part of DfD because recycling does not recoup the embodied energy and embodied carbon in the materials. They thus propose three broad areas to maximise the success of DfD: materials-related factors, design-related factors, and human-related factors. Materials should be durable, use nut and bolt fixings, avoid secondary finishes, and should be non-toxic and not composite. Design should be modular, elements designed for offsite construction, and adhering to a standard structural grid and a layering approach to building. Human factors concern communication, tools and training. O'grady et al. (2021) further exemplify these elements and disambiguate the terms disassembly and deconstruction. Deconstruction refers to the structural elements of a building as opposed to the components which comprise the fabric of a building, such as cladding, flooring and fixtures. Disassembly can occur at any stage during a building's lifecycle, for example during renovation, while deconstruction is an end-of-life process.

4. Results

From the initial screened 5000 papers at the initial stage, 80 publications related to AR are further studied. During the last filtering processes, 45 papers out of 80 full papers were filtered based on their relevance to circular design and are classified to related four circular design principles, publications which did not explicitly mentioned potentials related to the principles were discarded. Due to limited number of pages, only selected publications to exemplify this literature review is included in the bibliography. Figure 1 provides summary of mapped AR potentials with regard to the four principles and sub-principles.

AR and Designing out waste	
Planning materials reuse and recovery	√
Off-site construction	√
Optimisation of material use	√
Embracing a lean design	

AR and Materials selection	
Materials with low embodied energy and carbon	√
Transparent materials content	
Materials that can embrace disassembly and adaptability	√
Technical and biological materials that can be put in the industry or	

AR and Design for adaptability	
Standardisation, compatability, modularity and upgradability	√
Re-use and re-manufacturing	√
Multiple uses of structure	
Renovation, Rehabilitation and Conservation	√

AR and Design for disassembly	
Safe and accessible disassembly	√
Guidelines for deconstruction	√
Disassembly process using standarised tools	
Connectors and fixings allow multiple uses	
Minimal chemical connections	

Figure 1. Summary table

5. Discussion

5.1. AR AND DESIGNING OUT WASTE

Designing out waste emphasises selection of materials based on their reusability and overall recommends lean design. Three potential aspects related to designing out waste are highlighted in previous studies related to AR: 1) planning materials reuse and recovery, 2) off-site construction, and 3) optimisation of material use. Parry and Guy (2020) attempted to demonstrate how with help of mixed reality fabrication technology, it can promote effective re-use of scrap timber pieces. Scrap timbers were categorised by their cross-section profiles, an interface to show wireframe holograms were used to mark timber pieces was developed in relation to the developed parametric model of a small non-structural object. The final product was assembled in chunks using a series of holographic guides. The use of AR enables unskilled makers to assemble complex prefabricated structure as seen in Liu et al. (2022)'s study. Azambuja et al. (2022) developed an accurate and low-cost re-useable mould methodology to aid production of customised voussoirs using HoloLens© AR device to adjust 4-5 elements of the mould, giving rise to optimisation of the material use.

From the previous studies, it can be seen that in relation to designing out waste, the main architectural application from the use of holographic construction is highly attributed to the capability to assemble. Beyond this capability, it is noticed that how holographic construction can help to promote this circular building principle is still under-implemented. In addition, the current state of AR studies in this paper scope has not explored or has not identified how the technology can aid five potential ways to generate of building materials from waste, as elaborated in Section 3.2.1. For example, the first way is densification which involves compacting solid waste to solid blocks. One example is NewspaperWood™, a converted wastepaper to a wood like properties and aesthetics which can be re-introduced into paper recycling system once discarded. The maximum dimension of the panels is 140mm x 380mm and can be used for interior furnishing purpose. One potential area which can be AR-assisted is design process to simulate the artificial grain-based according to the design of furnishing implementation, potentially flooring or wall panels. Although the processes to manufacture recovered materials often is highly skilled in a controlled environment such as a factory, it is possible to promote processing recovered material for unskilled labour using AR technology.

5.2. AR AND DESIGN FOR ADAPTABILITY

Design for adaptability concerns at a macro level the portability of the building or façade for different purposes, and the modularity and upgradeability of components. Previous studies looked at three AR related aspects such as: 1) standardisation, compatibility, modularity and upgradability of building elements, 2) re-using and re-manufacturing, and 3) renovation, rehabilitation and conservation. Kontovourkis et al. (2019) investigates the potential implementation of AR in the assembly of a modular shading device and found that through AR usage, it is possible to improve users' assembly time performance. The study also highlights the potentials of AR use in unskilled users. They also posit that the technology can provide potential impact with regards to reducing construction time and cost.

Re-using practices using AR are related to adaptability of formworks. A project by Jahn et al. (2022), the Steampunk Pavilion, developed a simple timber formwork system with 100mm increments as a temporary scaffold for steam bending process of 133 timber boards using a holographic model. Similarly, a study by Lu et al. (2022) investigated a more sustainable concrete framework which eliminate wood waste in the casting process, by substituting it with adaptable metal sheets for a small scale structure. A guidance app was developed to support unskilled labour.

In another study by Sato et al. (2016), a marker-less AR system is proposed to aid outdoor renovation and maintenance projects at early design stage. Moreover, mobile AR was used for urban rehabilitation and conservation processes by facilitating the public/ users to obtain documental, constructional, functional, and historical information by pointing the device at a specific building. This study aims to be beneficial for architectural designers to help with project development, city council technicians will be able to know which buildings are currently approved by the municipality, and real estate agents may use the platform to access information of a specific building.

One aspect of AR applications which is not widely considered is to facilitate multiple uses of structure at the end of a building's life cycle. It is possible that in conjunction with a scanning system such as in material passport, AR technology can be used to aid on-site discussion to illustrate how existing structural elements can be simulated in its second building cycle and its lifespan layers, according to Brand (1995).

5.3. AR AND MATERIALS SELECTION

Material selection focuses on the environmental impact of materials in terms of embodied energy and carbon, biodegradability, and transparency of sourcing, in addition to their suitability for disassembly and recycling/disposal. The current AR application for this principle is quite limited. Most AR applications with regards to bamboo design are related to temporary structures and far stretching to be brought in mainstream construction. Goepel and Crolla (2020) are among the main scholars who continuously push the AR capability in design in conjunction with computational design.

AR systems can be used to raise awareness of circular material selection, in the form of educative applications. Potentially this can be achieved in mobile AR due to its lower investment. Having an interactive information bank of potentials of how building materials can be re-used would be beneficial. In addition, the concepts of materials passport perhaps can be included in the holographic construction system in accordance with the idea to keep circularity of building materials, for instance, a deconstruction manual for a certain material. This is related to the listed potential of 'transparent materials content'.

AR potentials in materials that can embrace disassembly and adaptability can be tailored specifically to simulating the adaptability in the form of holographic guidance. Lastly, with regards to the selection of technical and biological materials that can be put in the industry or nature, this potential is far stretched to be supported in AR. However, this can also be included in the educational support of circular design.

5.4. AR AND DESIGN FOR DISASSEMBLY

Disassembly should be safe and use standardised tools, and to assist with this process, reusable connectors and fixings should be used, and the use of chemical connections minimised. Within the scope of this paper, a couple of previous studies included disassembly aspect of a non-load bearing small scale projects (Hahm, 2019, Kontovourkis et al., 2019), however neither mention the strategy in great length. Both studies consider the disassembly process is to be conducted not long after the assembly. There is no previous study who explicitly studies Design for Disassembly (DfD) approach at the end of a structure's life cycle, which presents a significant research opportunity. The most related AR applications in circular design concern the assembly system which was mentioned in 44% of the total 45 analysed publications. This is due to the fact that disassembly and assembly (which is the most exploited characteristic of AR) are interrelated. As anticipated, to be able to understand disassembly potentials, there is a need to understand how construction materials were being assembled and there is a plethora of studies on holographic assembly guidance which is transferable to the disassembly guidance. The challenge can potentially be trying to comprehend a system which was already assembled a few years beforehand.

With the lack of previous studies related to AR potentials in DfD, the remainder of this discussion section will focus on mapping the feasible assembly potentials to its disassembly potentials. Five characteristics of DfD in circular building are safe and accessible disassembly; guideline for deconstruction; disassembly process using standardised tools; multiple uses of connectors and fixings; and minimal chemical connections (Rahla et al., 2021). From these five, the first two could potentially be AR-assisted.

Firstly, safe and accessible disassembly aspect of circular building is related to reducing 2D to 3D gap using AR and is a topic of concern in industry generally though thus far little research on its application has taken place in architecture. investigated the design of a gearbox for disassembly. They argued that firstly, accessibility of components to be reused should be prioritised- which is principle which could be adopted by the building industry- then tested three methods of disassembly to determine which was the most efficient. They noted that the use of AR can simplify the disassembly process because workers could follow the AR overlay in real time, without requiring training on how to disassemble a certain product. This may be particularly useful in the building sector where buildings tend to be unique designs, so training workers to disassemble a certain building may not be portable to the next project. It may be inferred therefore that efficiency can be improved if the AR deconstruction model is built concurrently with the construction model.

Secondly is guideline for deconstruction, perhaps where the main support from AR technology can take place. Chen et al. (2018) developed an AR assembly guidance system for Tou Kung structure, a 'five-storey insertion' traditional Chinese tectonic system. The study provided insights related to the assembly behaviour which are component classification and selection, and component searching. It was identified that classification methods should be based on the needed components in each stage, AR support can facilitate flexibility to show stages to be displayed, a progress bar being an interactive visualisation of each element. Bringing this to DfD system, as noted above, if the AR model for disassembly were created at the same time as the model for assembly, the efficiency of the deconstruction process could be improved by reducing the time required and thus the costs involved.

6. Research gaps

From this literature review in holographic construction related to circular design, it presents research gaps which can serve as points of departure for future research projects:

- Designing out waste: support for generating materials from waste
- Design for adaptability: illustrating the potential of adaptations in next building cycles
- Materials selection: educational and awareness raising tools
- Design for disassembly: disassembly guidance and deconstruction model

7. Conclusions

This study attempts to answer the research question which is "How can holographic construction aid the practice of circular design?" Using the four main principles of circular design, the author mapped out potential holographic supports related to the built environment based on published Augmented Reality studies in 2012-2022. It was observed that holographic construction can potentially aid in disassembly design and processes, AR-assisted design simulation for newly refurbished materials and educational tools. It is anticipated that due to AR potentials to aid assembly processes, the fourth principle, Design for Disassembly (DfD) provides the most relevant opportunities. In agreement with previous related literature review in XR by Stals and Caldas (2022), there is a necessity to consider the use of XR technology to support practices of professional architects; beyond the scope of experimental research. The use of holographic construction in providing assistance to transition to circular design is becoming imminent.

Acknowledgements

Research is supported by Xi'an Jiaotong - Liverpool University Research Development Fund RDF-22-01-003.

References

Akinade, O. O., Oyedele, L. O., Ajayi, S. O., Bilal, M., Alaka, H. A., Owolabi, H. A., . . . Kadiri, K. O. (2017). Design for Deconstruction (DfD): Critical success factors for diverting end-of-life waste from landfills. *Waste management*, 60, 3-13. https://doi.org/10.1016/j.wasman.2016.08.017

Azambuja Varela, P., Lacroix, I., Güzelci, O. Z., & Sousa, J. P. (2022). Democratizing Stereotomic Construction through AR Technologies-A reusable mold methodology to the production of customized voussoirs using HoloLens. Paper presented at the *The 40th Conference on Education and Research in Computer Aided Architectural Design in Europe (eCAADe 2022),* Ghent, Belgium.

BAMB (Producer). (2016). D1 Synthesis of the State-Of-The-Art: Key Barriers and Opportunities for Materials Passports and Reversible Building Design in the Current System. Retrieved from http://www.bamb2020.eu/wp-content/uploads/2016/03/D1_Synthesis-report-on-State-of-the-art_20161129_FINAL.pdf

Boulding, K. E. (1966). *The economics of the coming spaceship earth.* New York, 1-17.

Brand, S. (1995). *How buildings learn: What happens after they're built:* Penguin.

Chen, Z.-R., Liao, C.-J., & Chu, C.-H. (2018). An assembly guidance system of tou kung based on augmented reality. Paper presented at the *CAADRIA 2018*, Beijing, China

EllenMacArthurFoundation. (2022). *What is a circular economy?* Retrieved from https://ellenmacarthurfoundation.org/topics/circular-economy-introduction/overview

Frizziero, L., Liverani, A., Caligiana, G., Donnici, G., & Chinaglia, L. (2019). Design for Disassembly (DfD) and Augmented Reality (AR): Case Study Applied to a Gearbox. *Machines*, 7(2), 29. https://doi.org/10.3390/machines7020029

Goepel, G., & Crolla, K. (2020). AUGMENTED REALITY-BASED COLLABORATION-ARgan, a bamboo art installation case study. Paper presented at the *CAADRIA 2020*, Bangkok and online.

Hahm, S. (2019). Augmented Craftsmanship. Paper presented at the *ACADIA 2019*, Austin, TX.

Hebel, D., Wisniewska, M., & Heisel, F. (2014). *Building from Waste: Recovered Materials in Architecture and Construction.* In: Birkhauser Verlag, Basel. http://dx.doi.org/10.1515/9783038213758

Jahn, G., Newnham, C., & van den Berg, N. (2022). Augmented Reality for Construction From Steam-Bent Timber. Paper presented at the *CAADRIA 2022,* Sydney.

Kanters, J. (2020). Circular building design: an analysis of barriers and drivers for a circular building sector. *Buildings*, 10(4), 77. http://dx.doi.org/10.3390/buildings10040077

Kirchherr, J., Reike, D., & Hekkert, M. (2017). Conceptualizing the circular economy: An analysis of 114 definitions. *Resources, Conservation and Recycling*, 127, 221-232. https://doi.org/10.1016/j.resconrec.2017.09.005

Kontovourkis, O., Georgiou, C., Stroumpoulis, A., Kounnis, C., Dionyses, C., & Bagdati, S. (2019). Implementing Augmented Reality for the Holographic Assembly of a Modular Shading Device. Paper presented at the *eCAADe 37 / SIGraDi 23*, Portugal.

Liu, S., Wei, Z., & Wang, S. (2022). On-site Holographic Building Construction: A Case Study of Aurora. Paper presented at the *CAADRIA 2022*, Sydney.

Lu, X., Meng, Z., Rodriguez, A. L., & Pantic, I. (2022). Reusable Augmented Concrete Casting System-Accessible method for formwork manufacturing through holographic guidance. Paper presented at the *eCAADe 40*, Ghent.

Mazur, L. (2021). Circular economy in housing architecture: methods of implementation. Acta Scientiarum Polonorum. *Architectura*, 20(2). https://doi.org/10.22630/aspa.2021.20.2.15

McDonough, W., & Braungart, M. (2002). Design for the triple top line: new tools for sustainable commerce. *Corporate Environmental Strategy,* 9(3), 251-258. https://doi.org/10.1016/S1066-7938(02)00069-6

Munaro, M. R., & Tavares, S. F. (2021). Materials passport's review: challenges and opportunities toward a circular economy building sector. *Built Environment Project and Asset Management,* 11(4), 767-782.

Nasr, N., & Thurston, M. (2006). Remanufacturing: A key enabler to sustainable product systems. *Rochester Institute of Technology*, 23, 14-17.

O'Grady, T., Minunno, R., Chong, H.-Y., & Morrison, G. M. (2021). Design for disassembly, deconstruction and resilience: A circular economy index for the built environment. *Resources, Conservation and Recycling,* 175, 105847. https://doi.org/10.1016/j.resconrec.2021.105847

Parry, C., & Guy, S. (2020). Recycling construction waste material with the use of AR. Paper presented at the *The International Conference on Computational Design and Robotic Fabrication.*

Pearce, D. W., Turner, R. K., & Turner, R. K. (1990). *Economics of natural resources and the environment:* Johns Hopkins University Press. https://doi.org/10.2307/1242904

Rahla, K. M., Mateus, R., & Bragança, L. (2021). Implementing Circular Economy Strategies in Buildings—From Theory to Practice. *Applied System Innovation*, 4(2), 26. https://doi.org/10.3390/asi4020026

Sato, Y., Fukuda, T., Yabuki, N., Michikawa, T., & Motamedi, A. (2016). A marker-less augmented reality system using image processing techniques for architecture and urban environment. Paper presented at the *CAADRIA 2016,* Melbourne

Spreafico, C. (2022). An analysis of design strategies for circular economy through life cycle assessment. *Environmental Monitoring and Assessment,* 194(3), 1-33. https://doi.org/10.1007/s10661-022-09803-1

Stals, A., & Caldas, L. (2022). State of XR research in architecture with focus on professional practice–a systematic literature review. *Architectural Science Review*, 65(2), 138-146. https://doi.org/10.1080/00038628.2020.1838258

UNEP. (2022). 2022 Global Status Report for Buildings and Construction: Towards a Zero-emission, Efficient and Resilient Buildings and Construction Sector.

Wysokińska, Z. (2016). The" new" environmental policy of the European Union: A path to development of a circular economy and mitigation of the negative effects of climate change. *Comparative Economic Research,* 19(2), 57-73. https://doi.org/10.1515/cer-2016-0013

INTELLIGENT PACKING AND UNPACKING NAVIGATION OF PREFABRICATED METAL JOINT TIMBER FRAME

YANFENG WANG[1] and TAKU LIZUMI[2] and SOHEI YOSHIDA[3] and YASUSHI IKEDA[4]
[1,2,3] *Keio University.*
[4] *The University of Tokyo.*
[1]*yanfengwang007@gmail.com*
[1]*https://orcid.org/0000-0002-9257-2027*

Abstract. The packing and transportation of building materials is an important process in building construction. There are many small construction sites in Japan, where it is impossible to take all building materials out of the vehicles and store them on the site at one time. At the same time, many customized structural parts are often involved in building construction now. Their size and shape do not have fixed specifications like standardized building materials. In this article, automatic packing and analysis evaluation algorithms were written in Rhinoceros and Grasshopper environments to help users calculate and generate multiple results after loading the same wood in different order, and evaluate each result in multiple aspects. It is convenient for users to compare the advantages and disadvantages of different loading methods and finally choose the appropriate scheme. Finally, a packing case is analyzed by this algorithm, and the analysis results were obtained in a short time. Experimental results show that the algorithm can solve the problem effectively and successfully, and it has high practical value.

Keywords. Three-dimensional Packing, Constructive Algorithm, Irregularly Non-rectangular Cargo, Optimization, Evaluation

1. Introduction

1.1. RESEARCH BACKGROUND

This study is related to the IST Project (Advanced Information System for Steel Timber Frame Joinery Construction) of IKEDA Laboratory. The goal of IST project is to develop a comprehensive navigation system based on Mixed Reality (MR) technology and Hololens2 equipment for timber with metal parts, from factory assembly and processing, to packing and transportation, and finally to on-site building construction. This paper will introduce the development of the second part of the system, the Navigation system for packing and transportation of timber with metal parts through software Rhinoceros and Grasshopper and C# coding knowledge.

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 343-352. © 2023 and published by the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA), Hong Kong.

IST project is based on one specific case study from the building named MiraiSouzouJuku (Fig. 1). In the case of this building, the shape is a triangle with a certain degree of curvature. Therefore, the length of the timber construction components such as beams and pillars has a great change. It faces great difficulties in the timber processing in the factory, the packing and transportation, and the construction management.

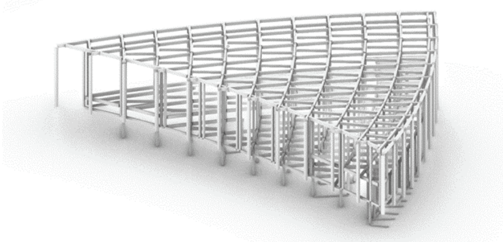


Figure 1. Model of MiraiSouzouJuku

There are different needs in different situation. The factory hopes to assemble timber with metal parts in the factory rather than on the construction site, so as to ensure the accuracy of assembly and save installation time. But in this case, the shape of timber will become irregular.

On the site, due to the limitation of the size of the site, it is difficult to store and manage all timber parts together. The on-site workers want the right timber at the right time, so there is a need for the packing order of the wood transportation package according to the on-site use sequence.

At the same time, due to the cost of transportation, it's also needed to consider the stability and occupied space of the package.

1.2. RESEARCH PURPOSE

In this research, according to these complex needs, not only developing an algorithm for packing irregular timbers automatically is needed, but also need to generate, analyze and evaluate different kinds of packing results so that the user can judge and select a proper proposal by balancing the actual on-site needs.

In order to ensure that the algorithm can be used in real construction, it is impractical for it to operate for several hours to find the so-called perfect one among thousands of packing methods. It should calculate the approximate optimal schemes in the shortest possible time and give the evaluation results based on different indicators, which are presented to the user through MR equipment in a visual form. The user can choose the appropriate scheme according to the needs of the ongoing construction and use MR equipment for construction navigation like people choose a suitable route from the feasible ones provided by the map navigation software.

2. Methodology and Technology

2.1. INTRODUCTION TO BASIC CONCEPTS

The Reference Point: It can be selected at any endpoint of the main wood part. Then the translation and rotation of the whole part are positioned based on the Reference Point.

Nesting Points: Nesting points are discrete points that are placed evenly within a container at a certain horizontal and vertical distance. When the packing algorithm starts to run, wood parts to pack will be translated form the Reference Point to the Nesting Points, also rotate based on the Nesting Points.

Nesting position: The wood moves to a nesting point based on the reference point, then rotates at an Angle around an axis.

Principle of minimum total potential energy: The principle of minimum total potential energy belongs to the category of physics, which means that when the potential energy of a system is the smallest, the system will be in a stable equilibrium state. The total potential energy (Π) is the sum of the elastic strain energy (U), stored in the deformed body and the potential energy (V), associated to the applied forces: $\Pi = U + V$. Since it is applied to the 3D packing problem, the elastic energy U is zero. And the V only related to the gravity of the packed wood parts. $\Pi = V = G*Z$, G is the force due to gravity and Z is the vertical coordinate of the center of gravity. Wood parts always try to reduce the height of their center of gravity through translation and rotation, so as to get closer and stable arrangement.

2.2. THE PACKING ALGORITHM

The Packing Algorithm includes two parts: The Feasible Placement Algorithm and The Proximity Algorithm.

The Feasible Placement Algorithm takes the packing order of timber parts as the variable. Parts are loaded into the container in a certain order (the height of the container is infinite and the bottom is rectangular), and guarantee no overlap between parts. Then the Proximity Algorithm starts to work to guarantee timber parts get as close as possible. Finally, after these two algorithms are executed successfully the packing result is obtained.

2.3. THE PACKING ALGORITHM WORK FLOW

1. Simplify the outline of the timber to be loaded. The metalwork model has a complex shape, direct calculation without appropriate simplification will significantly increase the calculation time.

2. At the same time, select the reference point for the translation and rotation later. Through the Grasshopper plug-in Center of Mass, wood and metal parts are endowed with different densities to find the center of gravity of the whole part.

3. Specify a packing sequence for all timber. Usually, sort the packing order according to the order of use during construction firstly to guarantee the best construction efficiency.

4. Set the length width and height in Grasshopper to create a Box object as the packing container. At this time, the height should be set as high as possible, so that the final stacking height of various package schemes will not exceed the height of the container.

5. Give appropriate nesting points distances in the X, Y, Z directions according to the size of the wood, and arrange multiple nesting points uniformly inside the container.

6. Start to run the Feasible Placement Algorithm. Translate the current timber part to each nesting point based on the reference point, and rotate around three coordinate axes respectively.

7. Calculate the height of the gravity center of the current part with different rotation angles at each nesting point, and find the optimal feasible layout attitude (in this attitude, the height of the gravity center of the packing part is the lowest, and does not overlap with the container and other parts already put in).

8. Since the nesting points are discrete, there would be a gap between the wood currently being packed and the woods already packed. The Proximity Algorithm can be used to make the currently packing wood close to the packed woods and container by moving vertically and horizontally.

9. Perform steps 6-8 for each wood until all woods are packed, then the Packing algorithm ends and the packing results is obtained and recorded.

10. Adjust the packing sequence of woods, and repeat steps 6-9 to get a new packing result. It can be repeated many times to get multiple groups of packing results.

11. Use the evaluation system created in Grasshopper to conduct multi-dimensional evaluation on each group of packing results and display the evaluation results.

2.4. THE FIVE EVALUATION FACTORS

The five evaluation factors include: Offset of Gravity Center, Top Surface Roughness, Package Volume, Valid Space Ratio, Packing Order Confusion.

1. Offset of Gravity Center: The system will find the actual gravity center of the package, and compare it with the ideal gravity center (volume center of package) to find the horizontal offset distance. The smaller the data, the smaller the gravity center horizontal offset of the package, and the more stable it is during transportation.

2. Flatness of Top Surface: When multiple packages are stacked vertically, the packages above need to be placed on a flat surface to ensure stability. Usually, some filling materials such as wood bar can be filled above the completed package to flatten the upper surface. If the original top surface is flat, less materials will be used to fill and improve the operation efficiency.

3.Package Bounding Volume: This result collects the Bounding volume data of the package to measure the volume that the whole package will occupy in actual transportation.

4. Valid Space Ratio: This result represents the space occupied by the wood parts in the package. Besides the top surface, gaps in the package also need to be filled with tools such as wood bar, the larger the valid space, the less filling objects are required.

5. Packing Order Confusion: This result indicates whether the wood in the package is loaded in the order required for construction. The larger the value, the less the construction efficiency value become.

3. Experiment

3.1. EXPERIMENT GOAL

1. Check the running results of the algorithm, whether it can successfully calculate the packing results under different packing orders.
2. Whether the evaluation system runs successfully so that user can get the information of each result.
3. Since nesting point density has a significant impact on the calculation steps and calculation time. Test the difference of results under different nesting point density.
4. Test the gap filling function of the experiment, so as to make better use of it during construction.

3.2. EXPERIMENT METHOD

This experiment will pack four types of timber parts. Under different experimental conditions, the number of each kind of timber part remains the same, but the overall packing order and the density of nesting points will be adjusted and changed.

In this experiment, as shown in figure 2, it is assumed to pack 7 pieces of timber Type.1 (length: 860mm), 7 pieces of timber Type.2 (length: 1380mm), 4 pieces of timber Type.3 (length: 2075mm), and 2 pieces of timber Type.4 (length: 2960mm), a total of 20 pieces of timber.

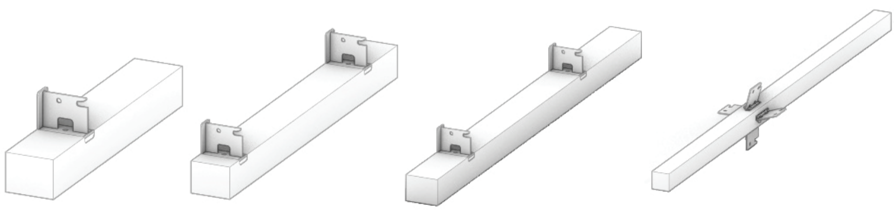


Figure 2. Timber Type.1 to Type.4 (left to right)

It is assumed that the packing sequence that can achieve the best construction efficiency is the sequence shown in the following figure and according to this best construction efficiency packing sequence, renumber the woods shown in the picture from left to right as 1 to 20 (Fig. 3). The best construction efficiency packing sequence is: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20.

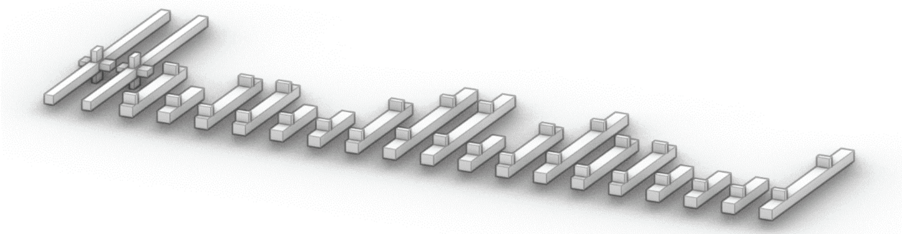


Figure 3. The best construction efficiency packing sequence

The experiment sorts the packing sequence in 10 groups, with each sort tested at two types of nesting points with different density: low and high density (Fig. 4).

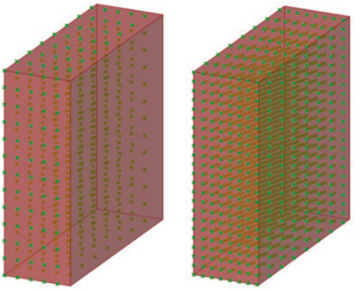


Figure 4. The low and high density of the nesting points

And a special experiment is also conducted. The longest timber No.4 is cut into four equal length parts to detect that with the same total timber volume (Fig. 5), the algorithm can use these small timbers to fill more gap areas.

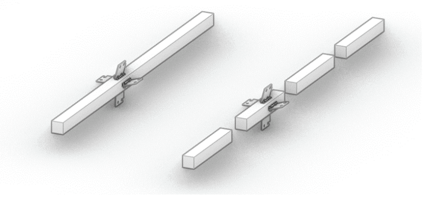


Figure 5. Timber Type.4 was cut into four equal length parts

4. Data Analysis

4.1. CONTROL EXPERIMENT

Different timber packing results will be calculated for the same packing order at different density of nesting points (Fig. 6). In the following analysis of experimental data, experimental differences based on density of nesting points in 10 groups of packing sequences will be shown. For example, Experiment 1-1 indicates results obtained in low nesting points density with the first type of packing order, while Experiment 1-2 indicates results obtained in high nesting points density with the same first type of packing order.

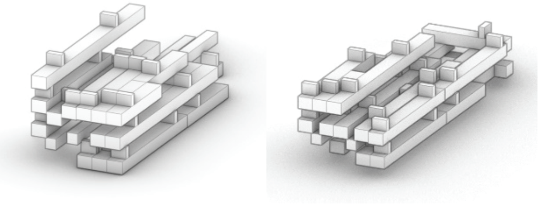


Figure 6. Packing difference with low nesting point density (left), high nesting point density (right)

4.2. ANALYSIS OF THE OFFSET OF GRAVITY CENTER

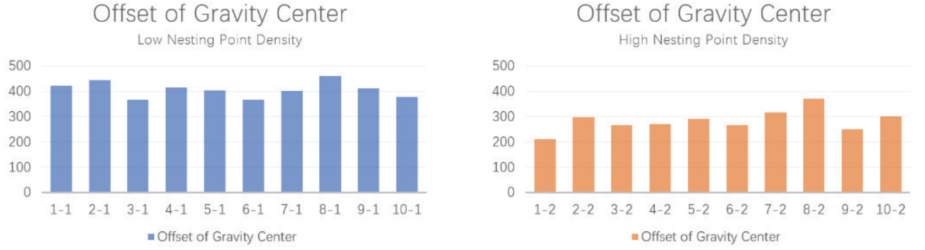


Figure 7. Result of the offset of gravity center (unit: mm)

4.3. ANALYSIS ABOUT THE TOP SURFACE ROUGHNESS

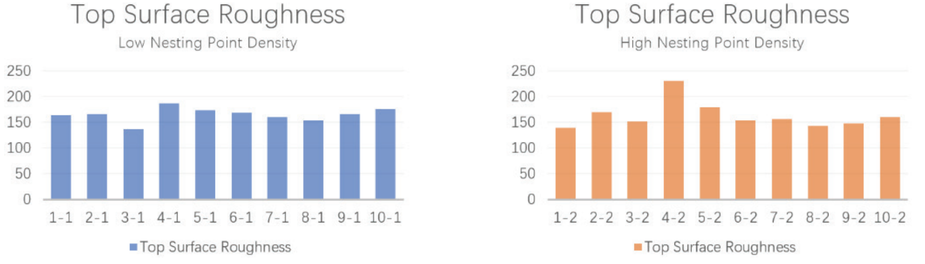


Figure 8. Result of the top surface roughness (unit: mm)

4.4. ANALYSIS ABOUT THE PACKING VOLUME

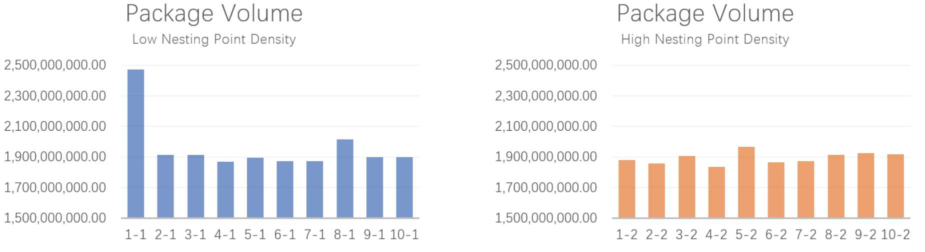


Figure 9. Result of the packing volume (unit: mm^3)

4.5. ANALYSIS ABOUT THE VALID SPACE RATIO

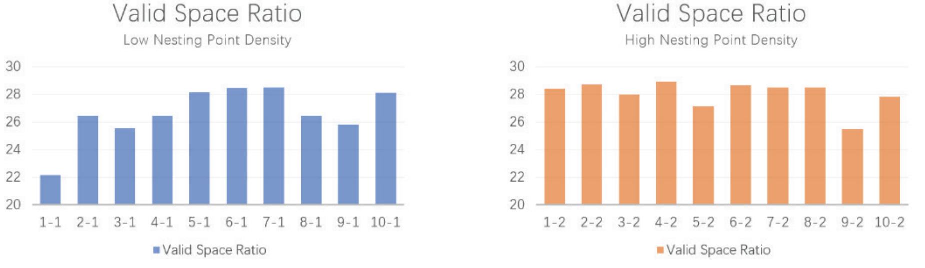


Figure 10. Result of the valid space ratio

4.6. ANALYSIS ABOUT THE PACKING ORDER CONFUSION

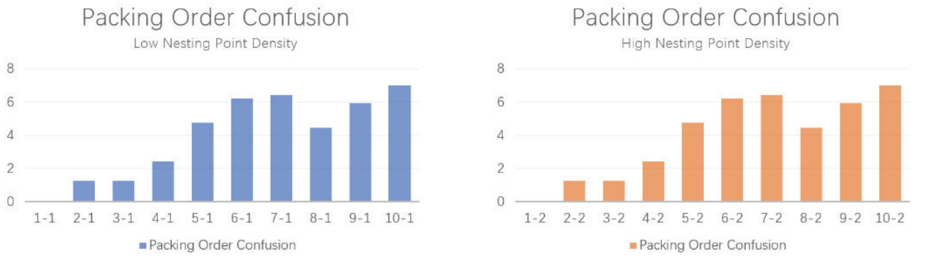


Figure 11. Result of the packing order confusion

4.7. ANALYSIS ABOUT THE AVERAGE CALCULATION TIME

This algorithm is designed to pack the timbers one by one according to different packing orders and nesting point density. The number on the X axis indicates the time to finish calculating the packing of this timber from the beginning of packing the first timer. For example, "10th" represents the time to finish calculating the packing location of the tenth timber from the algorithm starting to packing the first one.

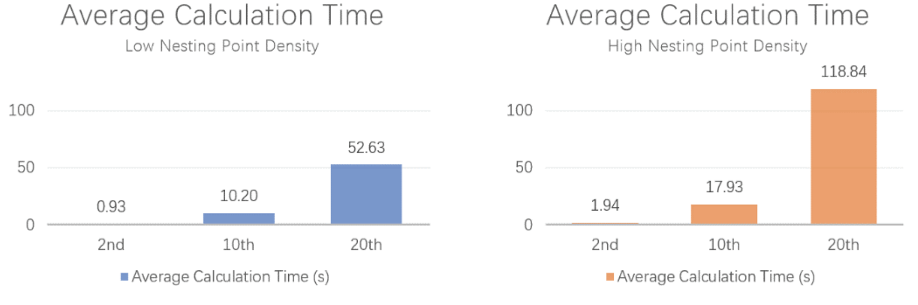


Figure 12. Result of the average calculation time (unit: second)

4.8. ANALYSIS ABOUT ALL THE PACKING RESULTS

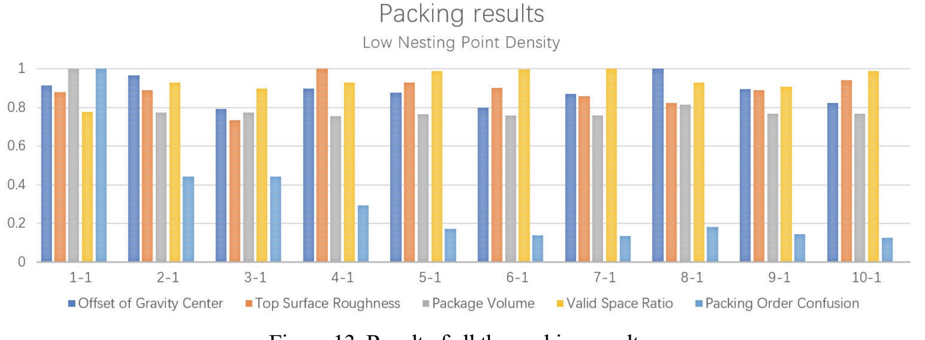


Figure 13. Result of all the packing results

4.9. ANALYSIS ABOUT THE SPECIAL EXPERIMENT

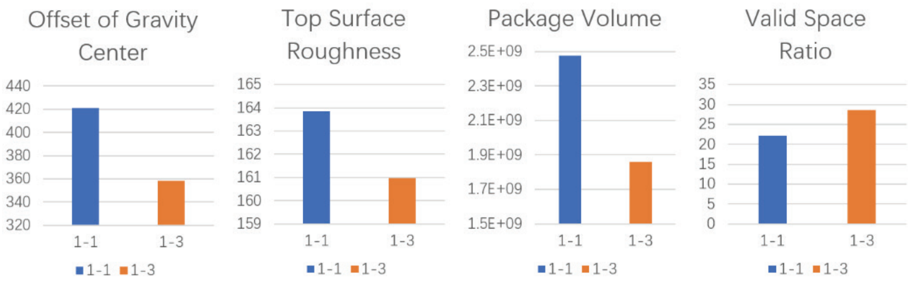


Figure 14. Result of the special experiment (1-3 is the results of special experiment)

5. the Experiment Conclusion

The algorithm can successfully calculate the packing results under different packing orders. From the experiment data, taking the packing order of wood as a variable, changing the packing order will affect the final packing result and the evaluation indicators. And the results of each evaluation index will fluctuate significantly according to the packing order.

5.1. ANALYSIS ABOUT PACKING CALCULATION TIME

As shown in the Figure 12, the calculation time of this algorithm will increase significantly with the increase of the number of timbers to be packed. This is because as the number of packed wood increases, the computational complexity of subsequent timber will significantly increase when it needs to detect the collision with the already packed timbers. In the actual packing work, reasonably controlling the wood packing quantity of each package can significantly improve the running speed of this algorithm, and improve the efficiency and user's feeling of use.

5.2. ANALYSIS ABOUT THE INFLUENCE OF NESTING POINTS

The Figure 7, 8, 9, 10, 11 shows the density of nesting points has a significant impact on the calculation time and packing results. It is obvious from the results that the increase of nesting points density will significantly increase the calculation time, because with the increase of nesting points, the steps of timber translation and rotation to find their best place will also increase. However, it will also significantly improve the experimental results, because the more steps to calculate the best location can lead to more precise result data.

5.3. ANALYSIS ABOUT THE FUNCTION OF GAP FILLING

As shown in Figure 14, this algorithm has the function of gap filling. In the special experiment, it can be seen that the space utilization rate and the packing volume in the experimental results are significantly improved. This is because after decomposing the large parts, the four small parts can be fully packed in the gap space of the package. Therefore, in the actual packing operation, if there are more small woods, the final packing result will usually have higher space utilization and smaller volume.

5.4. THE NECESSITY OF USER'S CHOICE

Due to the complex needs in reality, this algorithm is not designed to only show a best solution. As shown in Figure 13, the most important thing is to provide enough information for users to let them balance the needs and choose a proper proposal.

6. The Discussion of The Future Implementation and Improvement

This experiment will be the intermediate point connecting the whole IST project. The packing algorithm in this study combined with MR Navigation system designed by other members of IST team also based on Rhino and Grasshopper environment can make the whole construction navigation system work smoothly.

And in the future work, the IST team will complete the visualization function with MR devices to guide users' operation more directly (Fig. 15).

Figure 15. The future visualization navigation function

References

He, K., Huang, W. (2011). An efficient placement heuristic for three-dimensional rectangular packing. Computers & Operations Research, 38(1), 227–233. https://doi.org/10.1016/j.cor.2010.04.015

Liu, X., Liu, J.-min, Cao, A.-xi, & Yao, Z.-le. (2015). HAPE3D—a new constructive algorithm for the 3D irregular packing problem. Frontiers of Information Technology & Electronic Engineering, 16(5), 380–390. https://doi.org/10.1631/fitee.1400421

Ma, Y., Chen, Z., Hu, W. and Wang, W. (2018), Packing Irregular Objects in 3D Space via Hybrid Optimization. Computer Graphics Forum, 37: 49-59. https://doi.org/10.1111/cgf.13490.

Saraiva, R. D., Nepomuceno, N. &Pinheiro, P. R. (2015). A layer-building algorithm for the three-dimensional multiple bin packing problem: A case study in an automotive company. IFAC-PapersOnLine, 48(3), 490–495. https://doi.org/10.1016/j.ifacol.2015.06.129

Wauters, T., Verstichel, J. Vanden Berghe, G. (2013). An effective shaking procedure for 2D and 3D strip packing problems. Computers & Operations Research, 40(11), 2662–2669. https://doi.org/10.1016/j.cor.2013.05.017

Wu, H., Leung, S. C. H., Si, Y.-whar, Zhang, D., & Lin, A. (2017). Three-stage heuristic algorithm for three-dimensional irregular packing problem. Applied Mathematical Modelling, 41, 431–444. https://doi.org/10.1016/j.apm.2016.09.018

INFORMING USER-CENTRED APPROACHES TO AUGMENTED CUSTOM MANUFACTURING PRACTICES

ANTHONY FRANZE[1], GLENDA CALDWELL[2], MUGE BELEK FIALHO TEIXEIRA[3] and MARKUS RITTENBRUCH[4]
[1,2,3,4] *Queensland University of Technology (QUT).*
[1]*a.franze@qut.edu.au, 0000-0002-9099-455X*
[2]*g.caldwell@qut.edu.au, 0000-0003-0837-9310*
[3]*muge.teixeira@qut.edu.au, 0000-0003-0593-9597*
[4]*m.rittenbruch@qut.edu.au, 0000-0001-9279-1599*

Abstract. This practice-based research presents insights into the potential and challenges for augmented and mixed reality (AR/MR) technology to enhance Australian small-to-medium (SME) custom manufacturers' agility to overcome existing Industry 4.0 (I4.0) workforce productivity and efficiency challenges. Moreover, it seeks to understand the technology's ability to support custom manufacturers and the architectural, engineering and construction (AEC) sector transition to a more human-centric Industry 5.0 (I5.0) model, whereby the well-being of the fabricator is placed back at the centre of manufacturing processes. This qualitative study draws on interviews with eleven Australian custom manufacturing industry professionals to inform pertinent themes around fabricators' current use and perceptions of mixed reality technology. Results indicate benefits for fabricators in reducing 2D drawing and task-related ambiguities in fabrication and assembly practices and reveal factors surrounding underutilisation. Synthesising insights and reflecting on Teixeira et al., (2021)'s XR-PACT framework, key research areas are identified for future AR/MR development centred on fabrication users' distinct needs to improve accessibility, empower fabricators and ultimately assist the competitiveness of custom manufacturers and the AEC sector.

Keywords. Augmented and Mixed Reality, Custom Manufacturing, SME, User-centred Design, Industry 5.0

1. Introduction

The past decades have seen the architecture, engineering, and construction (AEC) industry and adjacent manufacturing sector experience rapid transformation through Industry 4.0 (I4.0) technologies such as robotics, artificial intelligence (AI), the internet of things (IoT), augmented and mixed reality (AR/MR). However, as computation and connectivity drive further automation and increased process efficiencies, often, the human costs of this optimisation can be overlooked (Nahavandi, 2019). In complement to I4.0, Industry 5.0 (I5.0) seeks to place human worker's well-being back at the centre

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 353-362.

of manufacturing processes, whereby human-machine collaboration extends worker capability, circular practices reduce industries' climate impacts, fostering resilient industries, and improved social conditions (Commission et al., 2021).

The current I4.0 paradigm is already challenging Australian custom manufacturing small-to-medium-sized (SMEs) who require support to counter severe skills shortages and increased agility to deliver high-quality products for evolving consumer preferences (DISER, 2020). With further production on-shored due to COVID-19 supply chain disruptions, lack of access to fabricators limits Australian manufacturers' capacity causing significant delays across sectors, including AEC. As such, industry peak body Weld Australia (2022) forecasts shortages of seventy-thousand welders by 2030 to meet Australia's future manufacturing objectives.

Operating at the core of the Australian AEC industry, custom manufacturers' workforces are hampered by challenges in fabricating and assembling increasingly complex products using limited 2D drawings and unintuitive labour-intensive manual processes. In response, some Australian custom manufacturers have embraced the assistive qualities of AR/MR technologies to improve their fabricators' productivity and ultimately contribute to the organisation's overall competitiveness. However, despite the significant benefits of AR/MR, SMEs can be reluctant or lack the time, resources and expertise to explore the technology's integration, contributing to low AR/MR adoption rates (AMGC, 2022).

In this paper, AR/MR is used interchangeably as spatially anchored virtual content, manipulated through natural gesture inputs afforded by the Microsoft HoloLens 2 see-through optical display, as defined in related publications (Liu et al., 2021).

This research seeks to understand how AR/MR technology can benefit custom manufacturers and the reasons preventing further adoption. In doing so, we foster insights around the human-centric theme of CAADRIA 2023 about 'the practical and real-world application of AR/MR technology to advance broader social well-being' by empowering fabricators and enhancing their workplace environments.

2. Toward Industry 5.0

As I4.0 innovation continues unabated, human workers can be left out of the loop in advanced manufacturing systems or unable to adapt to the evolving skillsets required—profoundly changing workforce dynamics and industry's role within society, with implications for broader society (Commission et al., 2021). To address these concerns, I5.0 proposes to extend I4.0s positive aspects by leveraging human ingenuity and machine efficiencies (Maddikunta et al., 2022). Whereby the human needs of the worker are at the core of sustainable processes, emergent technologies adapt to workers' needs and respect their fundamental human rights (Commission et al., 2021).

Increasingly I4.0 transformation impacts the fabricators at the heart of custom manufacturing practices producing low-volume and bespoke products such as architectural elements, facades and large-scale artworks. Increased computational power allows designers to create increasingly sophisticated parametric 3D computer-aided (CAD) designs. However, fabricators still rely on interpreting complex 3D digital designs using limited 2D drawings to materialise intricate 3D artefacts. Managing complex digital design data in this inefficient manner bottlenecks production and is among SMEs' and the AEC industry's greatest challenges (Côté et al., 2013). Moreover, fabricators often materialise complex artefacts using labour-intensive manual machine processes. These machines rely on the operators' tacit skills as they lack the assistive visual feedback that expensive computer numerically controlled (CNC) machines provide. Although CNC and robotics can be beneficial, the financial, process and human implications may not align with SMEs' current business strategy.

This paper addresses the following research question: "How can Australian custom manufacturers integrate AR/MR technology in their fabrication and assembly processes?" In exploring this question, we seek to understand the experiences of industry experts who have integrated AR/MR and those exploring its use to guide future research agendas which extend the technology's benefits to further SME workforces. We propose that AR/MR technologies are valuable for custom manufacturers as they enhance human workers' agency and productivity and contribute to more enjoyable working environments. AR/MR has been shown to facilitate positive outcomes for fabricators by improving task cognition and reducing reliance on 2D drawings, markedly lessening error, re-work and waste (Jahn et al., 2018) while synergising human and human-machine collaboration (Hahm et al., 2019). All critical factors to consider in moving to a more human-centric I5.0.

AR/MR technology has been researched across many facets of AEC (Davila Delgado et al., 2020) and manufacturing (Bottani & Vignali, 2019), addressing aspects such as prototyping, training and assembly. However, AR/MR use has been limited in practice by factors such as accessibility, user interface design, accuracy and latency (Cheng Jack C. P. et al., 2020). In the following sections, we describe the methods, findings and reflections on Teixeira et al., (2021)'s XR-PACT framework, identifying key areas for future research to focus AR/MR development on fabricators' requirements in response to the sentiments, needs and experiences of industry experts.

3. Methods

This paper reports on eleven semi-structured interviews conducted between February to June 2022 with custom manufacturing professionals from five Australian SMEs across AEC and complementary industries, summarised in Table 1.

Table 1. Summary of interview participants, accompanying SME, job role and

Participant ID	SME ID	AEC background	Role	Interview Format
1	A	AEC consultants in manufacturing, design development and management	Technical Supervisor	Virtual
2	A	AEC consultants in manufacturing, design development and management	Head of Fabrication	Virtual
3	B	AEC consultants in manufacturing, design development and management	R&D Manager & Robotics Engineer	Virtual
4	C	Custom manufacturing consultant	Senior Engineering Supervisor	Face-to-face
5	C	Custom manufacturing consultant	Project Development Manager	Face-to-face
6	C	Custom manufacturing consultant	Mechanical Engineer & CAD Drafter	Face-to-face
7	C	Custom manufacturing consultant	Mechanical Engineer & CAD Drafter	Face-to-face
8	C	Custom manufacturing consultant	Head of Fabrication & CAD Supervisor	Face-to-face
9	C	Custom manufacturing consultant	Managing Director & Chief Engineer	Face-to-face
10	D	AR/MR software development company/integrator	Co-founder & CEO	Virtual
11	E	Custom manufacturing consultant	Project Manager & Lead Engineer	Face-to-face

interview format.

Participants ranged in experience from 4-30 years. They included custom manufacturers' technical staff responsible for developing AR/MR responses, such as engineers, designers, CAD drafters and supervisors, managers solely responsible for the technology's adoption, software developers/integrators and lead fabricators responsible for undertaking and coordinating fabrication and assembly tasks. Noting that P4-P9 from SME C were exploring AR/MR adoption, whilst all other participants were experienced users of AR/MR technologies.

The interviews employed open-ended questions to understand users' experiences and perceptions of AR/MR technology. Examples of questions employed include: "What are some of the key learnings to adopting and integrating new technologies such as AR/MR in your design and manufacturing processes?" and "What are some of the main challenges you face when implementing new technologies?

3.1 ANALYSIS

A reflexive thematic analysis (Braun & Clarke, 2021) was used to interpret and analyse patterns across the qualitative dataset to develop themes based on the analysis of audio-recorded transcripts. Initial open codes were tabulated in NVivo software with refinement phases culminating in visual clustering of higher-order themes through affinity mapping.

In the results section, we outline findings from the reflexive thematic analysis by referring to participants' (P…) quotes around significant themes.

4. Results

In revealing Australian custom manufacturing professionals' experiences with AR/MR, the interviews raise significant questions about SMEs and the AEC sector's future adoption of the technology. For example, according to a software integrator, AR/MR is embraced as a valuable tool in fabricators' "toolkit to produce some high-quality goods" (P10), providing useful information during fabrication and assembly tasks, with fabricators reporting improved agency and confidence. However, fabricators exploring AR/MR adoption felt stymied by technical staff and management uncertain about the necessity for the technology's adoption.

4.1 FABRICATORS WELCOME AR/MR USE, BUT TECHNICAL STAFF ARE APPREHENSIVE

Among non-AR/MR users, some technical staff felt fabricators required considerable convincing of the technology's feasibility and "proving that it would work" (P6), anticipating sentiments of "this sort of stuff is futuristic, Sci-Fi… not in my lifetime" (P4). However, in discussing SMEs' experiences integrating the HoloLens, fabricators' responses appeared to be typically positive, "once they put it on, they can immediately understand how it is going to assist them in their fabrication approaches" (P1). In particular, "HoloLens is that one piece of technology that fabricators on the floor have always positively engaged with" (P1) as it "enhances worker capability across all skill levels" (P10). Improving the fabricator's capability in custom manufacturing is critical due to the prototype often being the product sold. As an experienced fabricator stated, “We have got one shot, and these glasses give us the best shot” (P2).

In contrast, management and technical staff appeared apprehensive about the technology correlating with P1s’ experience “usually, there is a real positive association with people who aren’t programming the technology”. According to AR/MR software developer and integrator P10, SMEs desire turn-key and off-the-shelf solutions and do not want to “do the time investment”. Currently, effective use of the technology requires on-site development support to iterate and adapt real-time responses to nuanced production processes, “if you aren’t able to adjust your workflow… suddenly, it becomes quite useless to you” (P1). In addition, while the initial device cost is relatively low, there is the “investment in someone actually to manage that and run it” (P1).

SMEs may also need to factor in issues around interoperability with existing CAD systems and potentially upgrade their digital infrastructure to leverage the technology’s potential. For example, custom fabricators may lack certain digital skills, as “often … the digital information that the manufacturers have… it’s not enough” (P10). Moreover, according to P11, technical staff may have difficulty or resist learning new applications due to feeling “it’s not in their best interest personally”. Fabricator P2 suggested management staff could fear AR/MR adoption due to their staff’s better understanding of the technology and not being “on top of it” enough to manage it.

4.2 AR/MR helps reduce numerous ambiguities fabricators face in the workshop

Discussions with participants highlight that the descriptive limitations of 2D CAD drawings cause miscommunication between designers and fabricators, who also contend with numerous task-related ambiguities to materialise complex artefacts. The following sections explore how fabricators identify AR/MR as a powerful tool for reducing these ambiguities.

4.2.1 AR/MR Removes 2D Drawing Ambiguities

AR/MR technology provides fabricators with at-scale, contextual visual information, improving comprehension of complex 3D CAD geometry; as P1 explains, “there’s no way you’re going to draw it and understand it better than just looking at it in AR” (P1). Using AR/MR, fabricators can overlay 3D content to compare virtual objects with their physical counterparts, which “shows you where there are any clashes… before we have to deal with it” (P2). Fabricators found virtual cross-references improved their confidence to move projects forward, as traditional 2D drawings did not alert fabricators to accumulating errors, leading to costly downstream impacts, whereas AR/MR “takes out that guesswork” (P2). An integrator explained how AR/MR helps fabricators by providing “digital information, in situ, on location, in reference to where you are (P10)” in doing so, it makes “the work environment… more enjoyable… by reducing that ambiguity and making workers’ lives easier” (P10).

Participants indicated AR/MR’s ability to reduce ambiguity as instrumental in fabricating amorphous and highly irregular artefacts, which can be difficult to describe given 2D drawings’ limited descriptive abilities. Resulting in fewer chances of fabricators “misinterpreting complex drawings” (P1) and avoiding the following scenario an integrator described, “the shop drawings didn’t really show me… I went and made it… the designer was like that's not what I meant" (P10). AR/MR allows fabricators to interpret better what is intended, spot issues earlier, and communicate concerns more fluently to design teams. Improving communication with design teams can empower fabricators, help them voice insights, and bridge transdisciplinary understanding. In AR/MR, fabricators can walk designers through fabrication issues in context, explain why certain design aspects are unachievable and propose alternate fabrication approaches. For example, P10 described how fabricators had walked designers through a particular issue, "that is going to be too weak of a joint. I am not going to be able to do that" as a result, designers developed a timely design response.

Collaborative AR/MR experiences can also help reduce ambiguities within fabrication teams by assisting collective problem-solving, as it is "easy to workshop how you're going to build something" (P1). Within a shared AR/MR environment, fabricators can better impart their tacit knowledge to co-workers, as it is "almost impossible to instruct someone … with traditional approaches" (P2). Also, AR/MRs' ability to synchronise multiple workers in the construction of artefacts can increase the "speed that you can produce things dramatically" (P10).

The fabricators using AR/MR in this study found that "AR is really good for getting things about right… good for approximation… it is not a precision-based tool" (P1). As such, the technology has been embraced by SMEs in the AEC industry, working with wide tolerances, where any improvements in tolerance through AR/MR are "hugely beneficial" (P1).

4.2.2 AR/MR Reduces Manual Process Ambiguities

Using AR/MR helped fabricators reduce laborious manual markup and set-out of difficult-to-datum curved-shaped objects, resulting in improved worker confidence due to less time spent double-checking, "It gets rid of all the squares, the levels, the laser lights, the plumb bobs; it is amazing" (P2). One task typically requiring an entire day now only takes P2 forty-five minutes to complete, with a "higher level of accuracy".

Fabricator P8 expressed a desire to adopt AR/MR to realise similar benefits for their practice, "to be able to mark something up quickly, double-check it… speed the process up immensely". For example, manual press brake machine limitations made it difficult for experienced fabricators to fold complex curve-shaped artefacts "if I'm doing something straight, it is fine… if it's not square, this one doesn't excel at it" (P8). Moreover, P8 had difficulties recruiting skilled fabricators to fold sheet metal parts, impacting the organisation's approach to product design "we try and avoid large milling jobs, we've got a lathe and mill, but only one person (P8) can operate them" (P5). As such, P8 suggested AR/MR is needed for training apprentice fabricators, "I can show them a couple of times… give them a digital twin of the overlays, more confidence, quicker turnaround, and quicker training".

P8 also wanted to improve the entire sheet metal folding process, beginning with manual scribing operations, a time-consuming, error-prone task relying on a "measure twice, cut once" philosophy. In contrast, P8 proposed using AR/MR to replace manually scribed lines at given fold locations with 2D virtual lines. Also suggesting sheet metal parts and the machine could have attached QR and ArUco markers to allow the HoloLens 2 device to track part orientation, position, and fold progress in relation to physical press brake operations. There was also a need to prototype and test part folding on the fabricator's machinery, which consumed large amounts of wasted material and time, "have buckets of test pieces… just got test piece, test piece, test piece" (P8). Finally, P8 sought an AR/MR overlay simulating the entire fold sequence in-situ to help fabricators pre-empt task requirements and use the machine more confidently, as the manual press brake offers workers little visual feedback or indication of progress during tasks, "it is the feedback… if I can get that feedback right here while I'm doing it…I would be keen to try it".

4.3 tHE use CASEs FOR AR/MR INTEGRATION are NOT CLEAR

While there is evidence that AR/MR is used for fabrication and assembly tasks (in the AEC sector), our participants reported that there are uncertainties regarding use cases for AR/MR in custom manufacturing processes, "no one knows what they are doing with the technology… no one knows the best practices for it…the best use cases for the technology" (P1). However, an area where AR/MR technology appears to excel at is assisting fabricators in assembling complex objects by providing easy-to-follow instructions "they build the entire apartment in a factory… ship it out, and like Lego bricks, stack it up" (P10). Assembly instructions are common to most projects and follow a similar methodology, "the most straightforward application… assisting the assembly of complex objects is easy to see the value in many projects" (P1).

In discussing the technology's benefits for fabrication tasks, AR/MR responses seemed best suited to highly repeated processes as significantly more development was required than augmenting assembly processes. Depending on the complexity of the fabrication process, expert teams could take several months to develop a suitable AR/MR response, "and after months of development, the AR application is not used" (P1). AR/MR fabrication approaches appeared most useful when they can be easily adapted to similar tasks "they're the same process where it's shipbuilding, building a railway car, you're all bending steel, they're welding things together… they all hit the same problem" (P10). SMEs can also experience distinct advantages by developing augmented fabrication approaches. For example, "if there is a modular construction company, there would be enormous benefits for them to develop a particular fabrication workflow for something they continually repeat" (P1).

4.4 tHE BUSINESS CASE FOR AR/MR INTEGRATION IS UNCLEAR

Given the limited knowledge around AR/MR use cases, participants expressed doubts about adopting the technology, posing questions such as "what's the return on investment" (P4) and "how can it be monetised" (P9), and "it's the justification… quantifying the ROI… that is the tricky part" (P4). This lack of clarity around the technology's use is apparent when talking to technical staff, "I feel it takes a little while before it's accepted and…needed" (P6) and "it's new, people don't have an understanding of the business model" (P9). The ambiguity around the AR/MR business case made getting management buy-in for the head fabricator impossible. P8 explained that integration was stymied largely due to management's previous negative experiences with emerging technologies "so many either over promise and under deliver…fail or they're just not there yet".

Fabricator P2, explained that companies may be unsure whether to invest resources now or "wait till the technology is better… we're going to spend all this money, and it's not going to be 100%". P11, who implemented the first digital transformation of an Australian custom manufacturing SME from 2D paper-based workflows, agreed that it is difficult for SMEs to adopt AR/MR technology when there is ambiguity around the use case, as "people just don't want to wrap their head around it" and suggest SMEs require further research and use case examples as "it helps to create pathways, …makes it easier for the industry to get on board".

AR/MR adoption is perceived as a risky investment for SMEs compared to larger enterprises; as P2 explains, "the R&D… is not as big an impact for large companies as it is for small or medium companies", along with the price point of current devices, "I think the price is going to be the biggest killer in the smaller market". In addition, technical staff expressed concerns about headset use on the factory floor and on-site due to potential health and safety risks around visual distractions. Whilst fabricators experienced issues around latency "as it can be glitchy, can get drift sometimes" (P2). Management staff also expressed concerns about how the technology will be supported to avoid extended downtimes and avoid the following scenario, "here are the glasses…good luck…we'll see you in six months when you've smashed them (P9).

5. Discussion and Conclusion

In discovering Australian custom manufacturing experts' challenges and opportunities in adopting AR/MR technology, the themes identified in this study raise important questions on how future AR/MR development can make the technology more accessible to fabricators in the manufacturing and AEC sectors. While technical obstacles with current-generation AR/MR devices require further development, discussions indicate significant cultural barriers contribute to low adoption rates. SMEs require further knowledge of AR/MR use case benefits to inform prospective business models that consider the unique and complementary nature of AR/MR technology to bridge fabricators' information and communication gaps in practice.

Fabricators identified AR/MR as a valuable tool for reducing ambiguity in their fabrication and assembly activities, reporting increased task confidence and improved work lives. AR/MR was perceived as less threatening to livelihoods than other I4.0 technologies, such as robotics; as P10 stated, "I am not worried about losing my job because someone still needs to wear the headset". However, as an emerging technology confined to small pockets of use, custom manufacturers' management and technical staff were uncertain about how to integrate the technology and whom to seek integration assistance from. SMEs require support to equip their workforce and acquire the necessary skill sets to implement the technology as, according to fabricator P2, AR/MR will be demanded by "the young people who are coming up through the trades now…they are really excited by this technology".

Based on findings and moving toward an I5.0 model, a user-centred focus is sought to ensure fabricators do not become passive adopters of AR/MR but rather play an instrumental role in SMEs and the AEC sectors' long-term use of the technology. As such, we consider Teixeira et al., (2021)'s XR-PACT framework employing Benyon, (2018)'s People-Activity-Context-Technology (PACT) analysis to advocate for user-centred extended reality (XR) development approaches in the AEC industry. In doing so, XR-PACT, described in Figure 1, outlines a comprehensive approach to user-centred XR development by considering and making explicit the intertwined relationships comprising the technology's use.

In line with sentiments expressed by participants and reflecting on Teixeira et al., (2021)'s XR-PACT framework, we propose further framework development centred on the AR/MR use needs of custom SME fabricators. As described in the shaded rectangles shown in Figure 1, we outline two prospective areas for consideration. First, we seek an improved understanding of the type of digital information fabricators require during tasks and how to convey best and express this information to enhance their practice. Second, explore further the opportunities and limitations of AR/MR solutions to assist fabricators across various custom fabrication and assembly practices.

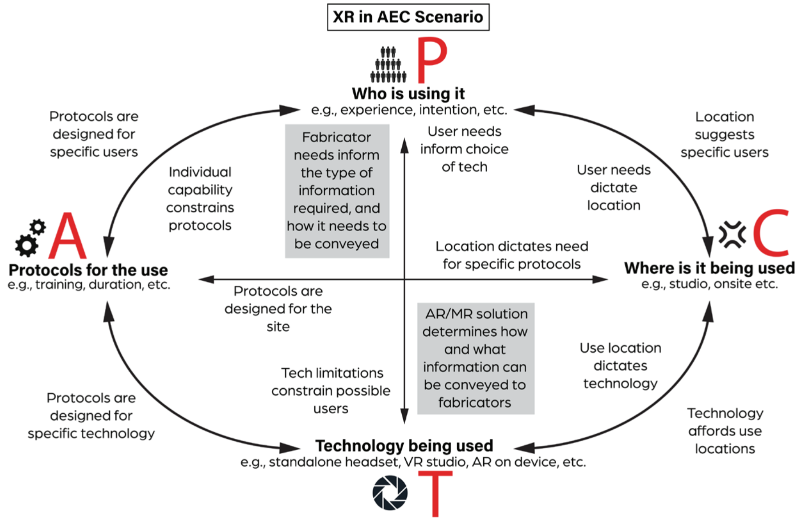


Figure 1. The author's recreation of Teixeira et al., (2021)'s XR-PACT framework. Key areas for future research are described in the shaded rectangles adjacent to Participants and Technology.

Historically, industrial process improvements have tended to spend significant amounts of time refined in the manufacturing sector before filtering through to the AEC sector. However, this study highlights how AR/MR approaches informed largely by AEC practice can also benefit fabricators in the manufacturing sector, outlining to the CAADRIA community the need for improved dialogue and cross-sectoral learnings to enhance fabricator practices and work lives across industry.

The next phase of this study aims to build upon Teixeira et al., (2021)'s user-centred XR-PACT framework in developing an AR/MR PACT framework, informed through a series of AR/MR custom manufacturing workshops with expert fabricators. The workshops will aim to foster insights around the two key areas outlined in Figure 1 and potential areas indicated by fabricators to help guide the development of future AR/MR responses centred on the specific user needs of fabricators.

Acknowledgements

We sincerely thank the participants for their contribution and the Queensland

University of Technology (QUT) School of Architecture and Built Environment for the scholarship assisting this PhD research. We also thank the Advanced Robotics for Manufacturing Hub (ARM Hub) and the QUT Design Lab for supporting our research.

References

AMGC. (2022). *Manufacturing Competitiveness Plan 2022-Transforming Australia from Lucky to Smart*. Advanced Manufacturing Growth Centre Ltd.

Benyon, D. (2018). *Designing User Experience : A Guide to Hci, Ux and Interaction Design*. Harlow: Pearson Education, Limited.

Bottani, E., & Vignali, G. (2019). Augmented reality technology in the manufacturing industry: A review of the last decade. *IISE Transactions*, *51*(3), 284–310. https://doi.org/10.1080/24725854.2018.1493244

Braun, V., & Clarke, V. (2021). *Thematic Analysis: A Practical Guide*. SAGE Publications, Limited. https://play.google.com/store/books/details?id=25lpzgEACAAJ

Cheng Jack C. P., Chen Keyu, & Chen Weiwei. (2020). State-of-the-Art Review on Mixed Reality Applications in the AECO Industry. *Journal of Construction Engineering and Management*, *146*(2), 03119009. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001749

Commission, E., for Research, D.-G., Innovation, Breque, M., De Nul, L., & Petridis, A. (2021). *Industry 5.0 : towards a sustainable, human-centric and resilient European industry*. Publications Office. https://doi.org/doi/10.2777/308407

Côté, S., Trudel, P., Snyder, R., & Gervais, R. (2013). An augmented reality tool for facilitating on-site interpretation of 2D construction drawings. *Proceedings of the 13th International Conference on Construction Applications of Virtual Reality (CONVR), London, UK*, 316=323.

Davila Delgado, J. M., Oyedele, L., Demian, P., & Beach, T. (2020). A research agenda for augmented and virtual reality in architecture, engineering and construction. *Advanced Engineering Informatics*, *45*, 101122. https://doi.org/10.1016/j.aei.2020.101122

Hahm, S., Maciel, A., Sumitiomo, E., & Rodriguez, A. L. (2019). *Exploring the human-material interaction in digitally augmented craftsmanship*.

Jahn, G., Newnham, C., & Beanland, M. (2018). Making in Mixed Reality. *ACADIA // 2018:Recalibration. On Imprecision and Infidelity*, 88–97.

Liu, R., Peng, C., Zhang, Y., Husarek, H., & Yu, Q. (2021). A survey of immersive technologies and applications for industrial product development. *Computers & Graphics*, *100*, 137–151. https://doi.org/10.1016/j.cag.2021.07.023

Maddikunta, P. K. R., Pham, Q.-V., B, Prabadevi, Deepa, N., Dev, K., Gadekallu, T. R., Ruby, R., & Liyanage, M. (2022). Industry 5.0: A survey on enabling technologies and potential applications. *Journal of Industrial Information Integration*, *26*, 100257. https://doi.org/10.1016/j.jii.2021.100257

Make it Happen, The Australian Governments' Modern Manufacturing Strategy. Australian Government. (2020). {Department of Industry, Science, Energy and Resources}.

Nahavandi, S. (2019). Industry 5.0—A Human-Centric Solution. *Sustainability: Science Practice and Policy*, *11*(16), 4371. https://doi.org/10.3390/su11164371

Teixeira, M. B. F., Pham, K., Caldwell, G., Seevinck, J., Swann, L., Rittenbruch, M., Kelly, N., Santo, Y., Garcia-Hansen, V., & Voltz, K. (2021). A user-centred focus on augmented reality and virtual reality in AEC: opportunities and barriers Identified by industry professionals. *Proceedings of the 26th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2021*, 273–283.

Weld Australia. (2022). *2022 Member Survey Results: Results Analysis*.

THE RE-CREATION OF A DEMOLISHED HOUSE BASED ON COLLECTIVE MEMORY

YISI LIU[1], NICK WEBB[2] and ANDRE BROWN[3]
[1]*University of Huddersfield.*
[2]*University of Liverpool.*
[3]*Victoria University of Wellington.*
[1]*Liuyisi1988@gmail.com, 0000-0002-7877-4618*
[2]*Nicholas.Webb@liverpool.ac.uk, 0000-0002-5998-1961*
[3]*Andre.brown@vuw.ac.nz, 0000-0002-0576-5112*

Abstract. This paper investigates how the collective memory of an architectural setting can be captured, re-created, and re-experienced. The method uses a workflow to create a Virtual Environment (VE) from diverse sources in an interpretational and iterative process. A case study was implemented to test this workflow, focusing on the re-creation of a now demolished family home based on information from multiple generations of a family. The workflow's main output is a VE in which the family members can re-experience the house and give continuous feedback. The essence of this workflow is the constant negotiation between remembered space in memory and its digital interpretation in the VE. The output was assessed and refined according to the users' feedback, which is used to loop back to specific stages in the workflow. This process revealed undisclosed histories and enhanced our understanding of this demolished site and its wider context. Moreover, the workflow provides a model for studying conceptually recreated spaces based on subjective sources and provides a user-centred experience in the VE.

Keywords. Demolished Houses, Collective Memory, Interpretation, Virtual Environment, User-centred.

1. Introduction

This paper discusses part of a wider research project which develops a workflow exploring multiple interpretations of textual spaces (Liu, 2022). The term "textual space" usually refers to the world presented in a literary work (Bushell, 2010, p. 48). In this research, the meaning of textual space is extended, referring to the space described in textual form or remembered space that has the potential to be textualised. For instance, the architectural spaces in literary works or a now demolished building exist in collective memory. When studying unbuilt or demolished architecture, researchers can mainly use primary materials such as drawings and photographs and mediating devices, including computer-aided design, to critique and enhance the understanding of these architectural works (Forte & Siliotti, 1997; Novitski, 1999;

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 363-372.

Webb & Brown, 2011). However, the use of subjective sources is underexplored in this area, although discourse has taken place in other related areas, such as literary work's openness for multiple interpretations (Pilz, 2003), reading images and the interpretations (Wu, 1992), and the relationship between a place, fictional re-creation, and the embedded collective memory (Burke, 2014; Eco, 1994; Said, 2000). Eco highlights the "adhesive function" of collective memory, which helps us "extend it back through time" (Eco, 1994, pp. 130-131). For example, we have seen Joyce's Dublin (2008), Pamuk's Istanbul (2006), and Nabokov's Saint Petersburg (1967). More importantly, these examples provide detailed descriptions and sometimes photographs so that these cities can be reimagined and re-experienced. This encourages us to investigate reviving remembered architectural settings which comply with the textual space's extended definition and consist of objective and subjective elements to be explored.

Regarding the interaction between human beings and the computer, user-oriented experience is a key feature in video games on which non-linear and environmental storytelling can be built (Brown, 2020). In addition, game engines have been used in the study of historic sites (Bozzelli et al., 2019) or to develop digital serious games (DSG) for educational purposes (Lorenzini et al., 2015). This gives precedent to develop a method enabling the users to experience the re-created house by freely exploring the VE and interacting with the objects rather than receiving information passively (Affleck & Thomas, 2005).

Consequently, this paper investigates multiple interpretations of a demolished home in China based on descriptions from four generations of the family who lived there for decades. This house – a traditional Chinese courtyard design (the site plan is shown in Figure 1) – was located to the southwest of the historic walled city of Tianjin.

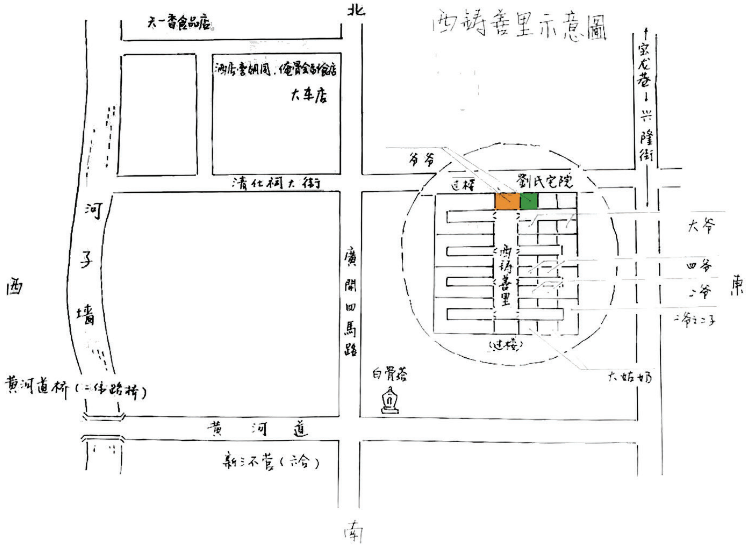


Figure 1. Sketch site plan provided by a family member shows the context before the 1970s. The green box is the single-storey house, and the orange box represents the first floor of the gate building where the family inhabited too.

As part of an urban regeneration project during the 1990s, a large number of historic buildings, many of which were in a poor state, were demolished, resulting in a significant change in the city's texture (Bricoleurbanism, 2010). Although the area's history (including this house) was relatively well known before 1949 (Shao, 1946; Zhang, 1966), what happened between 1949 to 2000 was unclear, even though official records showed that this area became over-crowded, unsafe, and in a state of disrepair on the eve of its demolition (Jianxing, 2000).

As family members do not have an architectural background, developing a methodology to re-create the demolished house in a VE that non-specialist users can experience is crucial, which requires the emergence of presence (Slater, 2009) and intuitive and understandable manipulation in the VE without complex tutorials (Brown, 2015). Moreover, as subjective sources are used, secondary interpretation is acknowledged to express the researcher's translation of the memory collected from the family members. For example, family members' textual descriptions are translated into drawings and digital models; and unlinked places in the description require filling in the gap to achieve continuous exploration in the VE. However, whether the secondary interpretation is exaggeratedly/insufficiently used is the core mechanism triggering the workflow to refine its product via multiple feedback loops.

Consequently, the questions below were raised:

- How can we collect and document the memory?
- How can the resident's interpretation of the demolished house be translated into a VE which can then be re-experienced?
- How can different versions of that architectural space be represented in VE?
- How can the digital representation be refined based on the user's feedback on the experience in the VE?

These questions are addressed using a methodological workflow. As the workflow collects memories as a database for creating digital representations, it requires constant communication with the family members to assess the workflow's output regarding its accuracy, diversity of information, and immersion level. These criteria are key to driving the workflow to review/refine itself based on the hypothesis that the *secondary interpretation* used during the translation process can be incrementally mitigated (but not eliminated), enabling a *user-centred* re-creation and experience of imagined spaces rather than a passive acceptance.

2. Methodology

The workflow shown in Table 1 articulates how the collective memory is extracted and translated into a VE. The workflow has a linear process with multiple feedback loops aiming to refine the VE and generate iterations. It sets out to collect the remembered spaces using a questionnaire, after which the collected data, including textualised memory, is then documented and analysed; this stage addresses question one. This is followed by a digitalisation stage addressing question two, in which the collected data is categorised and translated into digital representation regarding its potential embodiments in the VE. Furthermore, the workflow uses Unity3D to integrate all the translated data into stage four's product – a video-game-like application in which users can explore the re-created house; question three is also answered at this stage. As an incremental process, the workflow's product is experienced by the building users and receives feedback from them, informing the workflow to refine the VE regarding different factors. This refinement requires multiple feedback loops, which are the keys to testing the workflow's validity as the answer to the last question.

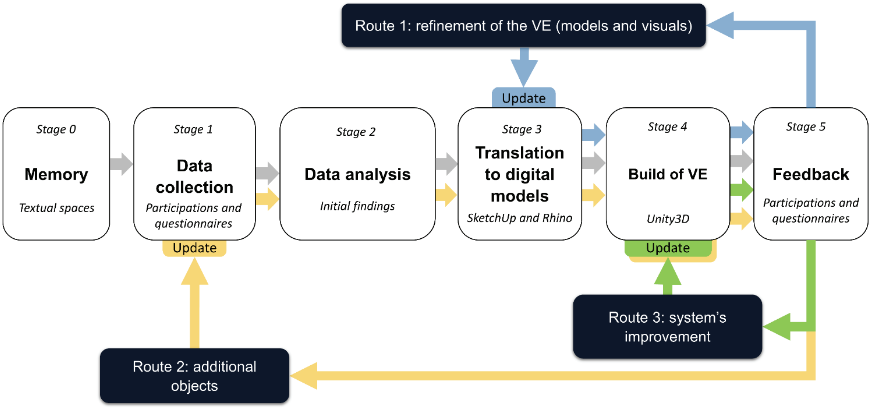


Table 1. The structure of the workflow.

2.1. DATA COLLECTION

A face-to-face workshop was conducted in China to collect data using a questionnaire. The workshop recruited twelve participants from three generations of a family who had lived in this house for five decades before its demolition. They textually described the house based on memory. The questionnaire aimed to help the family members without an architectural background articulate the space textually. Therefore, the questionnaire was split into three sections, from simple questions such as abstract feelings and keywords to complex spatial cognition regarding spatial sequence, dimensions, materials, and multisensory feelings. Additionally, existing photographs were provided to prompt significant locus of the house, and then the family members answered a series of questions to link the locus suggested by the photos and reconstruct the house and context. The questionnaires were finished individually to avoid potential influences from other participants. After all the finished questionnaires were collected, the data was summarised and analysed. Due to the pandemic of Covid-19, the testing and feedback sessions were conducted remotely.

2.2. DATA ANALYSIS

The data analysis aims to use the collective data to reveal key features and events and develop a timeline of this house. They referred to different periods of the house with some overlapped periods; these pieces of memory helped us complete a puzzle of the house's lifetime in half a century. More importantly, the finished questionnaires identified some significant and undisclosed findings in this stage. Time was crucial for identifying significant changes in the building's façade and interior space. In general, the timeline was divided into two periods based on a key change – the demolition of the gate building inhabited by this family (see Table 2).

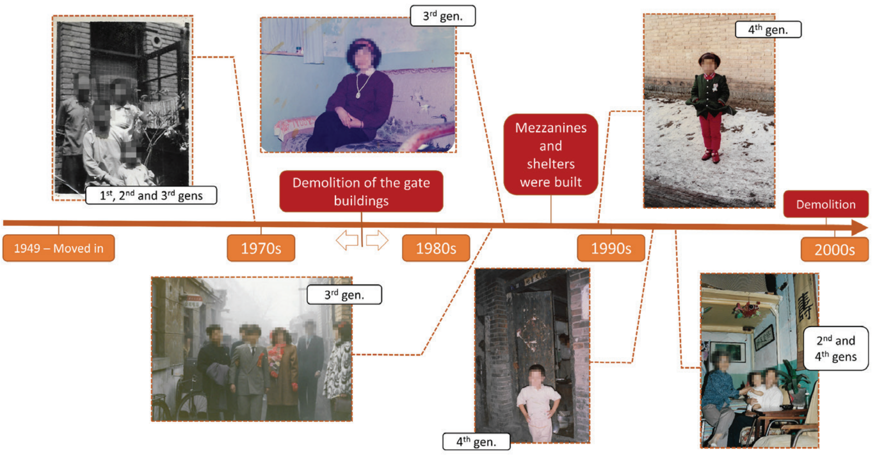


Table 2. The timeline of the house.

Consequently, the timeline is divided into two periods:

Late 1949 – the 1970s. The house's early history in this period suggested two gate buildings at both ends of a hutong (a type of narrow alley in China), flanking single-storey buildings and courtyards. The family's home was in the north-eastern corner of this area, and they also inhabited the north gate building. Based on the descriptions, they could access the gate building by a ladder from the single-storey house. Some places where only the names exist were still undemolished then. For instance, the Baigu Pagoda could be seen from the balcony of the gate building in the 1960s. Some artefacts were mentioned, but their locations, such as a sewing machine, a camphorwood chest, and a cabinet, remained unknown.

The 1980s – 1990s. In this period, two gate buildings were demolished because of the damages caused by the earthquake in 1976. To achieve more residential space, the family members built a mezzanine in the single-storey building for storage and temporary residence. Additionally, the residents there started to build temporary shelters to acquire more living space, resulting in this area's decay. Existing photographs provided consistent evidence and supported the family members' description. Additionally, some family members autonomously used sketching to express their spatial cognition, providing visual references to the interior setting. For example, a floor plan with the interior setting in an axonometric view demonstrates the scenario during the late 1980s (see Figure 2), which helped us extend and reconstruct the space suggested by existing photos.

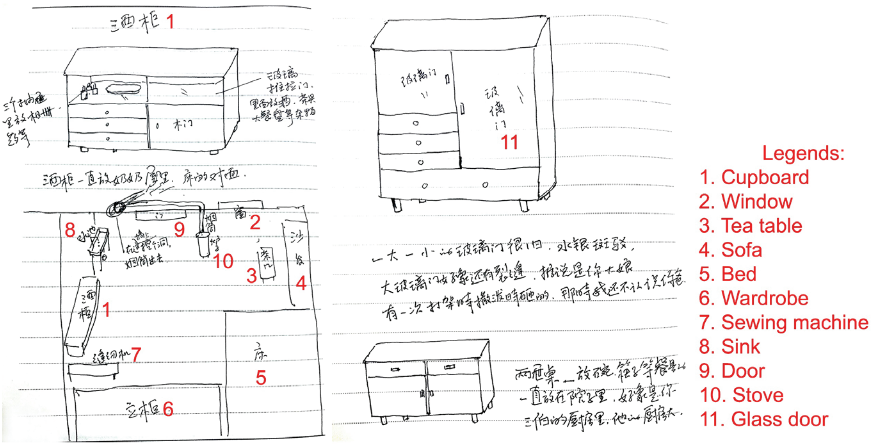


Figure 2. This image shows one family member's sketch of the interior setting in the 1980s.

As these two periods were identified, the translation into digital models could be created in two scenes in the final VE. One was set in the 1960s. Some now-lost buildings could be seen in this scene, such as the two gate buildings and the pagoda. The other scene was set in the 1980s. It showed the evolution of the house after the gate building's demolition, including the mezzanine in the single-storey building and the temporary shelters.

2.3. DIGITAL INTERPRETATION AND BUILD OF VE

The creation of digital models, including the house and context, was based on two scenes with universal parts and replaceable components, see Figure 3.

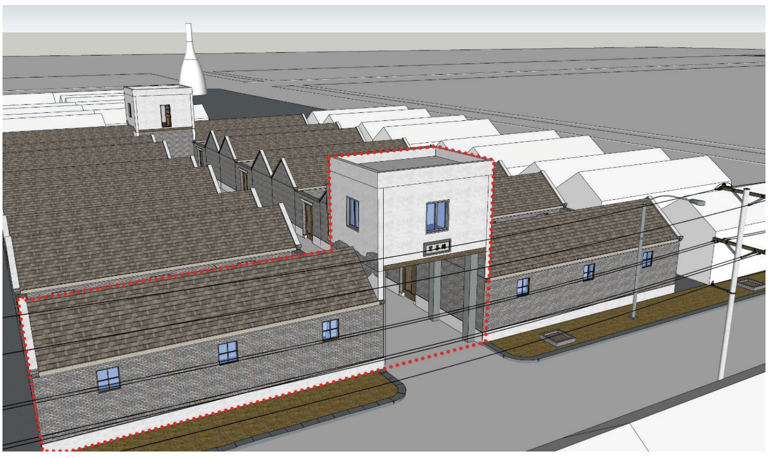


Figure 3. An aerial view of the digital model of this residential area. The pagoda can be seen in the distance. The red dashed box shows the replaceable parts.

This stage also developed a comprehensive method to create embodiments for visual and non-visual elements, see Table 3.

Architecture	Exterior space	Interior space	Context	User interface	Lightning	SFX	VFX
Walls	Courtyard	Furniture	Vegetation	Interaction systems	Directional light	Environmental sounds	Particle system
Floor ground	Infrastructure	Homeware	Terrain	Menu	HDR	Interaction sounds	Post-processing
Ceiling	Vehicles	Fabric	Waterbody	Windows Buttons Bars	Point light		Volumetric light
Window door		Artefacts			Spot light		
					Reflection probe		

Table 3. A diagram of how the collected data is classified and the embodiments in the VE.

The final product was a first-person-view with a game-like application; as a screen-based VE, the users could manipulate and navigate using a keyboard and mouse. Two scenes were separately built into the system, and users could experience them by switching between them.

It is noted that *secondary* interpretation was used in this stage to fill in the gap when insufficient information was collected. For example, the creation of the interior space before the 1970s needed interpretation because the locations of the mentioned artefacts were unknown. Moreover, multisensory feelings such as sound and atmosphere required embodiments in the VE. Therefore, the use of a games engine – Unity3D – aimed to embody visual and non-visual elements in the VE. Figure 4 compares a family member's sketch (1980s version) and the re-creation in the VE.

Figure 4. A comparison between the family member's sketch and the digital re-creation in the VE.

The key to validating the secondary interpretation was the constant negotiation between the family members and the researcher regarding the accuracy of the digital representation, the level of detail, and the credibility of the VE. Therefore, in the testing stage, the family members would experience the final product and provide feedback, which looped the workflow back to earlier stages to refine the product and produce iterations.

2.4. TESTING AND ITERATIONS

Due to the Covid-19 pandemic, face-to-face testing of the final product among family members with immersive virtual reality (VR) devices was impossible. Instead, a pre-recorded walkthrough video was used, and the session was conducted remotely. Even though the family members were actively involved in the remote session, the reconstructed house recalled nostalgic moments, resulting in a constant discussion regarding the feedback and more undisclosed information. Consequently, three feedback loops were identified, aiming to refine the final product regarding correcting the VE, adding more undisclosed objects, and improving the system to increase immersion and interaction. As a departure, the feedback loops informed the workflow to create new iterations.

For example, according to the feedback, the architectural space looked too 'clean' – the house in the 1990s was messy and shabby. Seamless texture maps did not represent the wear and tear quality. In response, UV mapping in Rhinoceros was used to improve the materials' quality and credibility (see Figure 5). Moreover, because the final product was not fully experienced due to the restriction of Covid-19, the interaction system was developed to provide an intuitive experience, including redesigning the user interface and VR availability.

Figure 5. The refined version of the gate building and the single-storey building in the 1970s scene using UV mapping in Rhinoceros to add wear and tear quality to the materials (left), and the screenshot of a user walking in the VE (right).

In addition, as the new iteration was demonstrated to the family members again, the improved visuals such as the remapped materials helped the users get to a deeper immersion, resulting in newly recalled objects that were undisclosed in the previous data collection. They also provided more detailed feedback for correcting the object's size or material. Overall, the feedback session was successful and drove the workflow well in generating new iterations. The development of the workflow's output informed us that the influence of *secondary interpretation* could be mitigated by refining the VE according to the feedback which also provides new information to be translated. In other words, *secondary interpretation* will be continually used in creating new iterations. Additionally, it is noted that more unidentified feedback routes can be found as more information will be included in future iterations.

3. Result and discussion

This paper used a workflow to re-create a now-lost house based on collective memory. The final product was a game-like application allowing the users to re-experience the visualised house. Although the collected memories were about different periods of the house, they created a "holography" consisting of a "network of references and influences" (Eco, 1987, p. 7). The unveiling of this lost history suggested unnoticed details and the social landscape (Burke, 2014), enhancing our understanding of this historical site and its context.

In response to the research questions, the use of questionnaires and participation managed to textualise and document the memory and keep them comparable. However, because of the ambiguity in the textual description, the translation stage required secondary interpretation as a mediating device to fill in the gap and create a VE with a sensible level of detail to maintain the immersion. Regarding the multiple versions of the house referring to different periods, the data analysis stage identified key changes to the house and informed the workflow to represent them in separate scenes. Moreover, the users could experience the house in different periods by switching between the scenes. Because of the intrinsic gap between the secondary interpretation and the original image in the family members' memories, the workflow was constantly critiqued and refined based on the users' feedback, aiming to mitigate the secondary interpretation's potential bias and distortion. Consequently, the feedback loops were the key to testing and refining the final product. Additionally, the limitation is noted that the workflow's product can be affected by the data source's availability, such as the original photos, and the participants were shown a pre-recorded VE video instead of using a rich VR headset or CAVE system, the immersion level was relatively low which could potentially impact the research results.

In conclusion, this paper contributes to knowledge in developing a consistent workflow incorporating analytical and representative tools to study a non-extant architectural site. Using subjective sources provides another perspective in enhancing our understanding of the now-lost building and the context. Furthermore, this workflow has the potential to be extended by importing more tools and being used in other fields, such as designing tools for communicating with architects and clients.

References

Affleck, J., & Thomas, K. (2005). Reinterpreting virtual heritage. In *10th International Conference on Computer Aided Architectural Design Research in Asia, CAADRIA 2005* (pp. 169-178). The Association for Computer-Aided Architectural Design Research in Asia (CAADRIA).

Bozzelli, G., Raia, A., Ricciardi, S., De Nino, M., Barile, N., Perrella, M., & Palombini, A. (2019). An integrated VR/AR framework for user-centric interactive experience of cultural heritage: The ArkaeVision project. *Digital Applications in Archaeology and Cultural Heritage*, 15, e00124. https://doi.org/10.1016/j.daach.2019.e00124

Bricoleurbanism. (2010). *Demolition of Tianjin's old city.* Retrieved November 23, 2022, from https://www.bricoleurbanism.org/china/demolition-of-tianjins-old-city/.

Brown, M. (2015). *Half-Life 2's Invisible Tutorial | Teaching Players.* 1:20-1:22 ed. Retrieved February 4, 2023, from https://www.youtube.com/watch?v=MMggqenxuZc&t=23s&ab_channel=GameMaker%27sToolkit

Brown, M. (2020). *How Level Design Can Tell a Story.* Retrieved November 23, 2022, from https://www.youtube.com/watch?v=RwlnCn2EB9o&list=PLObvYLWmb8WSc6LJm9g5WWm1lDaoLbqun&index=84&t=369s

Burke, P. (2014). *Eyewitnessing: The Uses of Images as Historical Evidence.* Reaktion Books.

Bushell, S. (2010). The Mapping of Meaning in Wordsworth's "Michael": (Textual Place, Textual Space and Spatialized Speech Acts). *Studies in Romanticism*, 49, 43-78. https://doi.org/10.2307/27867306

Eco, U. (1987). *Travels in hyperreality: essays* (W. Weaver, Trans.). Picador.

Eco, U. (1994). *Six Walks in the Fictional Woods.* Harvard University Press.

Forte, M., & Siliotti, A. (1997). *Virtual archaeology: Great discoveries brought to life through virtual reality*. Thames and Hudson.

Jianxing, Z. (Ed.) (2000). *Rebuilding in Tianjin: The reconstruction of dilapidated houses in the 20th century.* Tianjin Renmin Press.

Joyce, J. (2008). *Dubliners.* Oxford University Press.

Liu, Y. (2022, September 15). The development of a workflow in exploring non-linear interpretations of textual spaces. The University of Liverpool Repository. Retrieved March 1, 2023, from https://livrepository.liverpool.ac.uk/3164825/

Lorenzini, C., Carrozzino, M., Evangelista, C., Tecchia, F., Bergamasco, M., & Angeletaki, A. (2015). A Virtual Laboratory An immersive VR experience to spread ancient libraries heritage. In *2015 Digital Heritage* (pp. 639-642). https://doi.org/10.1109/DigitalHeritage.2015.7419587.

Nabokov, V. V. (1967). *Speak, memory: an autobiography revisited by Vladimir Nabokov*. Weidenfeld & Nicolson.

Novitski, B. J. (1999). *Rendering Real and Imagined Buildings: The Art of Computer Modeling*. Rockport Publishers.

Pamuk, O. (2006). *Istanbul: memories of a city.* (M. Freely, Trans.). Faber.

Pilz, K. (2003) Reconceptualising thought and Space: Labyrinths and Cities in Calvino's Fictions, *Italica*, 80(2), pp. 229-242.

Said, E. W. (2000). Invention, Memory, and Place. *Critical Inquiry*, 26(2), 175-192. Retrieved from http://www.jstor.org/stable/1344120

Shao, Y. (Cartographer). (1946). *Map of Tianjin*. Retrieved from https://www.geographicus.com/P/AntiqueMap/Tienjin-showa7-1932

Slater, M. (2009) Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments, *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 364(1535), pp. 3549-3557.

Webb, N., & Brown, A. (2011). Augmenting Critique of Lost or Unbuilt Works of Architecture Using Digitally Mediated Techniques. *In 29th eCAADe Conference: Respecting Fragile Places* (pp. 942-950). Education and research in Computer Aided Architectural Design in Europe (eCAADe).

Wu, H. (1992) What is Bianxiang? -On The Relationship Between Dunhuang Art and Dunhuang Literature, *Harvard Journal of Asiatic Studies*, 52(1), pp. 111-192.

Zhang, T. (1966). *Jinmen Zaji* (Vol. 57). Taiwan: Wenhai Press.

OF STONES AND WORDS

Computational Framework for Multifaceted Historical Narration of Wadi Salib

EYTAN MANN[1], JONATHAN DORTHEIMER[2], ALONA NITZAN-SHIFTAN[3], and AARON SPRECHER[4]
[1,3,4]*Faculty of Architecture and Town Planning, Technion, Israel Institute of Technology*
[2]*Ariel University*
[1]*eytan.mann@campus.technion.ac.il, ORCID: 0000-0003-0146-0677*
[2]*jonathan@dortheimer.com, ORCID: 0000-0002-7464-8526*
[3]*alona@technion.ac.il*
[4]*asprecher@ technion.ac.il, ORCID: 0000-0002-2621-7350*

Abstract. In this paper, we outline a computational framework to capture an intricate relationship between tangible and intangible cultural heritage - architecture and the multiple narratives pertaining to it, to unfold multiple histories as a means for a deeper, more comprehensive preservation of contextual heritage. Deploying a set of digital and computational tools, we present a cross-disciplinary method to produce environments infused with history, and at times overlapping narratives. The framework presented here aims to combine text and spatial data, using both Natural Language Processing and Semantic Segmentation, towards integrating seemingly divided epistemologies of heritage. We ask how we can use computation to enrich current cultural practices and what is at stake in deploying such tools. To explore these questions, we discuss a case study of Wadi Salib, an historical and conflicted neighbourhood in Haifa, Israel, and attempt to assess our framework's ability to render a historical tour through this multi-layered site. Finally, the paper identifies several pitfalls and key challenges for future research.

Keywords. Cultural Heritage, Archive, Architecture, Laser-scanning, Historical Narratives, Spatial Data

1. Introduction

In archaic and classical Greece, stones played an important role, both as a material object around which we can observe many ritual practices and as an 'image' – a symbol of something lost. Such ritual practices and mythical images involve a stone as a mnemonic object that bridges worlds – places, and times (Marrucci, 2008). The physical architectural remains of cultural heritage can potentially play a similar role. Multi-layered narratives challenge the preservation of cultural heritage because they

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 373-382.

pose the following question: Does any one layer of cultural heritage deserve to be preserved or rebuilt as that image of the past, or is cultural heritage valued in its 'stratified' state? Rather than looking at historical layers as simply written, erased, and rewritten, a critical approach to heritage seeks to expose "points of connection, proximity, and action between various pasts." (Witmore, 2007. p.547) Using a critical perspective on the materials to be preserved, this paper develops a conceptual and technological framework for digitally preserving built heritage, taking advantage of contemporary computational techniques to augment heritage sites as physical and virtual representations of multiple pasts.

We aim to adopt a critical heritage approach that accepts that history is not written exclusively by historians but involves a wide range of stakeholders and forces—acknowledging that there is never one history but always a multiplicity (Jarzombek, 1999). This approach stems from poststructuralist theories of the 1960s, which opened space for interpretation, challenging the division between material and narratives. Space must be understood not simply as a concrete, material object but also as an ideological, lived, and subjective one (Lefebvre, 1991). Through critical theory, space is repositioned from given to produced, bringing attention to the ways in which it constructs and transforms social life and its profoundly power-laden nature.

We intend to maintain spatial orientation and context by combining architecture and archival materials. In addition to contributing to a profoundly contextual heritage, such a model raises crucial questions regarding the representation of multiple narratives, the scale of historical evidence, and the accuracy of historical accounts. It is especially prevalent in places where conflict erases histories and reconstructing their traces is difficult (Nitzan-Shiftan, 2021). Since finding spatial intersections between archival materials and the heritage site is a complicated task, we claim that a partial automation of the analysis of large amounts of archival material, linking stories to spatial data, can offer a complimentary tool the production of new historical perspectives.

Material evidence, such as a building or an architectural component, can be linked with historical manuscripts that contain historical evidence in a dual process for bridging built heritage and historical archives. The result of the envisioned framework are 3D models that are informative, accompanied by metadata that describes: sources, interpretative processes, and geographic information. Ultimately, such a model could test multiple historical hypotheses and help capture the multiplicity of historical narratives that shape cultural heritage sites. Automation will allow large amounts of data to be processed, and unforeseeable associations between archival and architectural objects to emerge. Using this approach will potentially encourage collective engagement with heritage and can later be applied to extended-reality applications, in both VR and AR for historical tours through the site (Mann et al., 2022; Mann & Sprecher, 2022)

2. Computational approaches to analyzing language and space

As part of the framework, heritage materials will be processed simultaneously: manuscripts that present evidence in the form of natural language, and architecture - physical remains that are digitized and processed.

2.1. COMPUTATION OF NARRATIVES

Natural Language Processing (NLP) methods offer researchers from the humanities a way to "remotely read" archival manuscripts while picking up semantic meaning, locations within the site, and keywords. Specifically, using language models, we can identify textual elements that researchers might overlook when processing large amounts of text, including the language used to describe spatial attributes. For this, we followed research in Literary Geography, where scholars have mapped-out literary worlds described in books. We found that in such attempts, a "space of the text'" is often anchored in some form to the 'reality' of existing spaces and places (Zaman, 2020). Other works show how NLP techniques can be used for named entity extraction (Goyal et al., 2018; Nadeau & Sekine, 2007) and the advantages and challenges of extracting spatial terms from unstructured bodies of text using machine learning algorithms (Landau & Jackendoff, 1993; Zlatev, 2012).

2.2. COMPUTATION OF MATERIAL

The analysis of point clouds, captured by 3D scanning techniques, provides a different means of accessing physical evidence. The accumulation and analysis of point cloud data are often used today by research in the cultural heritage field. While wars, natural disasters, and climate change are destroying some of the world's most precious cultural sites, some are already recorded in three-dimensional spatial information via Terrestrial Laser Scanning (TLS). Remote sensing technologies for cultural heritage include three-dimensional documentation that allows the collection and reconstruction of digital models featuring a high degree of geometric accuracy (Remondino, 2011). The generation of 3D point clouds is an efficient way to manage tangible heritage assets in the Digital Cultural Heritage domain, to perform tasks such as morphological analysis, map degradation, or data enrichment. An informative model—could have the capability of becoming a scientific document.

Recently, machine learning came into play to enrich point clouds through semantic classification, allowing the addition of metadata for informing 3D models (De Luca, 2020; Pierdicca et al., 2020). Using neural networks in conjunction with Historical Building Information Modelling (HBIM) is emerging for managing multiple and various architectural heritage data transforming 3D models from a geometrical representation to an enriched and informative data collector (Bolognesi & Garagnani, 2018; Chiabrando et al., 2016). These methods are designed to recognize point clouds, classify, segment, and output categories.

3. Methodology

We explored techniques of augmenting data using historical archival research and Research by Design approach to develop the computer framework that was explored using a case study. By adopting a critical approach to cultural heritage, the study produces a testing ground for assessing the opportunities and challenges in automating the augmentation of a historical site.

3.1. CASE STUDY

We studied Wadi Salib, a neighbourhood in downtown Haifa, Israel, which has a painful history of violence and destruction. The site study was part of a collaborative research (founded by Israel Ministry of Science and Technology) during 2021-2022. The research team studied archives, including documents that provide evidence, such as testimonials, government reports, and articles.

In the first half of the twentieth century, Wadi Salib was damaged by wars and turmoil caused by grand ideologies. First, during the Arab-Israeli war of 1948, the Palestinian inhabitants fled the city during the 'Nakba,' the displacement of the Palestinian population during the 1948 Arab-Israeli War. Then, between 1960 and 1980, the government systemically cleared the neighbourhood of its Jewish immigrants, primarily from North Africa, who moved into abandoned Palestinian homes (Weiss, 1951; Yazbak & Weiss, 2011). Finally, as part of a strategic plan to "modernize the city," Wadi Salib's last residents were evacuated, and most of their homes were levelled, with some still standing empty (Weiss, 1951) (see figure 1). Wadi Salib is an urban heritage site with narratives embedded in the crumbling stones of architectural remains. While the site initially seems "silent," it buzzes with stories.

Figure 1. The mayor of Haifa oversees the destruction of the houses, 1964 (Photo Studio Dan)

3.2. COMBINING MANUSCRIPTS AND POINT CLOUDS

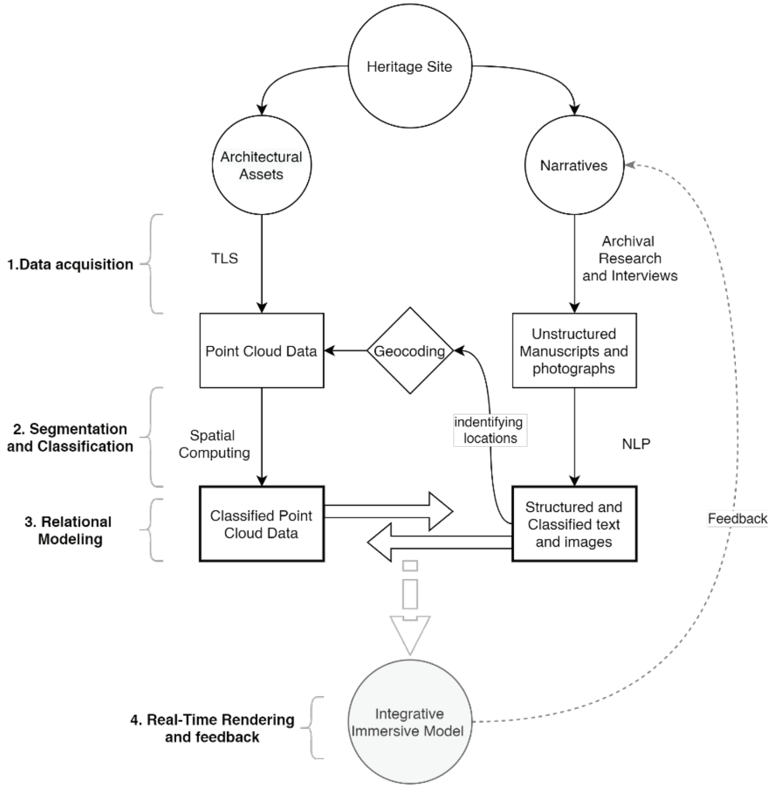


Figure 2. Simultaneous pipeline with four phases: 1) data acquisition; 2) Segmentation and Classification of data; 3) Relational Modelling; 4) Real-time rendering.

Developing the automation framework requires a multidisciplinary approach, combining historical research with computational methods to integrate architecture and archives. First, archival research was conducted to obtain various historical documents. Concurrently, 3D scans were produced with TLS to create point clouds of certain locations and buildings in Wadi Salib. Together, these data provide the basis for an integrative model of the site (see figure 2).

3.2.1. Data Acquisition (see figure 2. Phase 1.)

The collection phase followed a historiographical methodology and included the following archives: Haifa Municipality Archive, the Israel State Archive, private collections from Palestinian and Jewish refugees, and interviews with ex-inhabitants. The research team found primary sources, including testimonies, memoirs, and governmental reports, which tell two seemingly separate histories: one of a thriving Palestinian community, and another, after the 1948 declaration of the Israeli state, of a dilapidated Jewish-Moroccan immigrant neighbourhood. These two meta-histories are similar-yet-conflicting, as both consist of a tragic displacement and an explicit claim to the site.

Identifying the key locations mentioned in archival materials guided the 3D scanning of the site and the resolutions of scanning, using the TLS to produce a dense point cloud and record the global positioning of objects. Testimonies repeatedly mentioned a particular neighbourhood area during both historical periods. This area was the Casbah of Wadi Salib, the center of public life. Inside the Casbah, one residential building that was re-used as a synagogue since the year 1950. During the 1980s, the state of Israel demolished the area, all but one building, which had served as a synagogue since 1950. These archival findings guided our focus for scanning and attention to certain features, such as the architectural components mentioned and rooms with historical significance. The laser scanning resolution reflects an archival resolution.

3.2.2. Segmentation and Classification (see figure 2. Phase 2.)

For segmentation and classification of manuscripts recorded as text, we used a Large Language Model (OpenAI GPT-3) to identify locations and other spatial dimensions in the body of the unstructured text (Brown et al., 2020). We designed and optimized the model to predict the location and time of the described events in each sentence. To test the accuracy of the predictions, we compared the predicted locations and times to a manually labelled data set. Through such an iterative process, we fine-tuned the text prompt to improve the accuracy of the predictions. The manuscript segmentation process outputs an array of textual segments mentioning locations, spatial relations, and dates. The extracted locations were then globally positioned using a geocoder in the next step. In our case study, street addresses were resolved by the current open street map's geocoding service. However, some locations required human intervention to resolve the location and others were not identified.

The segmentation and classification of point cloud data were explored in both the GIS system approach and the deployment of a pre-trained neural network to classify point clouds using PointNet (Charles et al. 2017). We used a proprietary pre-trained model (pointly.ai). The method established the possibility of training a machine learning model to identify and classify several architectural components that hold historical importance: first, to identify (1) individual buildings within the recorded area, (2) long staircases going up the hill, (3) roofs, (4) roads. As a result, mentioning the location of events within testimonies can potentially be linked to physical objects.

3.2.3. Relational Modelling (see figure 2. Phase 3.)

As a result of the data augmentation, a relational model was produced, establishing a semantic linkage between the point cloud data and other data types through shared metadata. To assess the workflow this phase was done manually and included the usage of historic maps and photographs. After processing point clouds into textured realistic-looking mesh models, the team used BIM software to store and manage the relational model (see figure 3).

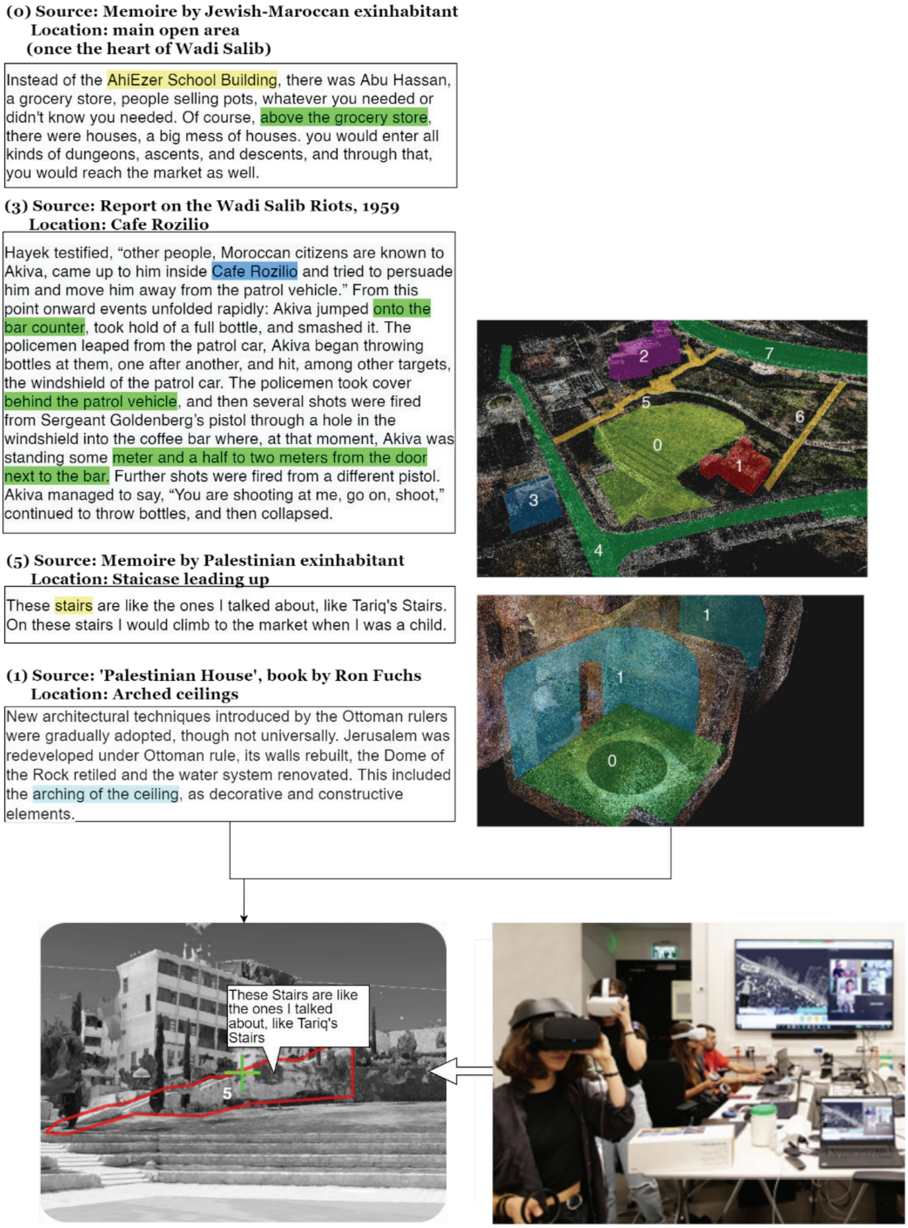


Figure 3: On the left column: segments of texts from various manuscripts mentioning place names, relational terms, or architectural components. On the right column: point cloud data segmented and augmented with metadata, including labels such as buildings' names, categories, and usage. Text segments and point clouds are automatically associated according to metadata. Finally, at the bottom center, the integrated model is rendered in real-time using the Unity gaming engine and can be experienced via VR and potentially further enriched by users' comments

3.2.4. Rendering (see figure 2. Phase 4.)

Rendering entailed importing the augmented model from a BIM software into a gaming engine (Unity 3D with Cesium Tilting Service). We used a geo-tiling system to geolocate and stream high-resolution models into the gaming engine. The augmented model was then rendered through a VR headset, allowing users to interact with model parts linked to historical materials.

4. Challenges and Pitfalls

The outlined framework above serves as a foundation for an automated tool to advance preservation of architecture and narratives. We see this experiment as a critical tool to identify the challenges and pitfalls of such automation and suggest some solutions. In our view, technological and historical challenges are inseparable epistemological concerns for this study. The digital architectural heritage community in the future should address those challenges.

4.1. NON-SPATIAL MANUSCRIPTS

Archival materials often do not refer to spatial features or locations. Such samples may contain language written in an overview format that relates to the site in general without mentioning specific locations or in a way that does not relate to the physicality of the site at all. While the information recorded in these documents may be of value, they still need a clear spatial context. For example, in an interview, the source offers a general remark on the whole neighbourhood, needing to be more specific to allow spatial extraction. This can be avoided to some extent by collapsing general remarks towards others down the text that might offer specific locations. A possible solution to this challenge is using an omniscient narrator on the narrative presentation layer that will provide information that may be relevant but does not pin down to a specific location.

4.2. UNCERTAINTIES

We notice that archival materials hold a range of historical uncertainty and varying integrity of the sources. Some are official government reports, and some personal accounts. These accounts differ in their objectivity and subjectivity, the amount of additional emotional descriptions, and rhetoric. To avoid a reductive presentation of the sources, we ought to integrate a labelling system into the application's user interface that may indicate validity levels. Such visual cues may assist the user in touring places augmented by textual and visual materials with some criticality - while maintaining awareness of the sources. To maintain historical validity in the framework context, we suggest using immersive storytelling while maintaining a connection to the original materials. This can be achieved by keeping a citation trail, indicating the types of sources. Balancing fiction and fact may be achieved by foregrounding the source while taking advantage of immersive effects. Here design of the user interface holds a critical role.

4.3. RELATIVE LOCATIONS

Related to the issue of uncertainty are situations where there are relative locations or several location aliases: Locations are sometimes expressed relative to other places, such as: "In the middle of the street lay a man struggling with a group of people who were trying in vain to remove him." The challenge of extracting spatial features from unstructured text based on spatial terms afforded the need to extract spatial terms and cross-reference other text segments that may indicate a landmark so that they serve as a spatial anchor. Thus, to overcome scenarios with relational terms only, we identified previously mentioned locations in other segments that can inform the landmark of relations.

4.4. TRANSLATION

Finally, there is the challenge of translation. Historical materials were translated from the original language, such as Hebrew and Arabic, to English since LLMs are not available for these languages. As a result, the original linguistic form and an essential cultural significance might be lost in translation. This challenge is also present in the display layer since most people do not understand both languages; they require some translation. To reduce misinterpretation and reduction, we store original text alongside translations for users to view and read texts in their original language and meaning.

5. Conclusion and Future Research

The paper outlines the building blocks for integrating archival texts and spatial data into a digital tour through the site. Segmentation and classification of both text and point clouds offer a new bridge. Our objective is to use this framework to create a VR tour through multi-layered heritage sites, exposing narratives embedded within the built heritage. The new conceptual-technical framework holds opportunities as well as challenges for cultural heritage preservation. Some of the challenges can be resolved with the advance of computer methods and some with the right design and presentation of information in the User-Interface. Storytelling – the way the histories are presented to the users, presents a compelling challenge as a result of these approaches.

Acknowledgements

This research stems from courses provided at the Material Topology Research Lab (MTRL), Technion, funded by The Azrieli Foundation Global Studio program.

References

Brown, T., et al. (2020). Language Models are Few-Shot Learners. In H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, & H. Lin (Eds.), Advances in Neural Information Processing Systems (Vol. 33, pp. 1877–1901).

Chiabrando, F., Sammartano, G., & Spanò, A. (2016). HISTORICAL BUILDINGS MODELS AND THEIR HANDLING VIA 3D SURVEY: FROM POINTS CLOUDS TO USER-ORIENTED HBIM. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLI-B5, 633–640. https://doi.org/10.5194/isprs-archives-XLI-B5-633-2016

De Luca, L. (2020). *Towards the Semantic-aware 3D Digitisation of Architectural Heritage*. 3–4.

Goyal, A., Gupta, V., & Kumar, M. (2018). Recent Named Entity Recognition and Classification techniques: A systematic review. Computer Science Review, 29, 21–43. https://doi.org/10.1016/j.cosrev.2018.06.001

Jarzombek, M. (1999). A prolegomena to critical historiography. *Journal of Architectural Education*, *52*(4), 197–206.

Landau, B., & Jackendoff, R. (1993). "What" and "where" in spatial language and spatial cognition. Behavioral and Brain Sciences, 16(2), 217–265. https://doi.org/10.1017/S0140525X00029733

Lefebvre, H. (n.d.). *The Production of Space* (1991st ed.). Blackwell Publishing. Retrieved December 7, 2022, from https://www.worldcat.org/title/22624721

Mann, E., Dortheimer, J., & Sprecher, A. (2022). Toward a Generative Pipeline for an AR Tour of Contested Heritage Sites. *https://doi.org/10.3390/app12199859*

Mann, E., & Sprecher, A. (2022). VR as critical historiography. *Difficult Heritage and Immersive Experiences*, 80–103.

Marrucci, L. (2008). Mythology and Rituals around the Stone: Applying Louis Gernet's Polyvalence des images" to "Greek Death." *Phasis*, *11*, 86–99.

Nadeau, D., & Sekine, S. (2007). A survey of named entity recognition and classification. *Lingvisticae Investigationes*, *30*(1), 3–26. https://doi.org/10.1075/li.30.1.03nad

Nitzan-Shiftan, A. (2021). Silencing palestinian architectural history in israel: Reflections on scholarship and activism. *International Journal of Islamic Architecture*, *10*(1), 233–243. https://doi.org/10.1386/ijia_00042_1

Pierdicca, R., Paolanti, M., Matrone, F., Martini, M., Morbidoni, C., Malinverni, E. S., Frontoni, E., & Lingua, A. M. (2020). Point Cloud Semantic Segmentation Using a Deep Learning for Cultural Heritage. *Remote Sensing*, *12*(6), 1005. https://doi.org/10.3390/rs12061005

Remondino, F. (2011). Heritage recording and 3D modeling with photogrammetry and 3D scanning. *Mdpi.Com*, *3*, 1104–1138. https://doi.org/10.3390/rs3061104

Weiss, Y. (1951). *A confiscated memory: Wadi Salib and Haifa's lost heritage*. Columbia University Press. ISBN: 9780231526265

Witmore, C. (2007). Symmetrical Archaeology: Excerpts of a manifesto. *World Archaeology*, *39*(4), 546–562. https://doi.org/10.1080/00438240701679411

Yazbak, M., & Weiss, Y. (2011). Haifa before and after 1948: Narratives of a Mixed City. In *Institute for Historical Justice and Reconciliation Series History of International Relations Library*.

Zaman, C. (2020). Spatial experience in humans and machines. In *MIT*.

Zlatev, J. (2012). Spatial Semantics. In *The Oxford Handbook of Cognitive Linguistics*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199738632.001.0001

TOWARDS HYPER-REALITY

A Case Study Mixed Reality Art Installation

GARVIN GOEPEL [1]†, GEORGE GUIDA [2]† and ANA GABRIELA LOAYZA NOLASCO [3]†
[1]Chinese University of Hong Kong, [2]Harvard University
[1]garvingoepel@gmail.com, 0000-0002-3274-7489
[2]gfguida@gmail.com, 0000-0002-2477-0064
[3]aga.loayza@gmail.com, 0000-0002-3943-8363
† These authors share first authorship

Abstract. This paper reflects on the notion of hyper-reality through the creation of ‘Self-Compass’, an immersive mixed reality art installation. By merging the physical with overlayed digital 3D content, this study proposes a view of current notions of the metaverse as an extension of reality rather than a digitized replacement of it. This was demonstrated by augmenting a modular installation with an immersive digital counterpart through an augmented reality (AR) application accessible through mobile devices. ‘Self-compass’ combines a timber structure and a digital AR overlay into a radial configuration that framed eight views, revealing an historical connection beyond the immediate context, and inviting reflections on the relationship between oneself and place. The AR overlay merges meaning with data, allowing one to rethink the physical through the digital, and providing awareness of our impact on place across time. The paper discusses and evaluates applied methods of merging digital and physical objects through a mixed reality (MR) installation. It expands on current workflows through the development of an AR mobile application and examines simultaneous localization and mapping (SLAM) techniques, essential in the alignment of digital content within real-world environments. The paper concludes by illustrating the potential applications and impact of AR technologies within design practices by augmenting the physical and revealing a new hyper-reality.

Keywords. Virtual and Augmented Environments, Mixed-Reality installation, Hyper-Reality

1. Introduction

We are about to enter a new era of how we interact with digital content. This era has most recently been labelled the metaverse, described as an immersive internet, and experienced through head-mounted displays rather than 2D screens (Reaume, 2022).

Prospects of the metaverse range from fully immersive virtual realities in which users will be completely detached from physical space (Meta, 2021) to fusions between the physical and digital worlds in a hyper-reality (Baudrillard ,1983). Coined by French sociologist Jean Baudrillard in Simulacra and Simulation Hyperreality describes a condition in which what is real and what is fiction seamlessly blend (Tiffin, 2005). This permits the blurring between reality and representation, where it is unclear where a simulation begins and ends (Baudrillard, 1983). Disney parks for example are such a hyperreal world that cultivate the fake or unreal until we give more value to the imitation than to the original (Baudrillard, 1983). Umberto Eco expands on this example through his travels in this hyperreality, stating that these imitations do not just reproduce reality but try to improve on it: "technology can give us more reality than nature can" (Eco, 1986). This study reflects on the notion of hyperreality by describing this fictional dichotomy as virtual content projected onto the real-world environment through augmented reality (AR). AR projections themselves are contextless, but if these projections relate to physical spaces or objects, we refer to them as mixed-reality experiences. By means of mixed-reality, Self-Compass, creates a hyperreal condition in which the physical and digital counterparts co-exist.

Keiichi Matsuda's dystopian short film vividly reveals such "Hyper-Reality" where one's daily life depends on these digital overlays (Keiichi Matsuda, 2016) (Figure 1). This paper presents a possibility of this future, by showcasing how Augmented reality technologies overlay information in the physical space and expand them into engaging mixed realities that revise the current collective perceptions of the built environment.

Figure 1. Holograms as part of our daily lives: Hyper-Reality by Keiichi Matsuda (Matsuda 2016)

1.1. CONCEPT

'Self-compass' is a mixed reality installation that consists of a physical and digital component (Figure 2 - 4). Acting as a material scaffolding for the digital, the physical pavilion is composed of 7 identical and concentrically arranged timber modules, which are unified by a circular ring on its top. The structure spans 12 ft diameter across the top ring beam and 12 ft high from the base. (Figure 4). The overall physical component embodies a metaphorical compass that frames eight views towards the surrounding context. The superimposed digital information consists of an animated compass over the top ring and 8 sets of interactive image interfaces with descriptive text, each revealing significant buildings from the past and present along the axis of each framed view and beyond the limits of the immediate site. These images reveal what we can and cannot see, perceive, and imagine. Self-compass invests into Hyperreality by treating physical surroundings, history, and digital displays as information with no hierarchy as equally valid. The scale of the pavilion acknowledges human and context scales, as well as present and past with the same importance, without no value statements, so that the visitor can shape their own perception.

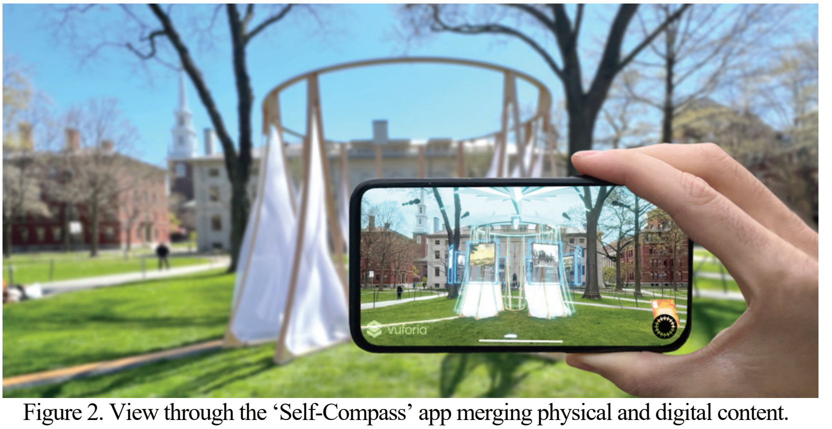


Figure 2. View through the 'Self-Compass' app merging physical and digital content.

Figure 3. Pavilion physical assembly on Site 01. Image by Joshua Widdicombe (Widdicombe 2022)

This public artwork was first exhibited at the Harvard Arts First Festival, which took place in Harvard Yard, in Cambridge, Massachusetts. The centre of the pavilion, point O (Figure 4), was marked at 40 meters from the Indicator P, which uses the statue of John Harvard as a site benchmark.

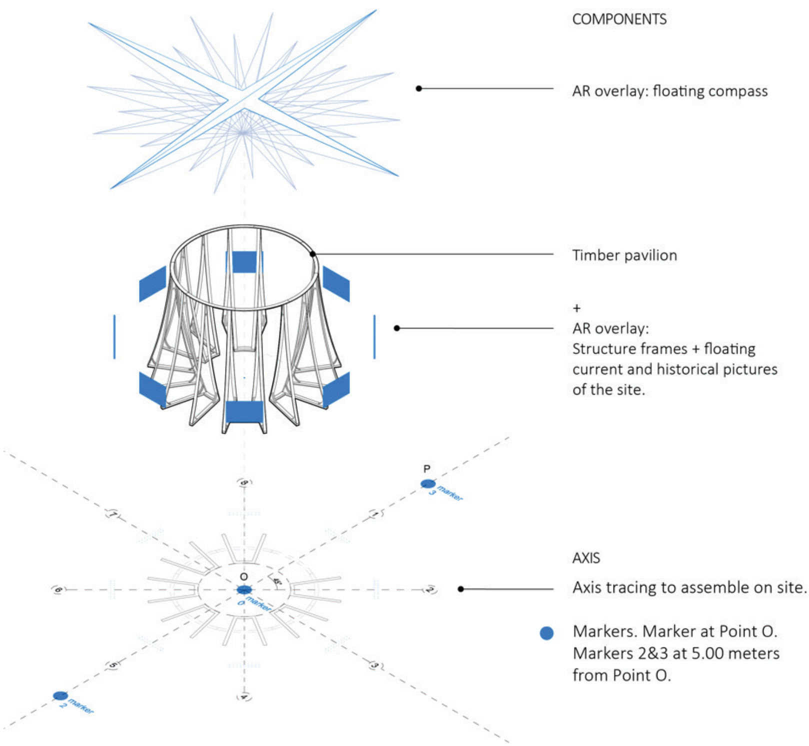


Figure 4. Pavilion Components: 'Self-compass' mixed reality content

At the centre of the pavilion the visitor is invited to rethink their connection to the site through the AR overlay on their smartphones. The exercise of situating the individual at the centre of a hidden geometry, pointing to specific alignments, means a reassessment of the symbols and meanings assigned to buildings and surrounding context, and a reflection of the role of oneself and the site.

2. State of the Art

Augmented reality is becoming progressively ubiquitous throughout multiple fields including construction, medicine, fashion, social media, education, and so on. In art, AR mobile applications allow users to display art pieces at any location (Acute Art 2022), whereas a trend can be observed where artists are enhancing physical art pieces such as 2D paintings with AR content (ArtForEveryone, 2021). Collective iheartblob expands the motion of 2D painting with `BLOB House´ (Belitskaja, James, and McCallum, 2021), by adding a digital layer of interaction onto a pavilion structure for the Vancouver Mural festival in Canada. Here, the colourful flattened 2D printed sheet walls of the pavilion reveal an animated 3-dimensionality once being triggered by a mixed reality app on handheld mobile devices.

Another mixed reality installation is "Resonance-In-Sight" (Crolla and Goepel, 2021) located in front of the Hong Kong Museum of Art (HKMoA). It consists of two curved and brightly coloured physical steel structures placed several meters apart and a digital element, which is echoing animated artworks between the two pieces in a tangible tension, visualized through an AR mobile application.

'Self-Compass' situates itself within a broader context of recent mixed-reality installations and artworks by expanding beyond physical representations and incorporating digital animation. It contributes by adding another layer of complexity through a 3-dimensional physical sculpture on which digital content is overlayed and experienced through multiple angles and within the installation itself. Visitors are fully immersed in a hyper-reality by passing through the installation and observing the AR overlay in 360° around them, rather than just from fixed point of views around the installation. The installation blurs the boundaries between the physical installations and the digital overlay by merging the edges with digital highlighted edges in MR. This effect precisely locates and seamlessly blends digital information towards the physical as built by creating a dialog between the two realities, a mixed reality, rather than just adding digital content. This study can be replicated in various locations, reacting to local context by implying a new site's object, and building mapping, as well as updating the AR display content.

3. Method

The overall design strategy was driven by physical fabrication parameters in order to create an immersive digital experience. Through a iterative design process, initial designs were 3D modelled to assess the structural stability and lightness of modules through timber mock-ups. This was then used to test with different content and displays that aligned with the main concept of the installation. The digital model would constantly get updated with the physical modifications of the timber pavilion until reaching the final built version (Figure 5). The digital model and AR application were adjusted to calibrate the size and projection of the AR component (Figure 7).

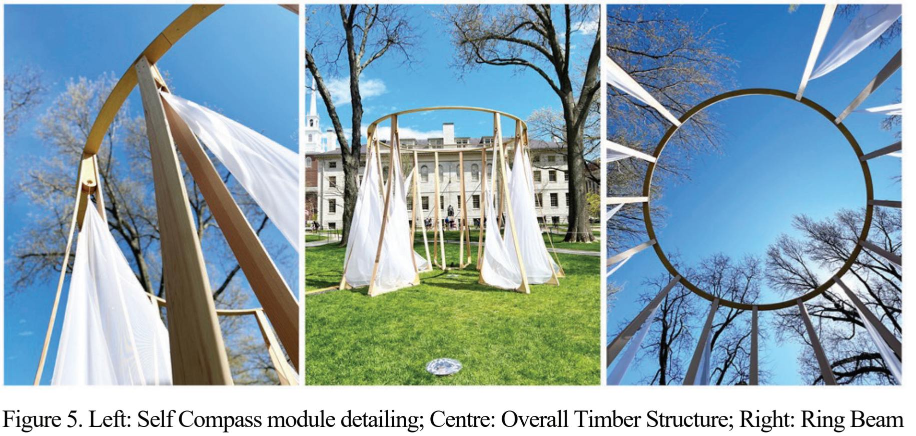


Figure 5. Left: Self Compass module detailing; Centre: Overall Timber Structure; Right: Ring Beam

3.1. AR APP

The physical installation was augmented by a mobile iOS app available for download on the Apple App store (Apple 2022a) called ‘Self-compass’ (Goepel, 2022) (Figure 6). The app’s main function is to permit users to interact with digital content overlayed on top of the physical installation. It, therefore, utilizes the built-in cameras of mobile devices to create a mixed reality experience. The digital content consists of a 20 ft compass with an animated rotating compass needle and radially animated spheres, which hover over the centre circle of the physical installation, a digital boundary edge of the installation, and 32 panels of which 8 are displayed at the same time within each physical module (Figure 6). The geometry was modelled in Rhino3D (Robert McNeel & Associates, 2021) and then exported as an FBX file to be imported into Unity3D (Unity Technologies, 2020) where the gameplay and user interface were created (Figure 7).

Figure 6. ‘Self-compass’ mixed reality content

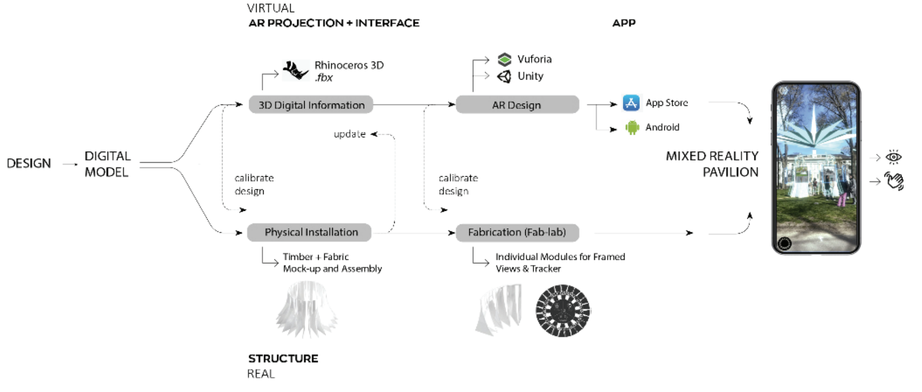


Figure 7. ‘Self-compass’ mixed reality architecture

The gameplay allows users to interact with the panels through a double-tapping motion on the mobile screen (Figure 8). This gesture sends a digital raycast on the digital panel surface which is set to be a mesh collider. If hit, the panel is deactivated and another one is activated. This cycle repeats up to eight times per panel until the starting image appears again.

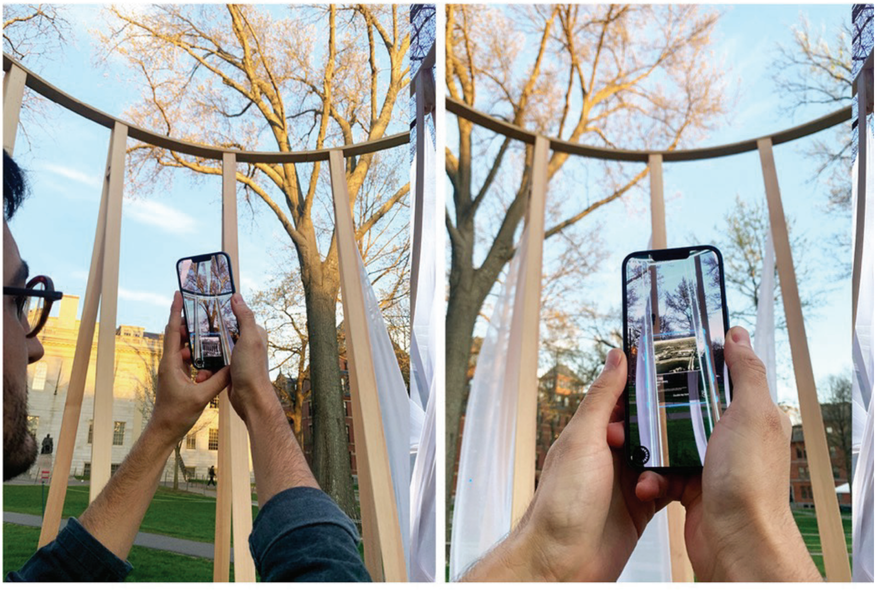


Figure 8. Image panel visible between each physical module

The sequence of flipping between panels displays old and new buildings in the respected compass direction, switching between past and present (Figure 9). The frames additionally change colour with each tap between present (white) and past (blue) to give users visual cues (Figure 9).

The minimal user interface consists of instruction panels on how to interact with and activate the augmented reality part of the app for the mixed reality experience (Figure 9) Once the instructions are read, a tap on the panel deactivates these and activates an image representation of the image trackers asking the user to scan these to activate the digital content. This scanning instruction disappears once a marker is detected by the camera, initiating the digital scene for the respected location of the image tracker being in front, behind, and in the centre of the physical installation.

The digital part was aligned with the physical installation through image tracking of three custom markers (Figure 9) by using the Vuforia engine (Vuforia, 2022) an AR software development kit (SDK) package for Unity3D (Unity Technologies, 2020). If the tracking is lost, a help icon on the top right corner brings back the instructions and resets the scene. The user can freely move between the marker positions and rescan the closet marker which will realign the digital content with respect to the user's location. The outer markers are meant to focus on the animated compass whereas the inner marker concentrates on the panels.

An icon in the bottom right corner of the app allows users to toggle between activating a digital twin of the installation and a depth mask, which is a digital material assigned to the 3D model to mask out digital geometry behind the physical installation to prevent digital content to appear in front of the physical objects. This accumulation of these elements makes the MR experience more accessible and therefore more immersive.

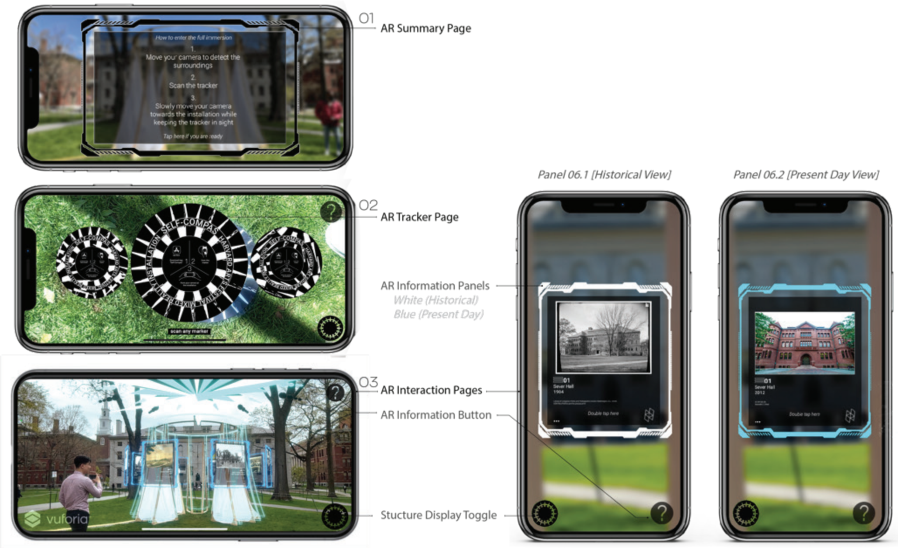


Figure 9. In-app and tracker instructions (left) and panel past + present (right).

4. Results and Discussion

4.1. DIFFICULTIES OF AR APP

Due to the unevenness of the site, the alignment of the AR app proved to be difficult to calibrate in each of the 3 locations. This resulted in an initial alignment on the first day, and a relocation of the Z plane height in Unity3D (Unity Technologies, 2020) of each one to provide a seamless overlay between physical to digital. The launch of the exhibition was coherent with the onsite assembly of the physical installation and the launch of the app in the App Store. This challenged the accurate alignment as any implementation of onsite detected indeterminacies needed to be compensated in Unity3D (Unity Technologies, 2020) and uploaded to the Apple App store with 24h-48h approval time. Due to these circumstances, the first day only had a mediocre alignment available in version 1.1 of the app, whereas the new app versions 1.2 and later 1.3 optimized the as-built installation with the digital model.

4.2. APP ANALYTICS

The 'Self-compass' app (Goepel, 2022) has been downloaded 188 times on the Apple App store (Apple, 2022a) during the Harvard Arts Festival period between the 28th of

April and 2nd of May 2022 (Figure 10). During this time the app's product page was viewed 103 times on the App Store and the app left 1.62K impressions by the app’s icon being viewed on the App Store, all on devices running iOS 8, tvOS 9, macOS 10.14.1, or later. This data has been collected and accessed through the Apple app store connect page (Apple 2022b) for the ‘Self-compass’ app (Goepel, 2022). In later periods between May 2nd to May 16th, and June 8th to present, corresponding to the re-location of the artwork in Harvard Graduate School of Design’s outdoor yard and the Studio 80 + Sculpture Grounds in Connecticut, the number of downloads experience a decay over longer periods.

Users demonstrated excitement and curiosity when discovering the virtual dimension of this built form, and while at times challenged by the app installation process, successfully engaged and walked around the pavilion with mobile devices and tablets

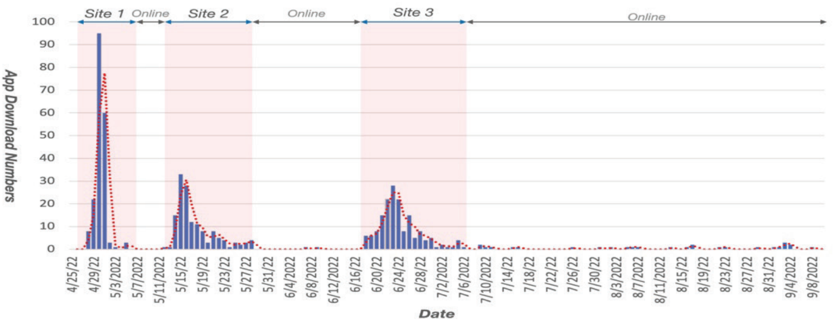


Figure 10. App downloads during the festival period

4.3. FUTURE DEVELOPMENT

Following the three separate occasions, the pavilion can continue to be relocated and include new site displays. Challenges of physical reassembly and adaptation of the virtual can be addressed through faster data collection, and automated mapping and alignment methods

5. Conclusion

The case study successfully demonstrated a fusion between a physical installation and an engaging and informative virtual counterpart. ‘Self-compass′ engaged visitors to experience glimpses towards the past and present and encourages discussion about the future of the built environment, where AR enables dissociations from the physical spaces and fosters hyper-real experiences. The intensity of use of the AR component, however, was related to the ongoing events of the site. Where there was an event, transit, or gathering, the number of users who experienced the app was greater than on other sites with fewer congregation of people. Such a case reveals the strict attachment of AR to popular or active sites to become successful. Through its potential to convey meaning, memory, and immersive digital media, AR illustrates the impact on expanding architectural design solution spaces through these new digital layers that behave as an augmentation of the physical, into a hyper-reality.

Acknowledgements

We would like to thank the entire Self-Compass project team including Thomas Kuei, Jeffrey Stevens, and Taeyong Kim. We extend our appreciation to the Harvard Arts First Festival advisory committee, especially Kathy King and Brian Dowling, for their support and funding. Our thanks also go to the Harvard Graduate School of Design, specifically Rachel Vroman and Burton LeGeyt, for their support in providing the necessary fabrication facilities.

Figure

Figure 1: Matsuda, K. (Director). (2016, May 19). HYPER-REALITY. https://www.youtube.com/watch?v=YJg02ivYzSs
All other figures by the authors.

References

Acute Art. (2022). Acute Art App—Virtual Reality & Augmented Reality Art Production. Acute Art. Retrieved from: https://acuteart.com/

Apple. (2022). App Store Connect. Retrieved from: https://appstoreconnect.apple.com/login

ArtForEveryone. (2021). ArtForEveryone. ArtForEveryone App - Hong Kong Museum of Art. Retrieved from: https://www.artforeveryone.hk/

Baudrillard, J. (1983). Simulations (P. Beitchman, P. Foss, & P. Patton, Trans.). Semiotext. Columbia University, New York City.

Belitskaja, S., James, B., & McCallum, S. (2021). Blob House – iheartblob. BLOB HOUSE by Iheartblob. Retrieved from: https://www.iheartblob.com/blob-house/

Crolla, K., & Goepel, G. (2022). Entering hyper-reality: “Resonance-In-Sight,” a mixed-reality art installation. Frontiers in Virtual Reality, 3. https://doi.org/10.3389/frvir.2022.1044021

Eco, U. (1986). Travels in Hyperreality (W. Weaver, Trans.; 1st edition). Harcourt.

Goepel, G. (2022). Self-Compass (V.1.3) [IOS]. Retrieved from: https://apps.apple.com/de/app/self-compas/id1621241119

Matsuda, K. (2016). HYPER-REALITY. Retrieved from: https://www.youtube.com/watch?v=YJg02ivYzSs

Meta (2021). The Metaverse and How We’ll Build It Together—Connect 2021. Retrieved from:https://www.youtube.com/watch?v=Uvufun6xer8

Reaume, A. (2022). What Is The Metaverse? Meaning and What You Should Know | Seeking Alpha. Retrieved from: https://seekingalpha.com/article/4472812-what-is-metaverse, Retrieved from: https://seekingalpha.com/article/4472812-what-is-metaverse

Robert McNeel & Associates. (2021). Rhinoceros (V.6.0) [PC]. Robert McNeel & Associates.

Stephenson, N. (1992). Snow Crash. Bantam Books.

Tiffin, J. (2005). HyperReality: Paradigm for the Third Millenium. Routledge.

Unity Technologies. (2020). Unity3D (V.2020.3.26f1) [Mac]. Unity.

Vuforia. (2022). Vuforia Engine AR (V.10.5) [Mac]. Retrieved from: https://developer.vuforia.com/

PLANTING DESIGN SCENARIO ANALYSIS WITH 3D VEGETATION MODELS GENERATED FROM L-SYSTEM ALGORITHM

WEI ZHANG[1] and JIAXUE HAO[2]
[1,2]Huazhong Agricultural University.
[1]zhang28163@mail.hzau.edu.cn, 0000-0003-4257-2659
[2]13932971357@163.com, 0000-0002-2708-3980

Abstract. Due to the complex branching structure and enormous details of the plants, it has always been a challenge to simulate the vegetation's 3D appearance. The currently widely used triangular surface models have multiple limitations in the representation of the tree structure for a large number of trees such as the low performance and low fidelity. As a generative algorithm, the L-system algorithm has been used to rapidly construct different vegetation models based on the branching characteristics and self-similarity of the foliage. In this study, a 160m×160m vegetation area in Qingshan Waterfront Park, Wuhan, China, was selected as the study area to construct a vegetation model based on the L-system algorithm to generate virtual 720° panoramic images. The virtual panoramic images were then compared with the on-site captured panoramic images by semantic segmentation method to verify the accuracy of the constructed parametric vegetation model. We calculated 3D vegetation volume during 3 different plant growth stages in the study area by converting the algorithmic vegetation model into a geometric voxel model. The results showed that the generated virtual panoramic visible green index was similar to the actual panoramic visible green index at the same location with an average difference of about 16% and the mean intersection over union (mIoU) of 50.18%. The 3D vegetation volume in this study area during the initial stage, the growing stage, and the mature stage was 17396m^3, 35679m^3, and 161007m^3, and the 3D vegetation volume per unit area was 0.68 m^3, 1.39 m^3, and 6.26 m^3, respectively.

Keywords. L-system, 3d Vegetation Volume, Panoramic Visible Green Index, Semantic Segmentation

1. Introduction

3D plant models are widely used in virtual scenes such as video games and computer graphics (CG) (Wan et al., 2019). The generative algorithms of complex structures has been widely adopted for the plant modelling, and the commonly used generative methods mainly include particle systems, fractal recursive algorithms, iterative function systems, and the L-system algorithm (Ben, 2019; Pu et al., 2016).

The use of generative algorithms for simulating natural plant morphology was

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 393-402.

originated in the 1960s and 1970s. Lindenmayer proposed the self-similarity of the plant growth based on fractal theory, and constructed a plant model using the L-system algorithm to simulate the growth of plants (Prusinkiewicz et al.,1990). The L-system algorithm follows the growth patterns of plants, and its simplicity and operability make it highly suitable for modelling and analyzing 3D plants in plant science.

In plant planning and design practice, plant modelling and visualization are still dominated by traditional modelling methods, but parametric modelling methods are rarely involved. With the development of the city information modelling (CIM) and digital twins, the accurate representation and simulation of the 3D trees have become an important research topic. In this study, we proposed a quantitative method to measure and validate plant growth of Wuhan Qingshan Waterfront Park at different stages based on the L-system algorithm.

2. **Research methods**

The study area is located in Qingshan Waterfront Park, Wuhan, China, which is adjacent to the Yangtze River and is an important green space for the surrounding residents. The park construction was completed in June 2017. A plot of 160m×160m in the Central Square and its surrounding areas within Qingshan Waterfront Park was selected for the study area (Figure 1).

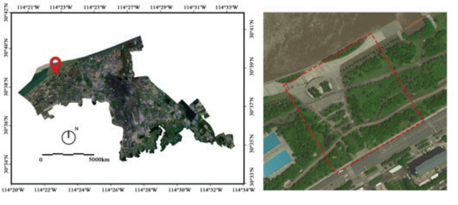


Figure 1. Study area (left panel, Wuhan Qingshan District; right panel, Central Square and its surrounding areas of the Qingshan Waterfront Park)

2.1. DATA ACQUISITION AND PROCESSING

2.1.1. 3D reconstruction based on UAV photogrammetry

The data of tree species, current canopy height and canopy width were collected by unmanned aerial vehicle (UAV) photogrammetry. The UAV automatic route planning tool in DJI Go4 software was used to obtain the 3D modelling data. The optimal UAV flight path was calculated, and high-quality modelling images were captured to achieve a high-precision 3D reconstruction model with detailed tree information (Chang et al., 2022).

The processing of the UAV acquisition data was carried out with Pix4dmapper. After setting up the image coordinate system, registering the UAV model, setting the camera parameters, we performed the automatic two-dimensional map and 3D model construction (Figure 2). Finally, we obtained 3D point cloud data models (Figure 3), 3D grid models, orthophoto, raster digital surface models (DSM), and other data during plant growth stage in the study area.

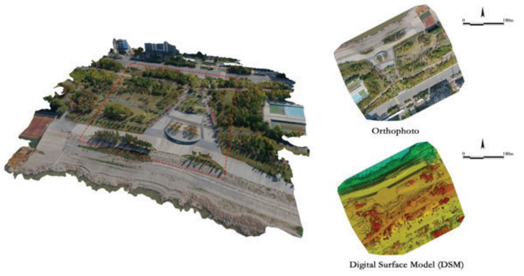


Figure 2. 3D reconstruction model of the study area

Figure 3. Point cloud data

2.1.2. 720° panoramic image capture

We used the Garmin VIRB 360° panoramic camera with a tripod to capture the panoramic images on 50 randomly generated locations in the study area. The Garmin VIRB 360° camera was equipped with a GPS + GLONASS satellite positioning system. The height of the camera was set as 1.6m, the average eye height of the adults. Each acquired image was in an equidistant cylindrical projection with a resolution of 5,640 × 2,820 pixels. The locations of the images were shown in Figure 4.

2.1.3. Plant geometry information extraction

The site topography and geometry data of the trees were extracted from the 3D reconstructed vegetation model. Based on the point cloud data, the orthophoto, planting construction graphs, and DSM site elevation data, we adopted the visual interpretation method to collect the physical attributes of each individual tree (Figure 5).

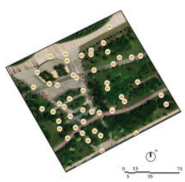


Figure 4. Location of the 50 panoramic images

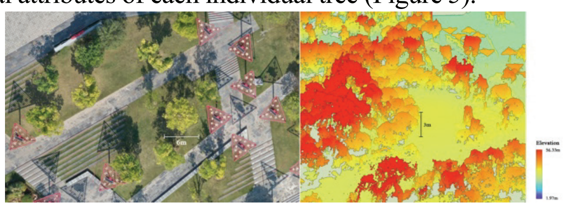


Figure 5. Extraction of information on trees in plant growth stage

2.2. L-SYSTEM MODEL GENERATION

The plant models in 3 different growth stages (initial, growing, and mature) in the study area were generated. The attributes of each tree in different stages were obtained from the planting construction drawing of Qingshan Waterfront Park, the 3D reconstruction model, and Database of Landscape Plants in Central China (Longfei, 2022) respectively.

2.2.1. Branching model

In the L-system algorithm, the plant branches were abstracted as travel paths from the starting point to simulate the branching process of a tree in the natural environment. The branching structures of the 20 plants in this study were formed by adjusting the L-system rules based on 3 basic plant branching types (false dichotomous branching, monopodial branching, and sympodial branching)(Chuangxing et al., 2012).

2.2.2. Parametric twig model

Different twig models were generated with the parametric tools in Houdini. 20 twig models were generalized corresponding to each tree species. The leaf shapes of these 20 twigs can be summarized into 6 basic categories (oval leaf, ovate leaf, inverted ovate leaf, palmate leaf, linear leaf, and round leaf). The leaf models were then constructed in parametric form as 20 twigs, which were used to construct leaf structures for the L-system algorithmic plant models (Figure 6). Each twig was then converted into an algorithmic symbol to construct the algorithmic plant model.

Figure 6. 20 Parametric twigs constructed by Houdini

2.2.3. Complete plant model

The L-system algorithm tool in Houdini includes around 40 parameters containing the iteration number of the algorithm such as the step size and scale, the angle and scale, the tube thickness and scale, the gravity factor, the random factor, and the constant value size settings. In the combination of relevant parameters and algorithm rules, a close-to-realistic tree model with a specified growth height and canopy size was then generated (Figure 7). 20 tree species on the site were parametrically generated (Figure 9). Eventually, the algorithmic tree models were placed to the locations in the study area to construct 3-stage vegetation models (the initial stage, the growing stage, and mature stage) (Figure 8).

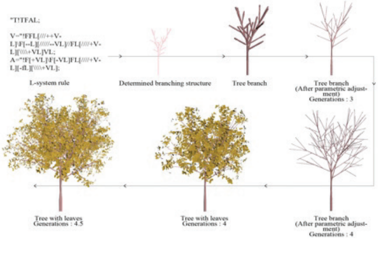


Figure 7. L-system tree model generation process

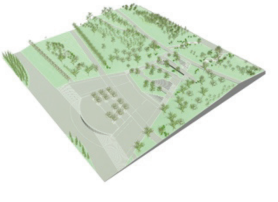


Figure 8. Algorithmic model

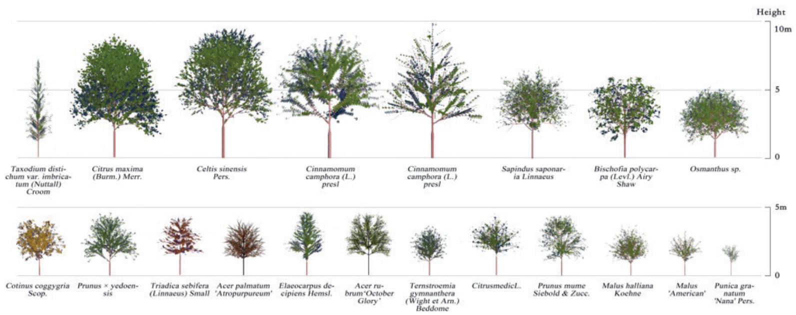


Figure 9. Summary of 20 L-system tree models

2.3. VOXEL VEGETATION VOLUME MODEL GENERATION

The voxel model is a geometric model which represents the spatial volume of a plant by voxels. We converted the L-system generated model into a point cloud model with an interval of 0.5m × 0.5m × 0.5m (Wei et al., 2022), and then constructed a vegetation voxel grid from the point cloud generated by Houdini (Figure 10). The 3D vegetation volume in the study area was calculated as the sum of the voxels (Figure 11).

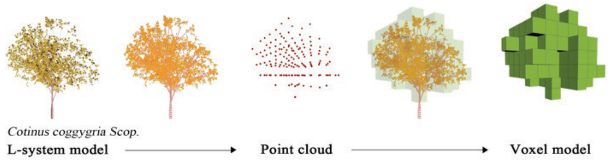


Figure 10. Plant algorithm model voxelization

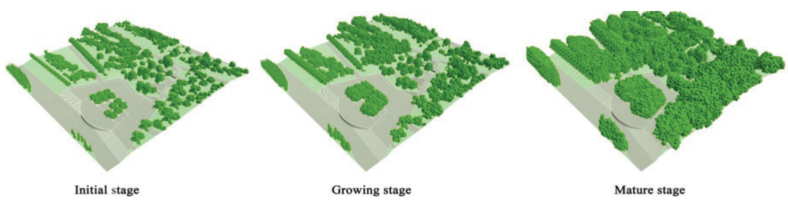


Figure 11. Voxel models in 3 plant growth stages (initial stage/ growing stage/ mature stage)

2.4. CALCULATION AND VALIDATION OF PANORAMIC VISIBLE GREEN INDEX

The panoramic visible green index (Wei et al., 2019) is defined as the percentage of the visible green area in the 720° spherical surface from a person's view, and it represents the human perception of green view in space.

2.4.1. Projection transformation

The distances between the parallel lines on the original equidistant projected images are not equal in real scenario, and thus the areas on images cannot represent the actual areas of vegetation. In order to eliminate the area distortion, the acquired panoramic photos were transformed from equidistant projection to equal-area projection. Specifically, the panorama was first converted from equidistant projection to spherical coordinates, and then from spherical coordinates to equal-area projection by the following formulae.

From equidistant projection to spherical coordinates:

$$\lambda = \frac{x_1}{(cos\,\phi_i)} + \lambda_0 \quad (1)$$

$$\phi = y_1 + \phi_1 \quad (2)$$

From spherical coordinates to equal-area projection:

$$x_2 = (\lambda - \lambda_0)\,cos\,\lambda_0 \quad (3)$$

$$y_2 = \frac{\sin\phi}{\cos\lambda_0} \quad (4)$$

Where λ and ϕ are the longitude and latitude of each pixel in spherical coordinates; ϕ_1 is the standard latitude line in spherical coordinates; λ_0 is the central longitude line in spherical coordinates; x1 and y1 are the horizontal and vertical coordinates of each pixel in the equidistant projection; x2 and y2 are the horizontal and vertical coordinates of each pixel in equal-area projection.

2.4.2. Semantic segmentation

To verify the accuracy of the vegetation model constructed by the L-system algorithm, we compared the virtual panoramic visible green index of the vegetation model generated by the L-system algorithm and the corresponding actual panoramic visible green index on the site. Figure 12 showed the on-site panoramic images and the virtual panoramic images at the same locations.

We used the semantic segmentation approach to identify vegetation elements in the images(Wei et al., 2019). We used the Dilated ResNet-105 convolutional neural network model provided by Wolfram Neural Net Repository to automatically compute the panoramic visible green index in 50 panoramic photos acquired from the site(Yu et al, 2017). The training datasets of Dilated ResNet-105 model were obtained from the Cityscapes Datasets, and training datasets contained 25000 manually labelled street images(Cordts et al., 2016). The Dilated ResNet-105 neural network model segmented the images into 16 categories (Figure 13). The panoramic visible green index in the study area was calculated as the ratio of the vegetation area to area sum of all elements (16 categories) in the equal-area projection images.

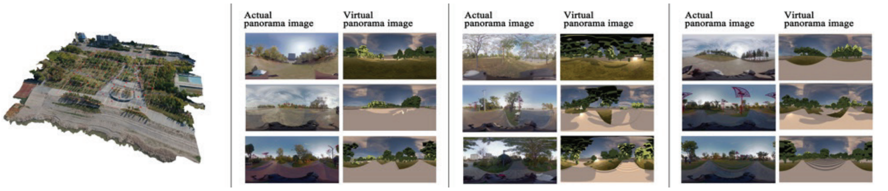


Figure 12. Actual and virtual panoramic images

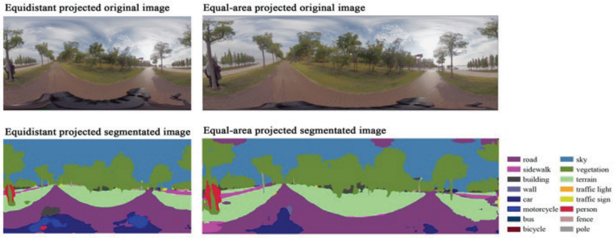


Figure 13. Semantic segmentation of equal-area projection images

3. Results

The results showed that the 3D vegetation volume in the study area was 17395.75 m3 in the initial stage after the construction was completed. After nearly five-year vegetation growth, the 3D vegetation volume was 35679.75 m3 in the growing stage, and it was estimated to be 161007.00 m3 at the mature stage.

In the growing stage, the 3D vegetation volume per unit area in the study area was1.39 m3, which was about twice of the 3D vegetation volume per unit area in the initial stage. Based on the traditional definition of green space rate, which referred to the proportion of the green area in total land area, we defined the proportion of vegetation volume in 3D space land area as 3D vegetation volume per unit area. The 3D vegetation volume per unit area in the initial stage, the growing stage, and the mature stage was calculated as 0.68 m3, 1.39 m3, and 6.26 m3, respectively (Table 1).

Indicator	Initial	Growing	Mature
Voxel unite size(m^3)	0.125	0.125	0.125
Number of voxels	139166	285438	1288056
3D vegetation volume(m^3)	17395.75	35679.75	161007.00
Study area (m^2)	25705.44	25705.44	25705.44
3D vegetation volume per unit area (m^3/m^2)	0.68	1.39	6.26

Table 1. 3D vegetation volume in the study area in three stages

Based on the semantically segmented panoramic images, we obtained a planar layout of the panoramic visible green index at the growing stage in the study area. The smallest panoramic visible green index was located within the Central Square, and the largest panoramic visible green index was located on the lawn or under the forest. The difference between the actual panoramic visible green index and the virtual panoramic visible green index of the L-system generated model was shown in Figure 14.

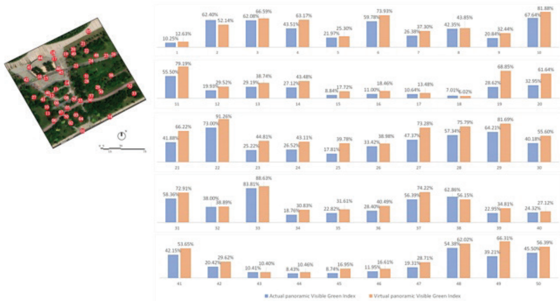


Figure 14 Planar visible green index distribution

In order to verify the accuracy of the virtual panoramic visible green index estimated by the L-system, we compared actual panoramic visible green index and the virtual one by calculating the IoU (intersection over union). IoU referred to the ratio of the green area intersection between the virtual panoramic image and the corresponding actual panoramic image to the green area union in two types of images (Garcia-Garcia et al., 2017), and IoU was calculated according to formula 5.

$$\mathrm{IoU} = \frac{A \cap B}{A \cup B} \times 100\% \qquad (5)$$

Where A and B are the target results and predicted results of green area in panoramic image. The green area of virtual panoramic image is predicted results. The green area of actual panoramic image is target result.

The result showed that the mean intersection over union (mIoU) of the visible green index in the study area was 50.18% (Figure 15), which indicated that the study area visual analysis conducted by the L-system-generated model was reliable (Rosebrock, 2016). The average difference between the two green indices (virtual and actual) was about 16%. Some equal-area segmented images were shown in table (Table 2). The difference in green index between the L-system generated trees and the actual ones might be attributed to the pruning of the trees, the effect climatic and environmental factors on plant growth, and the growth competition among the trees.

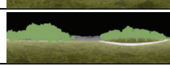
No.	Actual panoramic semantic recognition	Actual panoramic visible green index（%）	Virtual panoramic semantic recognition	Virtual panoramic visible green index（%）	Difference	IoU（%）
02		62.40%		52.14%	10.26%	54.56%
09		20.84%		32.44%	11.60%	43.81%
12		19.93%		29.52%	9.59%	47.48%
33		83.81%		88.63%	4.82%	76.93%
48		54.38%		62.02%	7.64%	77.95%

Table 2. Comparative analysis of selected panoramic visible green index

Figure 15 IoU distribution

4. Discussion

In this study, we adopted the L-system algorithm to generate the plant models and compared the panoramic visible green index between the generated models and the actual plant models with voxel modelling approach to verify the accuracy of the generated plant model. The automated generation of multiple types of trees based on the L-system algorithm makes it possible to model a large scale of vegetation and perform a visual analysis of the plants in the study area. The adjustable iteration parameters of this model allow the simulation of the different growth stages of the plants, the storage of the physical properties of plants in different stages, and the prediction of the visual changes in plant growth. Expressing plant morphology by a voxel-based model can avoid the adverse influence of plant porosity so as to achieve the consistency of visual analysis of plants in different growth seasons. The L-system algorithm-based vegetation modeling could provide a universal method for the construction of vegetation models in the urban planning and landscape design practice. It could help to solve the problems such as the inconsistency and complexity in plant modelling. The L-system generated models can also be used as a plant database for the city information modelling, spatial-ecological interactions simulation, and urban ecosystem services evaluation so as to support the building of the digital twins of our living environment.

There are still some limitations in the implemented L-system algorithms. Currently, the parameters of the L-system algorithm do not include environmental conditons such as climate conditions or plant growth competition. Future studies are suggested to integrate these factors into the generative algorithms so as to construct diversified and adaptable tree growth simulation models.

Acknowledgements

This study was supported by the National Key Research and Development Program of China (Grant No. 2019YFD1100401) and the Fundamental Research Funds for the Central Universities (Grant No. 2662021JC009).

References

Carey, B.(2019). Procedural forest generation with L-System instancing. [Masters Project Report, The University of Bournemouth University]. the Academic theme for Hugo.

Retrieved from: https://nccastaff.bournemouth.ac.uk/jmacey/OldWeb/MastersProjects/MSc19/02/MastersReport.pdf.

Changsong,W.,Yeping, P.,Guangzhong,C.(2022). A review of UAV aerial photography path planning technology for 3D reconstruction. Bulletin of Surveying and mapping, (8),98-103. https://doi.org/10.13474/j.cnki.11-2246.2022.0239.

Chuangxing, Y., Niande, Z., Wenbo,L.,Weiqiu, L., Hu,Y .(2012). Botany. Higher Education Press.

Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., ... & Schiele, B. (2016). The cityscapes dataset for semantic urban scene understanding. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 3213-3223).

Garcia-Garcia, A., Orts-Escolano, S., Oprea, S., Villena-Martinez, V., & Garcia-Rodriguez, J. (2017). A review on deep learning techniques applied to semantic segmentation. arXiv preprint. arXiv:1704.06857.

Hamon, L., Richard, E.(2010). Real-time interactive L-system - a virtual plant and fractal generator. In Proceedings of the International Conference on Computer Graphics Theory and Applications, GRAPP 2010(pp. 370-377). Computer Graphics Theory and Applications(GRAPP).

Hanan, J.(1992). Parametric L-systems and their application to the modelling and visualization of plants. (Publication No. 53757645) [Doctoral dissertation, The University of Regina].Semantic scholar. Retrieved from: https://www.semanticscholar.org/paper/Parametric-l-systems-and-their-application-to-the-Prusinkiewicz Hanan/b603b18e017da70c41262ed8fdb58d1f48afd9e0

Longfei, H. (2022). Database of Landscape Plants in Central China. Retrieved October 20, 2022. Retrieved from https://ito.clwy.cn/works/4 .

Peng, F. (2019). Growth model and visualization of a virtual jujube tree. Computers and Electronics in Agriculture,157,146-153. https://doi.org/10.1016/j.compag.2018.12.045.

Prusinkiewicz, P., Lindenmayer,A.(1990). The Algorithmic Beauty of Plants. Springer-Verlag.

Pu, L., Yuan, L., Xiaosheng, X.(2016). Three-dimensional modeling technology of virtual plant in digital landscape. In International Conference on Communications, Information Management and Network Security, CIMNS 2016 (pp.146-149).Communications, Information Management and Network Security(CIMNS).

Radomir, M., Przemyslaw, P. (1996).Visual models of plants interacting with their environment. In the 23rd annual conference on Computer graphics and interactive techniques, ACM. (pp. 397−410). The Association for Computing Machinery (ACM).

Rosebrock A. (2016). pyimagesearch . Intersection over Union (IoU) for object detection. pyimagesearch. Retrieved January 26,2023, from pyimagesearch.com/2016/11/07/intersection-over-union-iou-for-object-detection./

Shlyakhter, I., Rozenoer, et al. (2001). Reconstructing 3D tree models from instrumented photographs. IEEE Computer Graphics & Applications,21(3),53-61.

Wan Ying,W., Xiao, F.(2019). Exploration on modeling and application of 3D digital model of landscape plants. landscape architecture, 26(12),103-108. https://doi.org/10.14085/j.fjyl.2019.12.0103.06.

Wei, Z., Yuxing, Z., Mengqi,Y.(2019). Research on automatic identification and measurement of panoramic visible green index. Landscape Architecture. 27(10): 89-94.

Wei, Z., Ziqi, H., Xin, L.(2022).Voxel-based urban vegetation volume analysis with LiDAR Point Cloud. In 7th Fábos Conference on Landscape and Greenway Planning, Fábos (Vol:7. Iss:1. Art 67.) The Organizers: Dept. of Landscape Architecture and Regional Planning, University of Massachusetts, Amherst.

Yu, F., Koltun, V., & Funkhouser, T. (2017). Dilated residual networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 472-480).

AUGMENTED ENVIRONMENTS

The Architecture for the Augmented Era

ALVARO LOPEZ RODRIGUEZ[1]and IGOR PANTIC[2]
[1,2]*The Bartlett School of Architecture, UCL, London.*
[1]*alvaro.rodriguez.14@ucl.ac.uk, 0000-0002-5561-1365*
[2]*i.pantic@ucl.ac.uk, 0000-0001-8802-5592*

Abstract. Human imagination has played with the idea of an alternative technological world for years. From dystopian proposals like Neuromancer or The Matrix to more positive views like the recent Upload series, the exploration of the friction between the digital world and the physical world has entertained the imagination of our society for decades. Outside the fictional environments, the omnipresence of the internet and the development of “the cloud” are showing that the virtual world is possible and that the idea of a Metaverse is no longer part of science fiction but a very real future for human relations (Winters 2021). In line with the idea of the Metaverse, the intersection of the virtual and the physical world is being explored through the idea of Extended Realities. Technology is allowing humans to enhance their capabilities more than ever, and in fact, it has been proposed that we are entering the Augmented era (King 2014). This paper explores the opportunities and possible challenges that “Extended Architecture” has by analyzing a research project based on augmented reality as the media to explore these ideas. This project will propose a speculative approach to how the fact that in the recent future, everyone will have access to an AR device will change the way we perceive and understand our architectural environment.

Keywords. Work in progress, Virtual and Augmented Environments, Disruptive Modes of Practice and Pedagogy, Extended Realities, Machine Learning.

1. Introduction

Although accurate, Baudrillard’s idea of Hyper-Reality, in fact, can seem naive today if we analyze the evolution of society influenced by the current state of Mass Media, Social Networks, and the Internet (Baudrillard 1981). Simulation in our society is more prevalent than ever, and concepts like Instagram filters and “fake news” are making it almost impossible to separate reality and the “simulation” anymore. The shift toward a dystopian Hyper-Real future seems like a real possibility, whereas Keiichi’s Matsuda Hyperreality (Matsuda 2016) seems even a positive approach (Figure 1).

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 403-412.

Figure 1. Hyperreality by Keiichi Matsuda. 2016.

In fact, a major issue we are facing is not that what we see is altered or edited to look different or more attractive; the real risk is who or what is driving the simulation. In its origins, the internet was born for freedom and communication (Outside the military origins). The original Hackers, which was an early self-named group of programmers, envisioned the internet as a place outside the system that was not controlled by big corporations like IBM (Levy 1984). This scenario of freedom now seems all but a long-lost dream. The reality is that the newer internet platforms have conquered the Internet, and in fact, they have doubled their influence and size if we compare them with the previous generation. They are far more influential in our lives and poised to control the future of communications (Lainer 2013).

The omnipresence of smartphones and other portable devices with Augmented Reality capabilities is slowly blurring the barrier between the digital world and the analogue world. From early experiments like the Archeoguide Project (Vassillios et al. 2001) that overlapped virtual models of classic Greek buildings on top of the ruins for the visitors to experience the feeling of walking in ancient times (Figure 2), there is a will to enhance and mixing our world with the virtual one. More accessible and capable hardware has been released, and now any average user can access Augmented Reality (Coppens 2017). A good example of this is the Medusa Installation at the V&A Museum (2021), where an animated digital sculpture was showcased using Augmented Reality goggles (Figure 3).

The evolution of visualization is not the only driver for the change. Together with the visual capabilities of the AR goggles, most of the new devices like HoloLens© are incorporating a series of sensors that can read and interact with the analogue world. This idea is very relevant since that means humans can interact with the virtual world and the analogue world. Many professions and workers are now using these devices to guide and enhance their performance. Soon, these technologies will be mainstream in many areas like medicine or the manufacturing industry (Coppens 2017).

In this scenario, it seems paramount to really analyze and criticize how we use the Mixed Realities and how we will relate to them in the future. How the Augmented Human will perceive and interact with the urban environment and the architecture, and how this will change our creative view. In essence, what will it mean for architecture to enter the Augmented Era.

Figure 2. Medusa Project, V&A Museum. 2021.

2. The Augmented Environment

The idea of an augmented interaction with the environment is nothing new at this point. From its early development at the beginning of the XX century, the concept of an Augmented Reality that would enhance our view of the world has been a technological goal for the last 120 years (Sùnger, & Çankaya 2019). During the past 25 years, AR technology has accelerated its development and is becoming common in all sorts of devices (Clemens et al. 2015). In that sense, if the evolution continues at the same pace, it is expected that soon, many more objects like shop screens, cars (some like the new Mercedes Benz S Series already have it) or sunglasses will adopt this technology. In fact, it is expected that AR will be a ubiquitous thing in the near future (King 2014).

In this scenario, the speculation on how this augmented environment will be, brings two important questions to the discussion, what are we augmenting, and who is augmenting it? Several levels can be explored on what can be augmented. The first would be purely technical information, such as GPS guidance, weather conditions, safety alerts, or language translation. This augmentation could be helpful and seems innocuous, lacking any specific intention apart from real-time information. The second level is Practical Augmentation, such as general commercial advertisement, events guidance, historical facts or social interactions. These can be a bit more complex since they, by definition, have third-party influence, and their existence should require some control. The third level of augmentation would then be Perceptive Augmentation. This level touches on a deeper impact since it is the augmentation that would not only add information to the user but change the reality perceived by them in real-time. This could mean simple things like street decorations or styles to more controversial censorship or modifications. This situation makes this level the one that would need closer monitoring since it could mean a tangible risk (i.e. censorship).

Regarding the second question about who is augmenting the environment, a very intense discussion connects with current controversies about the ownership of the data, the internet, and the media (Lanier 2013). Again, to synthesize this, we could expect three possible scenarios according to the implication of the control and ownership the augmentation can have (Srnicek 2016). In the Open Source Scenario, in this case, the augmented content is produced by users for users in an open-source ownership state, avoiding control and management from big corporations or governments. The Centralized scenario. This scenario proposes the idea that a government/central neutral entity will manage the data flow and the augmentation so that the user can have access to the augmented content, ideally free and democratic, but with a certain bias based on the control entity. The Corporativist scenario. This scenario is the opposite of the first, where the tech corporations manage all the content with no or few regulations, so the users only have access to a completely brand-biased augmentation based on ownership and commercial transactions. Again, it is important to clarify that those three are ideal scenarios and are not exclusive to each other. Like the current state of the internet, we can find different levels of ownership and management models coexisting at the same time (Assange 2026).

3. Methods

Once the frame for the study is settled, this paper will introduce and analyze two speculative projects on how the challenges described before can be faced. The projects aim to propose specific scenarios where levels of augmentation and ownership are part of the discussion and the possible uses that such AR apps could have in the near or even the present. It is important to note that the projects experiment with existing AR technology and theorise about future technology. The main computing frame is based on Unity© and MRTK-Unity© tools to create the testing apps and functionalities.

3.1. THE MIXR PROJECT

The idea of a personalized augmentation for the urban environment is the starting point for the project. Trying to speculate with the proposal of a future where the Extended Realities can affect several aspects of our day-to-day life, this research explores the idea of an app that can change and configure the entire perception of the city according to the user control proposing at the same time three possible scenarios for its use, Positive, Neutral and Negative. To achieve this. The app would read the environment and detect the architectural elements around to replace them with themes chosen by the user. This ideally would give the opportunity to experience the city according to the user's taste, promoting an individualistic vision of the urban space in contraposition with the amalgamated noise we usually perceive.

Technologically speaking, the project uses an already existing technology called Image Segmentation to read and differentiate the surrounding environment and the features contained within (Zhang & Jinglu 2008). This type of computer vision allows for isolating urban features like façades or trees and then performing a change on them (Figure 4). Two actions were tested through this method, Remove and Replace. For the Remove, just an image deleting logic was used, so, for example, the user could get rid of small objects like street lamps or garbage bins. This action was limited by the technical development of the app during the research and was based on Image blur and layer segmentation.

Figure 3. Image Segmentation training.

For the Replace action, several algorithms from Pix2Pix (Isola et al. 2016) or Style Transfer (Gatys et al. 2016), were tested to give the user a wide catalogue of options to choose from. The option chosen for the prototype app was the Style Transfer. The model was trained with a collection of architectural images to create different styles, pre-classified per architectural style. As a result, the user could choose between several Machine Learning generated skins to replace the existing architecture in the environment.

Figure 4. Pix2Pix VS Style Transfer

In addition to the visualization functionality, The Add action was also added, so users could overlap new elements and information within the environment to complete the augmented visualization. A catalogue of features to add was created, and it used the same Image Segmentation technology to implement the selected components. The app also proposes the idea of a social interaction capacity based on creating communities of friends. This functionality would then add a chat and share geolocation information that the two users could visualise in real-time. An important option that the social interaction would also give is the possibility of sharing and combining views together, so a group of friends could share their own view of the city and experiment with it altogether (Figure 6).

Figure 5. App Visualization for smartphones.

The augmented environment proposal of the project positions itself into a hybrid level of augmentation (Cetralized/Corporativist). The app allows us to visualize practical information like events or communications with friends and modifies and changes the reality we perceive. The result is that Level Two and Three augmentations being the third most prevalent. Due to this fact, there was an interesting discussion within the project on how the app could be managed. The project then proposed that the app could look at three scenarios, Utopian, Neutral and Dystopian. In all three scenarios, the augmentation would be managed by a third party and for commercial use of the app. The Utopian scenario proposes a use with positive social impact, so styles and augmentation will be not only for improving the user´s perception of the city but to promote integrative values events. The Neutral scenario establishes that the central authority would manage the augmentation and would be only based on the user's taste and basic practical, unbiased information with it. The Dystopian will be a scenario where the central entity/corporation would allow the user a certain configuration but full of third-party content for commercial or political uses(figure 5).

3.2. THE XREF PROJECT

Dynamic spaces for changing uses are a concept that architecture has been speculating about for years. From the Archigram's Instant City (Cook et.al. 1961) to the Event Space proposed by Tschumi (Tschumi 1994) up to the Hyperreality (Matsuda 2014), there have been examples of this concept that shows the potential that this idea has. Opposite to the traditional immobile architecture, these projects explore how an ever-changing and mobile space can change our understanding proposing a user-based space that changes and adapts to the environment.

To create the Augmented Architecture, this project focuses on three main technical concepts to achieve the proposed goals. The first technical idea is environment recognition. To achieve this, A mix of two technologies is used to survey the environment and generate a base geometry to fixate the augmented space. For environment sensing, a ZED® camera is used, and for interaction tracking, a Leap

Motion® camera is combined. The ZED® Camera senses the environment by using a stereo camera system enhanced by AI to produce an accurate 3D model of the area recorded. This information was used to generate a digital twin of the environment to locate the Augmented Architecture (Liu, Zhao & Li 2012). The Leap Motion is a tracking device that uses two infra-red monochromatic cameras to locate and track objects on a spherical projection in a short distance. This capability allows for hand and gesture tracking to allow for human interaction. These two sensors were installed into an Oculus Quest VR device to upgrade it into an XR device. Combining the three technologies, it is possible to produce a precise and low-lag XR environment for an interactive and immersive experience.

Figure 6. Experimental XR Device.

The augmented space is generated digitally by using a series of geometry block collections and a Wave Function Collapse Algorithm (Kim et.al., 2019). Based on the mesh border generated by the 3D scanning, the WFC distributes the blocks and generates the space based on predefined premises established by the user. As a result, every space is different for every user and can be modified easily, just changing the rules or the seed for the generation process. The option of choosing different block catalogues also gives the space a more bespoke feeling to the user creating a dynamic personal augmented space that is connected and can be experimented with in conjunction with the physical space. (Figure 7)

Figure 7. WCF Model.

In order to connect both technological approaches and give the user easy access to the augmented space configuration and experimentation, a digital platform app is also proposed. This app, which could work either with smartphones or goggles, allows the user, through a simple interface, to scan the surroundings, choose from the catalogue, configure the space option and generate it automatically.

From the scenario perspective, XRef focuses on an “All Out” situation where a third party defines the app, the augmentation and the content for a mainly commercial and advertising product (Corporativist scenario). The result is a virtual space that gives access to the user to virtual shops that overlap and complement the urban environment and caters to the user's likes as well. The space, therefore, constitutes a series of different event spaces where the user can mix visual experiences like art exhibitions with direct shipping for products. Depending on the user's will, the augmented space will show different configurations, from purely augmented space proposals that can interact with the user in a playful manner, more informative spaces like galleries and multiplayer experiences, or full-on virtual shop street that adverts and sales a product based on the users collected data. (Figure 8)

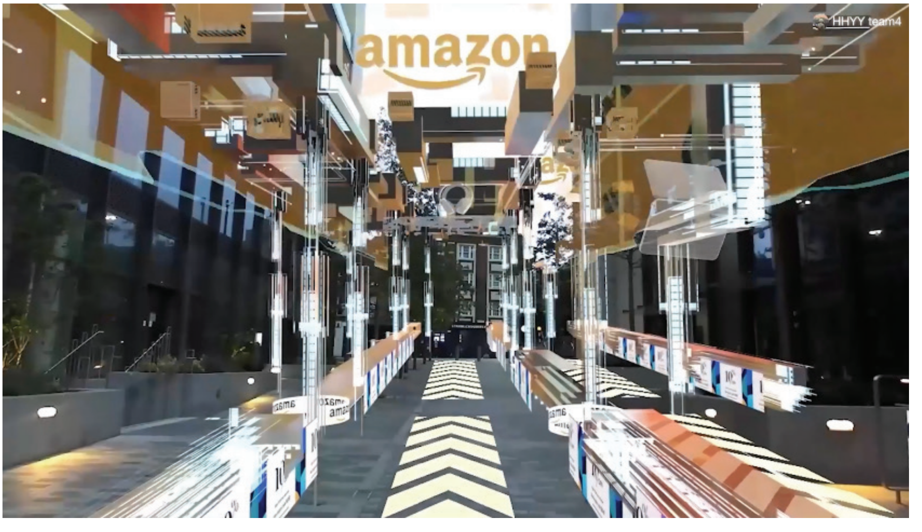


Figure 8. Augmented Space.

4. Results and Discussion

The exploration of the boundaries for Augmented Architecture is a long but paramount process. The future of architecture will be linked to the idea of the Metaverse, and the intersection between the two worlds is a relevant area to study. The projects described a speculative approach to how this interaction could be in the near future.

From the pure technological perspective, even though the capabilities of Augmented Reality hardware have rapidly increased during the past years, it is true that for now, its real impact is limited and mainly niche to the research world. The projects were primarily tested using a series of smartphones which allowed us to see a hint of the potential, but at the same time, they were limited in sensors and computing power. Regarding the software area, the technology seems to be ahead, being able to perform the actions explored with sufficient results. It is expected that in the near future, the software algorithms used for the project will improve and gain capabilities for more precise and agile results and interactions (Jordan & Mitchell 2015).

The proposed apps and scenarios show two of the possible uses that Augmented Architecture could have soon. From the functionalities to the ownership, the projects develop a case study on how this technology could affect and interact with us as users. In that sense, the Apps showcase a pretty realistic approach for the possible ownership of the augmentation and its possible scenarios of use.

From a comparative perspective, the two projects show different positions. While MiXR proposes more of an open case where the possible uses of the app, whether positive, neutral or negative, are tested. Problems like the ownership of the augmentation and what that would mean regarding social censorship are questioned and explored with the three scenarios. On the other hand, XRef explores a more commercial and, in that sense, realistic approach where this technology can be used for mixed situations benefiting the user (with a new and configurable experience) and the private sector (by promoting shopping). What is interesting to point out from both cases is the fact that because this technology is based on simple mobile phone apps, access to it will be quickly mainstream and has the potential to reach a wide public.

The use of Augmented Architecture presents a number of challenges. Although some of them, like third-party ownership or urban editing, are explored here, more research will be needed to understand the real potential as well as the risks of this future area of architecture. It will be essential to understand not only how it works but the social impact it could have in the very likely near future.

5. Acknowledgements and Credits

This paper is based on the B.Pro AD Research Cluster 9 Agenda. The research corresponded with the 2020-2021 academic year. The RC9 was led by Alvaro Lopez Rodriguez, Igor Pantic and Jose Pareja Gomez. MiXR project team comprised XuanYang Chen, Pengfei Mu, Congyuan Yuan and Shuhui Yu. The XRef project team comprised Haonan Wang, Haoya Chen, Yuling Guo and Yingxuan Tang. The graphic material referencing the projects was produced entirely by the students.

Figures

Figure 1: Hyperreality. 2016. http://hyper-reality.co/assets/HQ_images/hyper-reality_03.jpg
Figure 2: Medusa Project. 2021. https://www.tindrum.io/

References

Assange, J., Appelbaum, J., Müller-Maguhn A., & Zimmermann J. (2016). Cypherpunks: Freedom and the future of the internet. OR Books.

Baudrillard, J., (1981). Simulacres et Simulation. Paris. Galilée

Cook, P., Crompton, D., Greene, D., Herron, R., & Webb, M. (1971). Archigram: Instant City, 1961-1971. London.

Coppens, A. (2017). Merging real and virtual worlds: An analysis of state of the art and practical evaluation of Microsoft Hololens, Master's Thesis, University of Mons.

Gatys, L.A., Ecker, A.S., Bethge, M. (2016). "Image Style Transfer Using Convolutional Neural Networks," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 2414-2423, https://doi.org/10.1109/CVPR.2016.265.

Jordan, M., & Mitchell, T.M. (2015). Machine Learning: Trends, Perspectives, and Prospects. Science (New York, N.Y.). 349. 255-60. https://doi.org/10.1126/science.aaa8415.

King, B. (2016). Augmented: Life in the Smart Lane. Singapore: Marshall Cavendish International.

Kim, H., Lee, S., Lee, H., Hahn, T.,& Kang, S.(2019). Automatic Generation of Game Content using a Graph-based Wave Function Collapse Algorithm. 1-4. https://doi.org/10.1109/CIG.2019.8848019.

Matsuda, K. (2016). Hyperreality. Accessed May 16, 2016. http://km.cx/projects/hyper-reality

Milgram, P. & Kishino, F. (1994). A taxonomy of mixed Reality visual displays. IEICE Transactions on Information Systems, E77 (12).

Merrell, P. (2007). Example-based model synthesis. I3D '07: Symposium on Interactive 3D graphics and games. 105-112. https://doi.org/10.1145/1230100.1230119.

Lanier, J. (2013). Who Owns the Future? New York: Simon & Schuster

Levy, S. (1984). Hackers: Heroes of the Computing Revolution. New York. Doubleday.

Liu, S., Zhao L., & Li, J. (2012). "The Applications and Summary of Three Dimensional Reconstruction Based on Stereo Vision," International Conference on Industrial Control and Electronics Engineering, pp. 620-623, https://doi.org/10.1109/ICICEE.2012.168.

Srnicek, N., & Williams, A. (2016). Inventing the Future: Postcapitalism and a World Without Work. London & New York: Verso Books.

Unity Technologies. (2019). Unity engine. Version 2019.3.5f. http:// www.unity.com.

Vlahakis, V., Karigiannis, J., Tsotros, M., Gounaris, M., Almeida, L., Stricker, D., Gleue, T., Christou, I. T., Carlucci, R., & Ioannidis, N. (2001). Archeoguide. Proceedings of the 2001 Conference on Virtual Reality, Archeology, and Cultural Heritage. https://doi.org/10.1145/584993.585015

Winters, T. (2021). The Metaverse: Buying Virtual Land, NFTs, VR, Web3 & Preparing For the Next Big Thing! Self-Published.

Arth, C., Grasset, R., Gruber, L., Langlotz, T., Mulloni, A., & Wagner, D. (2015). The History of Mobile Augmented Reality. Inst. for Computer Graphics and Vision. Graz University of Technology, Austria

Sunger, İ., & Çankaya, S. (2019). Augmented Reality: Historical Development and Area of Usage. Journal of Educational Technology and Online Learning. 2. 118-133. https://doi.org/10.31681/jetol.615499.

Isola, P., Zhu, J-Y., Zhou, T., & Efros, A.A. (2016), Image-to-Image Translation with Conditional Adversarial Networks,arXiv, https://arxiv.org/abs/1611.07004

Tschumi, B. (1994). Architecture and Disjunction, The MIT Press, Cambridge MA & London Uk.

Zhang, J., & Hu, J. (2008). “Image Segmentation Based on 2D Otsu Method with Histogram Analysis.” 2008 International Conference on Computer Science and Software Engineering 6 (2008): 105-108.

FABRICATION OF COMPLEX CLAY STRUCTURES THROUGH AN AUGMENTED REALITY ASSISTED PLATFORM

JIAXIANG LUO[1], EFTHYMIA MASTROKALOU[2], RAHAF AL-DABOUS[3], SARAH ALDABOOS[4] and ALVARO LOPEZ RODRIGUEZ[5]
[1,2,3,4,5] *University College London*
[1]*luojiaxiang0610@foxmail.com, 0000-0002-6982-3692*
[2]*efimastrokalou@gmail.com*
[3]*rahafaldabbous@gmail.com*
[4]*skaldabbous@hotmail.com*
[5]*alvaro.rodriguez.14@ucl.ac.uk, 0000-0002-5561-1365*

Abstract. The relationship between clay manufacturing and architectural design has a long trajectory that has been explored since the early 2000s. From a 3D printing or assembly perspective, using clay in combination with automated processes in architecture to achieve computational design solutions is well established. (Yuan, Leach & Menges, 2018). Craft-based clay art, however, still lacks effective computational design integration. With the improvement of Augmented Reality (AR) technologies (Driscoll et al., 2017) and the appearance of digital platforms, new opportunities to integrate clay manufacturing and computational design have emerged. The concept of digitally transferring crafting skills, using holographic guidance and machine learning, could make clay crafting accessible to more workers while creating the potential to share and exchange digital designs via an open-source manufacturing platform. In this context, this research project explores the potential of integrating computational design and clay crafting using AR. Moreover, it introduces a platform that enables AR guidance and the digital transfer of fabrication skills, allowing even amateur users with no prior making experience to produce complex clay components.

Keywords. Computer Vision, Distributed Manufacturing, Augmented Craftsmanship, Augmented Reality, Real-time Modification, Hololens.

1. Introduction

Digital craftsmanship describes the involvement of the digital aspect in the process of making (Parisi, 2019). Craftsmanship on its own leaves us at the mercy of individual talent and skill. With the introduction of machine logic, the fundamentals of making have shifted into a more in-depth understanding of the matter (Wang, 2009). The

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 413-422.

certainty represented by the machine and the physical skill of the creator is combined through digital craftsmanship by developing a harmonious dialogue between them (McLuhan, 1964). Based on this concept, this paper will explore Augmented Reality as a tool that engages the maker in the fabrication process, builds a stronger relationship between creator and matter, and strengthens their interaction.

This study argues for an Augmented Reality assisted crafting process focused on both the fabrication of single components and on the generation of bigger assemblies. The components are developed through an innovative method of ceramic making that proposes a skeleton of sticks and woven rope to which clay is applied.

From the beginning, it was determined that this system would be accessible to all whether they were high-skilled workers, designers, artists, or users enthusiastic about ceramic making. As a result, a digital platform (ceARamics) was created that was divided into parts: "design" and "make". In the design section of the application, the designer/architect can import a volume, control specific data inputs, generate clay parts, choose among weaving styles and control the overall density. When the process is finished, a request for fabricating one's design can be made. In the "make" section of the application, users can implement clay components by following simple holographic instructions. Those instructions are embedded in the platform from other users that have recorded their weaving styles and techniques, thus transferring their skills. The whole idea of the ceARamics application is based on distributed manufacturing, enabling users with only the use of their phones and 3d printed nodes (which can either be 3D printed locally or be delivered to them) to fabricate without the use of expensive gadgets. The process is multidisciplinary and expands the production chain allowing for a fully democratised manufacturing process that is enabled through AR technology.

2. Methodology

This research aims to find an accurate and fast method for manufacturing and assembling complex clay structures. Under the guidance of holographic introductions, a method of making and assembling clay model structures is developed for users of various skills, from inexperienced to experts.

Material testing preceded digital model simulations. After studying different clay moisture ratios and structural options, the research team established a material workflow that allows for rapid modelling that does not produce cracks and affect firing, as well as a template system that allows for guided assembly using AR tools. The AR system was developed based on Hololens2 and mobile devices, making it accessible.

To enable large-scale model stitching, the geometrical representations that CeARamics works with are polyhedrons that can be replicated in an infinite number of stitches. The generative logic is provided by the Grasshopper© plugin WASP©. Components can be stitched together according to custom rules. Growth simulations that were used include growth along curves or surface constraints and aggregations through point clouds. The defined rules can be changed at any time during the procedure. These geometries can be replaced by digital clay models using computer simulation, so that designers can get a real-time simulation of the result via AR.

In terms of applying the system in architectural design, the research was mainly focused on the development of ceramic building facades such as shown in Figure 1.

Figure 1. Facades

Although ceramic facades exist today they are limited in terms of form and volume. Therefore, a system is introduced that can produce volumes and elevations defined by their complexity and plasticity, which is possible by the ceramic manufacturing method proposed in this paper.

3. Digital and Augmented Craftsmanship

After the industrial revolution, technology was introduced into the manufacturing process, changing the norms of crafting (Carpo, 2017). Even though automation processes in manufacturing reduce production time to a certain extent, there are cases where digital tools can become a burden due to the weakness they demonstrate in tasks that include human-like movements. A robot or an automated function may sometimes lack the ability to handle or manage an object in settings that require fine muscle control (Dellot and Wallace, 2017). Although specific tasks of production can be automatable, machines demonstrate inability to undertake a complete process because it is challenging for them to shift between tasks. In this context, human presence is required to oversee this shift or to complete the rest of the process.

Augmented craftsmanship is an approach to digital crafting that combines both digital tools and physical craftsmanship. Augmented craftsmanship can be seen in newly developed tools (such as HoloLens) that overlap virtual and real environments. It provides the opportunity to include more of the human factor within the fabrication process by broadening the creator's experience and developing further connections between the designer and the fabricated matter.

Subsequently, AR technologies may be the next step toward digital crafting by establishing a dialogue between automated processes and craftsmanship. Nowadays, more and more designers and architects are implementing AR technologies in their projects. Such an example can be seen in the woven steel pavilion implemented in the CAADRIA 2018 Workshop. The pavilion represents how designers can be holistically engaged in the process of making, by developing a digital platform that enables them to create interactive holographic instructions that translate design models into intelligent processes enabling unskilled construction teams to assemble complex structures in short time frames and with minimal errors (Jahn et al., 2018).

In the fabrication process presented in this paper, weaving through the frame of the components was an essential step. However, because of the scale of the components and the small distance between the sticks weaving as a task was too complex to be performed by a robot (robotic arm) so the human factor needed to be involved. Furthermore, for the success of the physical models, the maker was required to have a good understanding of their form and materiality. At the same time, it was important to generate a discipline through which the components could be mass-produced while maintaining their highly customisable features. Augmented Reality was selected as the tool that would enhance the user's natural crafting abilities and would accelerate the manufacturing process by facilitating mass production and customisation using human labour guided by HoloLens.

On the weaved frame of the components, sprayed clay was applied. This method, as opposed to using moulds or 3d printers, allowed the clay element to maintain its organic features and unpredictability. The idea was to be true to the matter and to generate a method that best represents its identity by respecting and testing its limits (Picon, 2019).

4. CeARamics Application

The components that were explored in this process were based on geometries of tetrahedrons and pentahedrons because compared to orthogonal shapes they provide numerous possibilities for combinations in all three axes and can be attached by all their faces. In order to establish the fabrication method presented, material testing was implemented as well as a trial-and-error process for the development of a rigid and feasible substructure that could be removed during the firing phase. As illustrated in Figure 2, for each component, nodes are placed on the corners of the polyhedrons that hold the sticks in place. These nodes are open from both sides, making it possible to pull away the sticks once the clay is dry. Weaving in patterns between the positioned sticks follows, creating the faces of the polyhedrons. After that is completed, clay is applied.

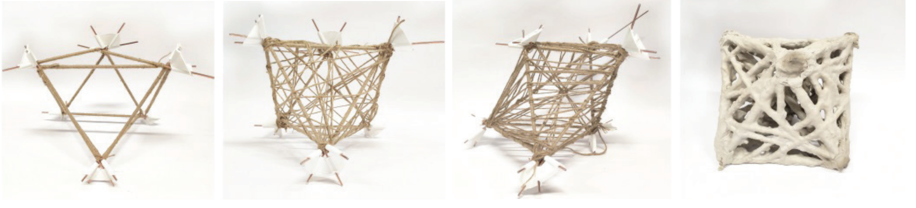


Figure 2. Component fabrication

4.1. AUGMENTED REALITY FOR THE DEVELOPMENT OF COMPONENTS

In the "make" section of the application dedicated to the manufacturing of components, users are able through their phones or Hololens to launch the Augmented Reality simulation displayed in Figure 3, that positions the outline of the component in space as well as where the sticks and nodes should be. After the user has assembled the frame, weaving zones with lines within them appear so that it is easy to follow the weaving patterns with the rope. This simplifies the process of assembling the component and makes it faster, even for someone who is doing it for the first time. After having assembled and weaved the total number of components that each user was assigned to, the users collaborate with designated ceramist workshops to apply clay and glazes.

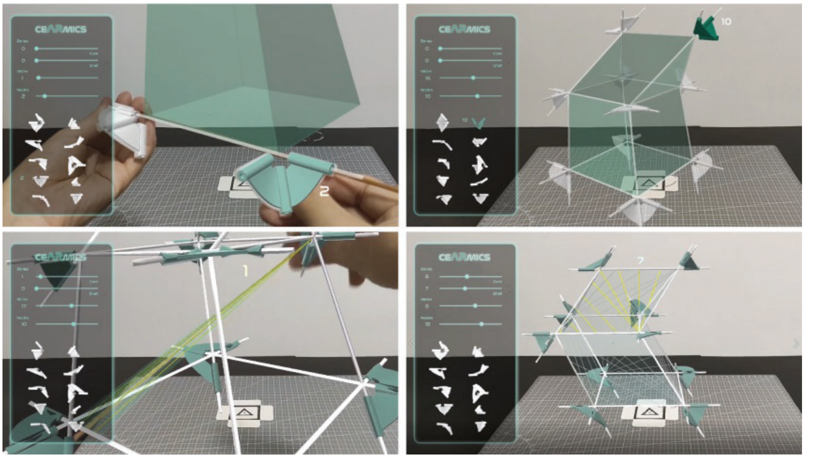


Figure 3. AR simulation for component manufacturing

4.1.1. Application of weaving styles

Computer vision is a field of artificial intelligence that trains computers to interpret and understand the visual world by replicating parts of the complexity of the human visual system. Using digital images, machines can accurately identify and classify objects — and then react to what they "see." (Jahn et al., 2019).

In this research, computer vision was explored as a way to identify weaving patterns-styles that would then become digitised and easily followed by users. In a scenario where different people weave on the same physical component, it is understandable that each individual would intuitively weave differently. Therefore, in the design part of the app, the option of weaving styles was incorporated. This option was crucial for architectural applications such as facades, where the components must exhibit some degree of uniformity. Two methods were explored. The first was movement tracking and the second was drawing curves on a physical model with the guidance of AR. The goal for both cases was for the information to become digital and be incorporated into the application so that the users could follow their desired weaving styles.

Object tracking is a field within computer vision that tracks objects as they move across a series of video frames (Wang, 2019). This works by creating a unique ID and specifying a position for the object which will be identified in each frame. Tracking was used as a way to determine weaving motion, by placing a sphere of distinguishable colour at the end of the rope as shown in Figure 4. This process generates the coordinates of the sphere while it moves, thus composing the "weaved" curves. However, this process wasn't entirely successful because when a person weaves, their hands or even the rope are very likely to interfere, disrupting the process.

For the second method, the faces of a component are unfolded digitally so that it becomes 2-dimensional. This is because the application allows drawing on each face only in 2D. Figure 5 demonstrates how the user is enabled through AR to virtually draw "weaving lines" through the screen of a mobile phone on the physical model.

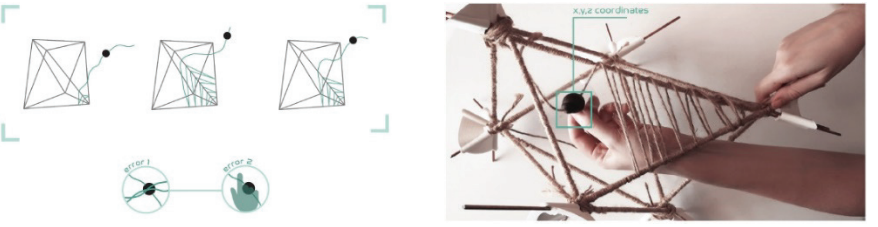


Figure 4. Sphere tracking

Those lines are directly transferred as curves to the design program and can then be attached to the unfolded 3D model through the app. This allows users to rapidly decide on the weaving style they prefer and instantly have the digital model of it.

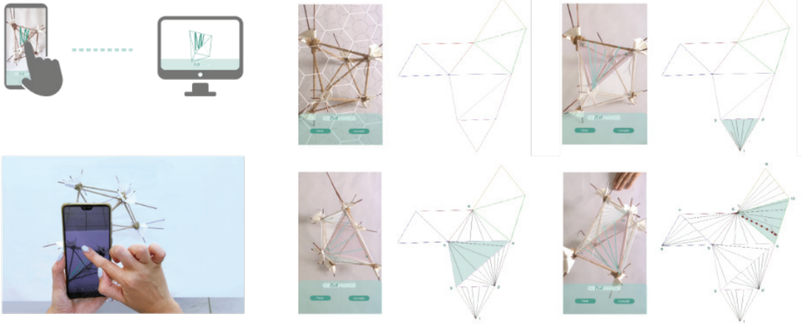


Figure 5. Weaving style recognition

Nine basic weaving styles were developed for the style catalogue, from which users can select one for their design. In the making part of the app, users are guided through AR to weave based on the specific style the designer has chosen.

4.2. AUGMENTED REALITY FOR ASSEMBLY

The assembly of components to create bigger parts was studied both physically and digitally. Augmented reality through the application supported both and one was dependent on the other. The design section of the ceARamics platform allows the population of digital clay elements based on a generative algorithm. The generative logic of component growth is based on custom rules and includes all information necessary for the aggregation process (geometry, location of connections, and orientation). Therefore, the type of components and their position when producing a bigger aggregation is determined by this algorithm. The types of components used were:

- The basic tetrahedron and pentahedron named Type-V + Type-T displayed in Figure 6.

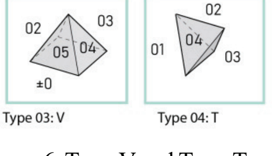


Figure 6. Type-V and Type-T

- Geometrical representations of orthotetrahedrons and pentahedrons named Type-L and Type-S, displayed in Figure 7. These types allow for a faster and more random generation of aggregations compared to the basic tetrahedron and pentahedron since they contain more faces that enable a wider variety of connections. However when the growth simulation was performed, it was found that the results performed directional convergence. The main reason is that due to the specificity of the geometry there is only a 20% possibility to change the growth direction of the components. In order to apply directional variety it was necessary to add Type V and T as a transformation medium.

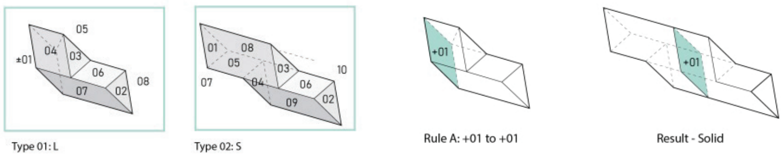


Figure 7. Type L and S combination

In the end, four sets of rules that used all types of components were selected to generate different discrete design results. For example, the wall generation results displayed in Figure 8 are the outcome of these different set of aggregation rules.

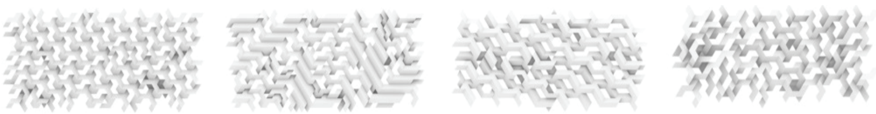


Figure 8. Walls according to sets of rules

In the design section of the platform, designers can import their geometries (surfaces, voxels) to the app and generate solid components on them based on the generative algorithm. Using the WASP© replacement algorithm, the solid components can be replaced by digital clay parts of varying weaving densities. This process relies on mechanical simulation analysis as well as functionality. Depending on the type of space the geometries will enclose (public, semi-public, private) the porosity of clay parts is adjusted accordingly. In terms of mechanical analysis, the generated part obtains the corresponding force areas, which are replaced by clay modules of relevant strengths to obtain a reasonable design result. This process is described in Figure 9 below.

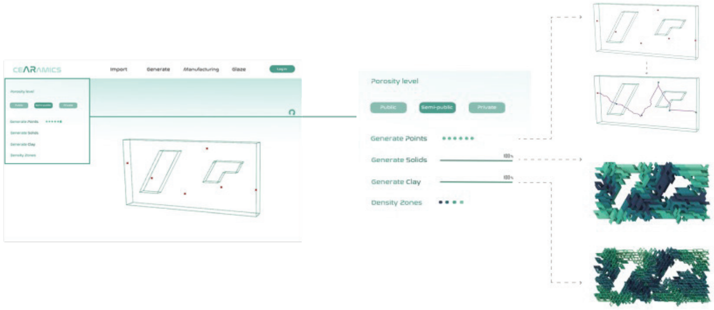


Figure 9. Replacement of solid components with clay parts

Users can finally choose from a variety of weaving styles for their design and then view the glazed version of it by selecting a colour from the palette. The designer can then apply through the application for the generated part to be manufactured. The clay elements will be assigned to manufacturers worldwide that will produce them guided by the application. Once fabrication is complete, the components are shipped to the building site. Construction workers can use AR to view the digital aggregation within the real environment. They can then select a portion of the aggregation to assemble and begin construction. The prospect for this platform is that it is meant to be accessible on different devices (Hololens, tablet, phone). Thus, on their chosen device, manufacturers can view the aggregation in segments consisting of three different colours of solid components as shown in Figure 10. Each colour corresponds to a different density of components. The combination of clay parts is a cladding system positioned on special panels that are attached on a building like a curtain wall. The components themselves are non-structural. In the occasion when the edges of the components are too irregular because of the organic form that clay produces, they are sanded and smoothened in order to be well-adjusted and fixed on the panel.

Figure 10. Assembly of aggregations

5. Making the platform accessible

When digital tools were first developed, they were handled in a way in which users could process, alter and communicate information individually rather than in a joint effort (Carpo, 2011). However, design and production started to overlap at the turn of the 21st century with the emergence of digital tools that allowed for generative and automated processes through information modelling and contemporary fabrication methods (Carpo, 2011). Modern technology allows architects through digital communication to control what is being implemented and cultivates solid connections between the different disciplines involved in the process of making (Celanto, 2007). Furthermore, digital architectural software has evolved in a way that emphasises its participatory nature and, except for the professionals involved, creates platforms for citizen and community participation as well (Carpo, 2011).

The digital era brought the shift from mass production to mass customisation, which entailed variability and a process of specialist and social involvement in all stages. Therefore new models of shared ownership and authorship are developed (Michalatos, 2016). An important part of the holistic system that was introduced is open-source networks that allow interaction and citizen participation (Sanchez, 2017).

Within this context of a holistic system, the process that is analysed in this paper aims to create a platform that offers the tools and means for decentralised forms of production within a system that enables social empowerment.

When it comes to manufacturing through the ceARamics application for the building industry, it was critical to invent a way for mass production. Instead of following a centralised manufacturing procedure, it was questioned whether Augmented Reality could activate a system of distributed manufacturing. In that context, the application users are assisted throughout all the needed steps with an AR simulation which they can launch through their phones. Whether fabrication is implemented by specialists, workshops that engage communities, or even individuals fascinated by ceramic making, the aim is definitely for a decentralised procedure that expands the production chain and democratises the process. The nodes, sticks and rope that are needed for component assembly are shipped as a kit to the users, and then as illustrated in Figure 11, users only need their phones to operate. The fact that this process does not require expensive equipment to function adds to its accessibility and makes it more inclusive.

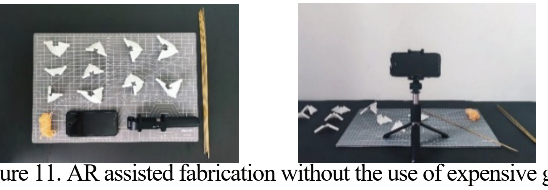


Figure 11. AR assisted fabrication without the use of expensive gadgets

In addition, it is understood that since the fabrication process could be implemented with the participation of people that could be anywhere on the planet, what was previously introduced by Michalatos (2016) about shared authorship is also applied here. The outcome is a collaborative effort of various contributors which also adds to the development of a genuine democratised system.

6. Conclusion

The use of clay as a building material has been a constant throughout architecture history. In recent years, its integration with digital and automated technologies has collaborated to keep it as one of the more relevant materials, as well as sustainable. By introducing Augmented Reality assisted manufacturing into the process, the catalogue of possibilities is expanded, and traditional and new techniques, can now be combined together with the prevalent industrial manufacturing of the material.

This paper presents the idea of a platform that allows for an AR-assisted crafting method to develop clay components and their assemblies. The platform has the potential to guide the fabrication process with holographic instructions and allows for component population through a generative algorithm that provides various design options. This solution for merging physical craftsmanship with the efficiency of technology allows non-skilled workers to participate in the process, enabling a larger community to access and share manufacturing skills and preserve them on a database.

Although the project is still in the development phase, and more research and testing would be required, the capacity to record and deliver skills to untrained users has been tested through prototypes that show the potential of this AR assisted crafting system. In contrast to other mainstream automated manufacturing processes, the physical crafting characteristics of clay enable this method to be more intuitive, participatory, and community conscious while preserving the heritage of clay craftsmanship, demonstrating how craftsmanship can continue to adapt to future technologies. The presented manufacturing system is a model seeking to explore the limits of a concept like this and stimulate thought about how it can be applied to other production models as well.

Acknowledgements

The research was developed in The Bartlett School of Architecture, UCL 19/20, within the Research Cluster 9 Agenda of the B.pro AD M.Arch. The project was supervised by the cluster leaders Alvaro Lopez Rodriguez and Igor Pantic (0000-0001-8802-5592).

References

Carpo, M. (2011). The alphabet and the algorithm. Mit Press

Celento, D. (2007). Innovate or perish: new technologies and architecture's future. Harvard Design Magazine,26(Spring/Summer).

Dellot, B., & Wallace-Stephens, F. (2017). The age of automation. *Artificial intelligence, robotics and the future of low-skilled work*, 44.

Driscoll, T., Farhoud, S., & Nowling, S. (2017). Enabling mobile augmented and virtual reality with 5G networks. *AT&T: Dallas*, *TX, USA.*

Jahn, G., Newnham, C., & Beanland, M. (2018). Making in Mixed Reality. Holographic design, fabrication, assembly and analysis of woven steel structures, DOI:10.52842/conf.acadia.2018.088

Hahm, S., Maciel, A., Sumitiomo, E., & Lopez Rodriguez, A. (2019). *FlowMorph-exploring the human-material interaction in digitally augmented craftsmanship. CAADRIA 2019* (pp. 553-562). The Association for Computer-Aided Architectural Design Research in Asia (CAADRIA), https://doi.org/10.52842/conf.caadria.2019.1.553.

Laing, R. (2018). *Digital participation and collaboration in architectural design*. Routledge.

McCullough, M. (1996), Chapter 7: Medium. In: *Abstracting Craft, The practiced digital hand*, Cambridge, Massachusetts, London, England, pp. 191-220, https://doi.org/10.4324/9781315161488.

McLuhan, M., & Fiore, Q. (1967). The medium is the message. *New York, 123(1)*, 126-128.

Michalatos, P. (2016). Design signals: The role of software architecture and paradigms in design thinking and practice. *Architectural Design,86(*5*)*, 108-115. https://onlinelibrary.wiley.com/doi/abs/10.1002/ad.2096

Parisi, L. (2019). Media ontology and transcendental instrumentality. *Theory*, *Culture & Society,36* (6), 95-124, https://doi.org/10.1177/0263276419843582.

Picon, A. (2019). Digital Fabrication, Materiality and Nostalgia. In *Robotic Building* (pp. 112-116), DOI:10.11129/9783955534257-024.

Sanchez, J. (2017). Combinatorial Commons: Social Remixing in a Sharing Economy. *Architectural Design*, *87* (4), 16-21.

Wang, X. (2009). Augmented reality in architecture and design: potentials and challenges for application. *International Journal of Architectural Computing*, 7(2), 309-326, DOI:10.1260/147807709788921985

Yuan, P. F., Menges, A., & Leach, N. (2018). *Digital Fabrication,* Tongji University Press

DESIGNING FOR THE LIVING BUILDING CHALLENGE

An Integrated Material Environment Visibility Framework for New Zealand Designers

CHITRAKALA MUTHUVEERAPPAN[1], MARC AUREL SCHNABEL[2] and RYAN SIMPSON[3]
[1,2] *Victoria University of Wellington, NZ.*
[3] *Frequency, NZ*
[1]*chitrakala.muthuveerappan@vuw.ac.nz, 0000-0003-4460-1193*
[2]*marcaurel@gmail.com, 0000-0002-2923-1609.*
[3]*ryans@frequency.nz*

Abstract. To achieve rigorous performance in sustainability toward zero-carbon in the Built Environment, the New Zealand government also promotes Living Building Challenge (LBC) in design. The LBC seeks detailed material specifications till disposal with environmental impacts to measure sustainability in design and insists 75% of the project materials supply within 90 km of the project location. This study explored the constraints related to LBC material requirements and found the major barriers for designers are a) the sustainability imperatives namely red list and material lifetime are not considered by online material catalogue providers, b) no material data is shared among the primary stakeholders for acceptance and compliance evaluation, and c) no facility to identify the supply chain and suppliers within a certain distance of the project location. As a solution, this research designed a hybrid material catalogue having all possible material parameters classified into 5 major categories and developed an integrated material environment visibility framework linking the processes among the designers, contractors, suppliers, and the client. This proposed integrated framework provides a shared environment on material viewing, similar alternatives, accurate compliance evaluations in design, material suppliers availability, and other supply chain details at design stage itself. This new framework is currently under industry evaluation for compatibility, reliability, and industry professionals' acceptance.

Keywords. Building Material Supply Chain Management, Centralised Construction Material Catalogue, Integrated Building Material, Living Building Challenge, Sustainable Design Materials.

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 423-432.

1. Introduction

New Zealand is the sixth emitter of greenhouse gases in the world emitting 17.2 tonnes of carbon dioxide equivalent per person should play a vital role in environmental protection and climate change (Net-Zero, 2022). To achieve rigorous performance in sustainability, the New Zealand (NZ) government insists on zero-carbon and promotes Living Building Challenge (LBC) in the built environment. Subsequently, due to the geographical position of NZ on earth, building material production and availability are very much scarce and the recent pandemic worsens the material supply chain and designers are suffocating to find alternatives (Pointon, 2022). The challenges and barriers to implementing zero-carbon in the built environment are multifarious. One of the challenges is the material selection and its supply chain providing visibility in protecting the environment which serves as the base for sustainability needs immediate attention. Hence, the LBC implementation issues being raised at the design phase should be collaboratively solved by the designers, contractors, and suppliers to reduce the impact at a later stage and this will ensure all stakeholders' involvement in achieving LBC adoption in projects. This paper explains the research developed an integrated material environment visibility framework on material selection and its supply chain that sophisticates the NZ designers under one roof on cloud to achieve sustainability in design is detailed.

2. Living Building Challenge Implementation Issues in Projects

Living Building Challenge (LBC) is an advocating tool to measure sustainability in the built environment for building certification. Under the sustainability context, the LBC criteria have been comprised of seven 7 petals and are expressed as 1) Place - assures the relationship with nature, 2) Beauty - expresses the value of the human spirit, 3) Health - provides physical and psychological well-being of humankind, 4) Water - balances with climate and environment, 5) Energy - enhances solar usage, 6) Materials - of long last and biodegradable, and 7) Equity - to humankind (LBC 4.0, 2016). Evidence found that the first five criteria of LBC have been considered in design through various design processes and philosophies (EVCI, 2015; Isah. et al., 2018). The sixth criterion is the material used in the design and the seventh criterion insists on nations' governing policies in implementation.

This LBC certification of a building project can have any one of the following out of three categories namely a) Living building certification complies with all types of imperatives, b) Petal certification satisfies three or more petals achievement towards living building certification, and c) Net zero energy building certification (BXG, 2017).

Towards sustainability, the NZ government initiated the Zero-Carbon policy and announced a net zero carbon excluding methane by 2050 (MBIE, 2017.) Since Designers have not been given sufficient material information with certain technical details making it is difficult to know whether expected compliance is achieved (AIA, 2016; Dowdell. et. al., 2017). To endorse sustainability in projects, many NZ building projects are keen to obtain any one of the LBC certifications. While designing projects for LBC certification, one of the major issues among designers is the selection of bio-degradable materials or proving the red zone imperative documents on materials that could add visibility to the environment (Ryan, 2021).

In practice, designers are selecting materials by referring to suppliers' catalogues or word of mouth, and by surfing local online material libraries. Unfortunately, adequate and sufficient parameters of material specifications related to LBC are lacking in the current referencing practice for accurate decisions in the material selection that supports government policies. These multiple search ways in identifying materials, analysing their LBC compliance, and identifying alternatives due to lack of supply during construction is not only time-consuming but can cause delays at any stage throughout a project impacting overall productivity (Studer and Mello, 2021).

Though the net zero energy-certified building numbers are significant in NZ, very few studies have examined the adoption challenges and barriers, especially on materials in NZCS. One of the LBC petal-certified NZ projects called the Glenorchy project has tabulated a detailed percentage of red-list imperatives of the materials used. The Glenorchy project (2022) informed that the material criterion is the core competency of LBC and is highly challenging for NZ Designers and Contractors.

Above all, the LBC certification demands 75% of the materials should be supplied within a minimum of 90 km of the project location (MBIE 2017). This increases the material selection complexities to comply with local policy on the local material use percentage in design. In addition, while contracting, identifying suppliers within a certain distance whether urban or rural is highly challenging to the project's success. Hence, material supply chain consideration becomes significant.

In conclusion, local material use and materials' environmental visibility should negate the materials' supply chain, expecting the Architects to regain the architectural creativity that is lost through material constraints. Furthermore, this added visibility could enhance the likelihood of new developments achieving tougher sustainability certifications supporting the LBC. These issues need to be collaboratively solved by the design team, contractors, and suppliers to ensure the local policies that all stakeholders involved in the project find optimized ways to speed up the project's completion.

3. Research Aim and Objectives

Having identified current challenges in material for LBC certification, this research aims to develop an integrated environment visibility framework among primary stakeholders of Designers' material selection is proposed. To achieve this aim, the following objectives are set forth:

- Explore and identify the design material constraints under the NZ context.
- Design and propose an integrated material environment visibility framework among the primary stakeholders' material selection processes.

Since a project's material selection is prior to council approval, the concerned people in material use and acceptance in design are the designers, contractors, and the client who are involved in the design, and the supply availability and distance requirements included Supplier involvement and are the primary users of this work. Identifying all adequate and sufficient parameters of existing materials in NZ for LBC certification is out of this research scope.

4. LBC Design Material Constraints to Designers

Designers are the predominant users of materials around the world nevertheless of size, units, and categorisation of material formats. The pertinent literature review on design materials informs that NZ and Australia use Coordinated Building Information (CBI) format whereas countries namely the USA, Canada, Europe, Scotland, and many other countries use Master Specification Divisions (MSD) (Allison and Hartley, 2020; TFG,2018; EVCI, 2015).

To analyse the material parameters on LBC requirements, NZ's major online material catalogues namely EBOSS, Product-Spec, and Master-Spec that all follow the CBI format and are considered for the review. The CBI classifies all materials into 8 categories namely general, site, structure, enclosure, interior, finish, services, and external and the MSD classifies their materials into 35 categories that reduce considerable material exploring time.

Table 1 indicates that the green-coloured columns provide the format provides that relevant parameter details and the red-coloured columns do not consider that parameter in providing details. MSD is lacking acoustic parameters whereas CBI is not considering many parameters, especially on the red list and demolition shows environment visibility closely associated with sustainability.

Parameters	MSD	CBI
Product_Name		
General category		
CSI category		
Collection/Type		
Size		
Standards		
Installation manual		
Brochure		
Colour range		
Additional information		
Fire safety		
Flame and smoke		
Fire certification document		
Acoustic performance		
Life expectancy		
Warranty document		
Phone number		
Email		
URL		
Address		
Light reflectance value (%)		
Environmental product declaration (pdf format)		
Cleaning and maintenance		
Material safety data sheet		
Material information sheet		
Health Product Declaration		
Sustainability certification (LEED, LBC, WELS, etc)		
PVC free		
VOC content (g/L)		
Tag words (Alternative search words)		
End of life 1 = recyclable 2 = landfill		
LBC imperative 10 Red list 1= compliant, 2 = non-compliant		

Table 1: The CBI and MSD format Material Parameters Comparison

Since CBI is lacking many material sustainability parameters, it is difficult to identify appropriate LBC compliance. Moreover, the material parameters are randomly listed, and the designers require extra effort to consider instead of within a glimpse. The CBI format requires environment visibility and additionally a rational approach to grouping the parameters that suit each stakeholder's needs. Hence, this research recommends a hybrid material catalogue of CBI and MSD formats having all the parameters grouped according to the nature of their description.

The recommended hybrid material catalogue comprises CBI and MSD material parameters (Table 2). Besides the hybrid of MSD and CBI parameters, additional parameters namely material origin, and supplier's geographic are added for local policy compliance. For clients' sophistication, material prices and stock information are included.

S.No	Category	Parameters
1	General Information	Product Name
		General category
		CSI category
		Collection type
		Size
		Colour range
		Material information sheet
		Tag words
		Country of origin
		Wholesale dealer
		Geographical constraints
2	Technical Data	Standards
		Fire safety
		Flame and smoke
		Fire certification document
		Acoustic Performance
		Light reflectance value
		Material safety data sheet
3	Sustainability	Environmental Product Declaration
		Health Product Declaration
		Sustainability certification
		PVC content
		VOC content
		End of life content
		LBC imperative 10 – Red List
4	Additional documentation	Installation Manual
		Brochure
		Additional Information
		Life Expectancy
		Warranty
		Cleaning and maintenance
		BIM Model
		CAD Model
5	Contact	Company URL
		Country
		Province/state
		City
		Postcode
		Address 1
		Address 2
		Phone
		Email

Table 2: The recommended Hybrid Material Catalogue Structure

For grouping of parameters, general information and technical details of material categories are common in practice. Since every designer's focus is on sustainability, relevant parameters should be grouped under this category. Additionally, any documentation that supports sustainability, technical, or general compliance should be compiled under a documentation category for additional referencing. Finally, if any material suitability is considered by the designers, then the suppliers' details are necessary to address the supply chain distance constraint.

Based on the practical sense of data access, the new hybrid catalogue's material parameters have been grouped into five major categories namely 1) General information, 2) Technical, 3) Sustainability, 4) Additional documentation and 5) contact information.

5. Integrated environment visibility framework among the primary stake-holders' processes

A collaborative contribution from all stakeholders ensures the success of a project. The designers, contractors, suppliers, and the client are the predominant players and become the primary users of a project. All remaining parties namely the council, project team, legal, Facility manager, education, etc., and whoever is associated with the project become tier-two users (Figure 1).

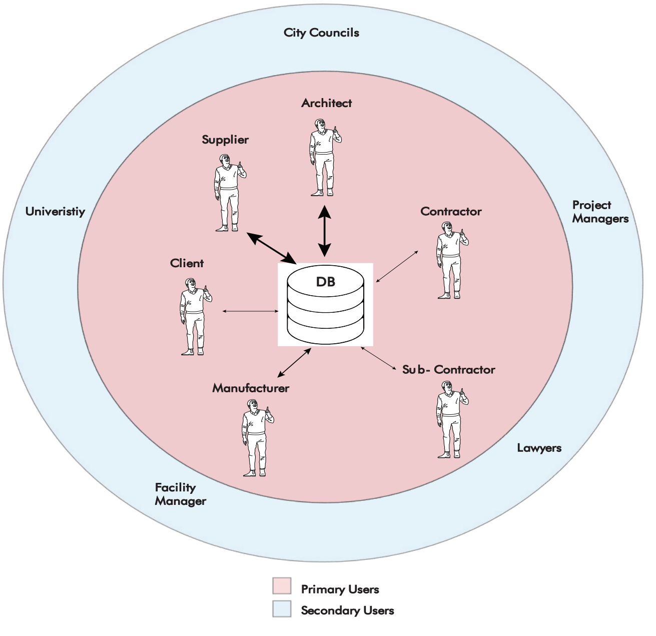


Figure 1: Two-tiered project stakeholders' involvement in a project

Since the client gives his acceptance of design material use in project for further process by the rest of the primary users, there is no specific process role for the client in the Framework. The integrated material environment visibility framework proposed by this work shares data via the primary user's processes. Each prime stakeholder's material environment visibility integration is carried out via their process and is explained as separate sub-sections.

5.1. DESIGNERS

The Designers have six main functions which are 1) material keyword search, 2) analysing material specifications, 3) visualising applied material in design, 4) customising project materials personal folder and checking its supply availability among the Contractor(s) and Client, 5) contacting the supplier for ordering, and 6) providing feedback for future reference (Figure 2).

The first three functions are standard practise between the designer and client however the innovation starts with the fourth function which enables the architect to customise their own folder of interested materials to be shared with the contractor and client for their comment and acceptance in design use. This function reduces the time taken for gaining producer statement approvals at building completion. Since entering the pandemic, stock availability is one of the main concerns when specifying materials. This eliminates the ability to purchase materials from outside the region which could cause significant delays and supports the local government policy. This additional visibility provides options for all businesses irrespective of their size and intended purchase value. This framework enables visualising all suppliers of the same and similar materials within the given area around the project location.

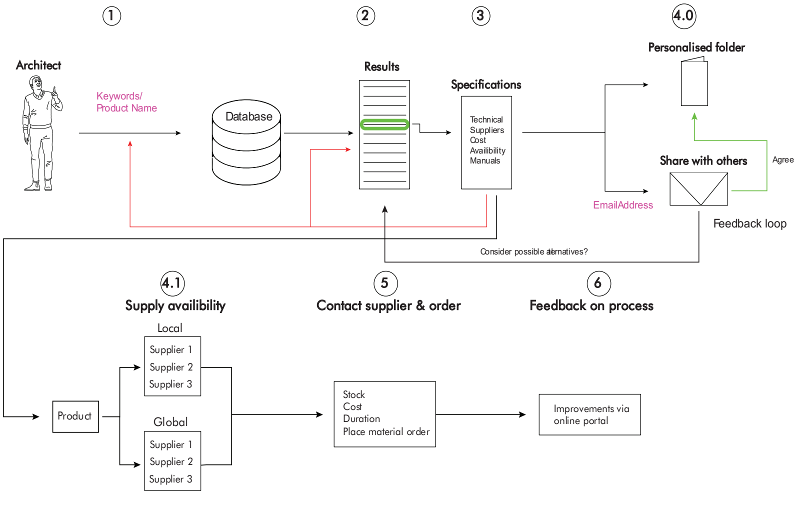


Figure 2: Designers Process Framework

5.2. SUPPLIERS

The supplier's role is pivotal to the framework's success as this connects both supply and demand. The supplier has 4 main roles namely 1) Assigning sub-suppliers, 2) Input/updating material data and change visibility graphics, 3) Issuing material update notifications, and 4) Responding to material queries (Figure 3). The aim of assigning sub-suppliers cater to show supply chain links in rural locations for contractors and designers. Also, It is the supplier's responsibility to update the material details whether minor or major changes have occurred for their product. Any new updated features or information release of materials can be sent to all stakeholders through a notification alert. Suppliers can also respond to the questions raised by other stakeholders within the framework.

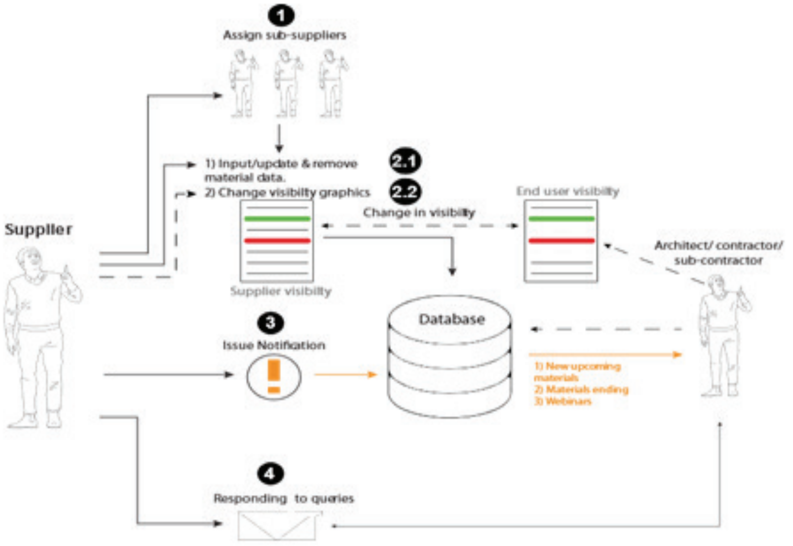


Figure 3: Supplier's Process Framework

5.3. CONTRACTOR

The function of contractors (and subcontractors) is to communicate with the architect/designers, analyse their material catalogue, and provide constructive advice that is either to approve, decline, or propose an alternative material. Though contractors and subcontractors have a similar function to the architect, their predominant function is to provide buildability advice through the feedback analysis loop (Figure 4).

The contractor will receive the URL link listing all proposed materials from Architect's customised project material personal folder for analysis. Upon review, the contractor can submit their response back to the architect for acceptance or material substitution. This enables synchronisation in material selection among the designers and contractor(s) that strengthen the sustainability evaluation accuracy at design phase. This process binds contractor(s) with the design team and their entire supply chain.

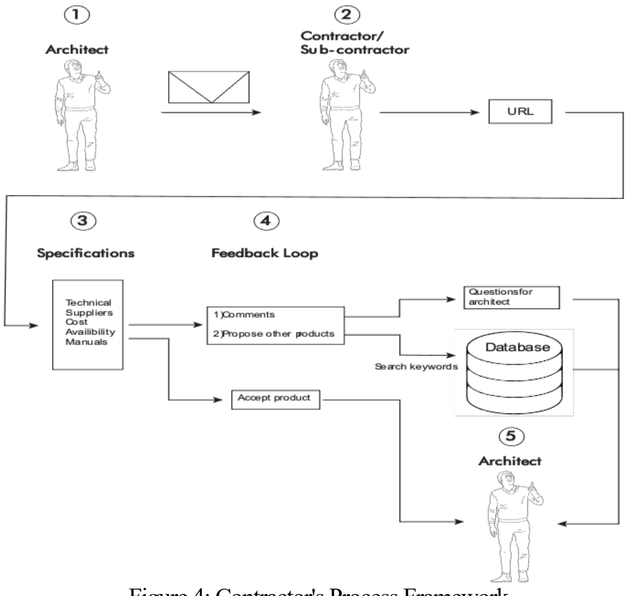


Figure 4: Contractor's Process Framework

6. The Proposed Framework Development and Testing

The proposed new hybrid material catalogue has been virtually visualised for real-time use (Figure 5). PHP was used to develop the Graphical User Interface (GUI) and MYSQL as a Database to store the material catalogue and other relevant information. The entire setup has been placed on the cloud at Victoria University of Wellington, managed by their Information Technology Service (ITS) department.

To access the digital catalogue, each user needs a secured username and password to log in to the website. prototype. According to their registered role (Client / Designer / Manufacturer / Supplier / Contractor), the proposed framework shows their dashboard

At present, the integrated environment visibility among the primary players in material selection and the suppliers' availability around the project location is under development.

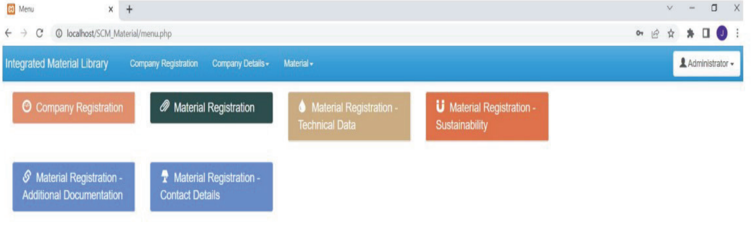


Figure 5: Prototype Dashboard

7. Conclusion

Designers are hardly provided with comprehensive material parameters and even if so, the provided parameters are inadequate and insufficient to evaluate environment visibility imperatives for LBH certification compliance in NZ. Limited consistency was found within the sustainability pertinent of CBI format needs synchronised addressing over MSD format and local policies. This research recommends a hybrid of CBI and MSD format's material parameters. The established two-tier framework identifies the predominant and beneficiary stakeholders in the project material supply chain. The functions of each stakeholder were detailed, and their process relationships are then generated into an integrated framework.

At present, the conceptualised hybrid material catalogue has been visualised digitally with limited functionality for real-time implementation and is under testing. The integrated material visibility environment framework is under development. The current research output is expected to test on a real-time project by the designers, contractors (and suppliers), and clients to gain constructive feedback on its adaptability and for further research and development.

References

Allison, N., & Hartley, G. (2020). *Digital product data for lifting productivity*. Wellington: BRANZ.

AIA. (2016). *Material transparency & risk for architects: An introduction to advancing professional ethics while managing professional liability risks*. Washington DC: American Institute of Architects

BXG - The Building Excellence Group (2017). *The Living Building Challenge - Building Rating Tools.* Retrieved September 12, 2020, from https://www.bxg.co.nz/lbc

Dowdell, D., Page, I., & Curtis, M. (2017). *Electronic traceability of New Zealand construction products: Feasibility and opportunities*. Wellington: BRANZ.

EVCI, M., & Ciravoglu, A. (2015). An evaluation of the role of environmental, social and economic factors in architects choice of building materials. *Megaron / Yildiz Technical University , Faculty of Architecture E-Journal.* Pg.139-147. DOI: 10.5505/MEGARON.2015.02886.

Isah, A., Muhammad, I., & Katun, I. (2018). Optimizing architects' building materials' specification strategy through knowledge management principles. *Journal of Construction Innovation*, Vol.18. Issue 4. Pg: 394-410.

LBC 4.0 (2016). *Living Building Challenge 4.0 - A Visionary Path to Regenerative Future.* The International Living Future Institute.

MBIE (2022). *Climate Change Response Act 2002 Updated 2022*. Ministry of Business and Employment, New Zealand Government.

Net-Zero. (2022). *Net-Zero - The Road to 2050 - Podcast.* Retrieved December 7, 2022, from https://netzero-nz.org/?p=321.

Pointon, N. (2022, February 1). Construction material shortages, price hikes plague building industry. Retrieved February 10. From https://www.rnz.co.nz/news/business/460620/construction-material-shortages-price-hikes-plague-building-industry

Ryan. S. (2021). *An integrated supply chain framework between designers and material library under a digital environment for efficient project management*. Victoria University of Wellington.Studer, W., & Mello, L. C. (2021). *Core Elements Underlying Supply Chain Management*. MDPI - Buildings, 1-20.

TFG -The Future Ground (2018). *Mullet Hall - LBC Materials Red Lists*. Retrieved November 3 (2021).

SEARCHLIGHT SPACE- A SENTIENT SPACE HAVING EXPLORATORY BEHAVIOUR

PAN-PAN SHIUNG[1] and JUNE-HAO HOU[2]
[1,2]Graduate Institute of Architecture, National Yang Ming Chiao Tung University.
[1]ppenny37tw@arch.nycu.edu.tw, 0000-0001-5539-3137
[2]jhou@arch.nycu.edu.tw, 0000-0002-8362-7719

Abstract. If a space is a sentient being, how would it recognize the environment to establish interactive relationships with people and objects? The modern living environment is full of ubiquitous computing and data flows, the smart space is equipped with rich technologies allowing data exchange between devices for the purpose of meeting human demands. However, the lack of development for a basic environment with cognitive abilities presents a challenge for creating an adaptable space, which can grow and control our surroundings to enhance our behaviour. Thus, the goal of this study is to propose a basic cognition of a sentient space. Firstly, the author designed the methodology for a sensorimotor through learning biological mechanism. Three adaptive strategies of exploratory behaviour are then developed to operate the cognitive map of the space. Lastly, the process is visualized to enable human interaction. Through the on-site installation, usability of the prototype is verified and positive user feedbacks were obtained. In conclusion, the ultimate goal of the research is to develop a space with the characteristic of a sentient being through a prototype, also, it allows users to understand how human activities affect the cognition embodied in the space.

Keywords. Autonomous Exploration, Adaptive System, Spatial Information, Cognitive Map, Human-Computer Interaction, Steering Behaviours

1. Introduction

Sentient beings have long been a fixture in human imagination due to our natural tendency to rely on and connect with the natural world. For example, the magic house Casa Madrigal in the movie Encanto serves as a vivid character with sentient and thinking abilities that play an important role in the story. Likewise, biophilic design, which aims to enrich our contact with the natural world, has also become a popular approach in the design of both objects and spaces, including homes and buildings.

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 433-442.

1.1. SMART SPACE AND SENTIENT SPACE

As society evolves into the post digital era, smart technologies constantly improve all aspects of human life through the development of novel technologies, such as artificial intelligence (AI) and the Internet of Things (IoT) — connected "things" that equipped with rich technologies to sense and react to specific requirements. By utilizing Simultaneous Localization and Mapping (SLAM), for example, smart devices are able to map and understand their environment in a dynamic way, leading to a multitude of new and improved applications such as autonomous navigation and augmented reality. In the field of architecture, smart spaces from homes and offices to smart cities are often linked to energy, sustainability, comfort, and convenience. A smart building operates through automation of operations in order to maintain its occupants' comfort, and it is a way of minimizing energy consumption (Lee et al., 2021).

Compared to the smart space, the consideration of spaces as conscious entities through pervasive computing and the public realm provides another significant perspective in architecture. Dana Cuff (2003) coined the term "cyburgs" to refer to "an environment saturated with computing capability." This concept reflects the idea that digital technology is becoming increasingly intertwined with our physical surroundings, and it further enhances agencies and people who live in the cyburg. "It is a world where we not only think of cities but cities think of us, where the environment reflexively monitors our behaviour." Crang and Graham (2007) further developed the concept of cyburgs in the framework of sentient city, where the environment is recursively influenced by social activity. When the environment transcends human thinking and become a provider of intelligence, it can be a part of ecological practice as well. Thrift (2014) argued that the sentient city is not aware of itself in any human way. We need to think about conglomerations of all kinds of entities but not from single human perspective. Consequently, it is important that we let things have their own forms of energy, tenacity and magnetism built out of forces of "outstinct" (as opposed to instinct). Therefore, this paper aims to answer the research question: "how can we build an environment for a space to embrace its sentient characteristic?".

The architecture framework for the behaviour of the adaptive system enables occupants to sense, experience and understand smart spaces (Lee et al., 2021). Nowadays, a number of adaptive prototypes have developed, such as the physiologically-driven "ExoBuilding" that can change shape in real time and provide immersive experience (Schnädelbach et al., 2016). This research focuses on how to build a cognitive map of a sentient space in an unfamiliar environment, particularly on making its action adaptive as well as purposeful.

1.2. COGNITION AND EXPLORATORY BEHAVIOURS

The research reviews the understanding of the behaviours of sentient being from the perspective of cognitive science. In general, exploitation and exploration are determinant adaptive processes that allow animals to build cognitive maps in ever-changing environments. These behaviours are generated as animals spontaneously switch between different modes such as perception, attention, memory, and motivation. For instance, zebrafish's (Danio rerio) exploratory behaviours, through repeated simple actions, can balance between novelty and predictability conditions, and finally achieve the spatial cognition of elementary animals (Zheltova & Nepomnyashchikh, 2020 ; Red'ko et al., 2015). In human child development stage, motor babbling is a process where the brain learns the relationship between muscle activity and actions, with this kind of repeated cycles of perception and action, allowing infants to gains an understanding of object properties and interact with them (Corbetta et al., 2018).

Moreover, animals' attention ability can be aware of specific conditions. This physiological basis has been widely compared to a "spotlight" that enhances attention to sensory information within a specific region of perception space, making information processing faster and more effective (Brefczynski & DeYoe, 1999). This utilization of limited resources, particularly in visual perception, requires the scanning of the spotlight to capture information. In this paper, we take this concept further by representing the behaviour strategies proposed as a searchlight metaphor, emphasizing the addition of autonomous motor scanning.

1.3. PROJECT GOAL

Information and Communication Technology (ICT) is driving our society towards widespread data and integration of smart objects. While spaces are closely related to humans, they are frequently only utilized as containers for IoT and ICT. We prefer to think like Cuff's (2013) idea that the Internet, digital communicating, data processing, and sensing increasingly actuate the world around us, and we will exist simultaneously within both abstract and physical space.

This paper presents a concept for the creation of a "Searchlight space" that responds to the sentient space with basic biological characteristics. Within this space, people will be able to find new ways to interact with and ascribe this new character to a biophilic design of biological behaviours rather than to a framework for smart spaces. We propose an internal exploratory behavioral strategy of a sentient space, gradually constructing dynamic cognition within limited resources, and a spatial installation was planned and deployed to highlight the concept and associated technologies.

2. Concept and design

2.1. PROPOSED SYSTEM

2.1.1. System concept

The proposed sentient space system mimics natural exploratory behaviour by seeking changes and responding to stimuli, like animals exploring their surroundings. It actively perceives interior changes in brightness through the "Searchlight", the autonomous navigable and purposeful moving light beam, and the "Reactor", the movable sensors for responding to light beam and data collection, that progressively update the contents of the scan over time.

2.1.2. Composition

The construction of the Searchlight Space consists of two parts: (1) The connection between equipment and software is similar to sensorimotor integration in neurology, which refers to the coupling of the sensory and motor systems to allow animals to use sensory information to make useful motor actions (i.e., recursive procedure). (2) The computing system creates a spatial heatmap of the physical space and uses it as its cognitive map to navigates the Searchlight by generate positions according to the adaptive exploratory strategy. The generated positions thus form a path on the heatmap.

Specifically, a projector inside emits the Searchlight, constantly changing brightness of the environment to stimulate the reactors. The computing system processes the input data to learn about the spatial features, such as locations of the reactors and overall brightness, to recognize the environment. As the cognitive map of the space is gradually refined, the Searchlight's moving path and focus become clear. This communication is crucial for sensorimotor mechanisms, enabling the creation of self-regulating mechanisms for the Searchlight space. Furthermore, the cognitive map is visualized and the projected light pattern is post-processed for better understanding and an immersive user experience (see Figure 1).

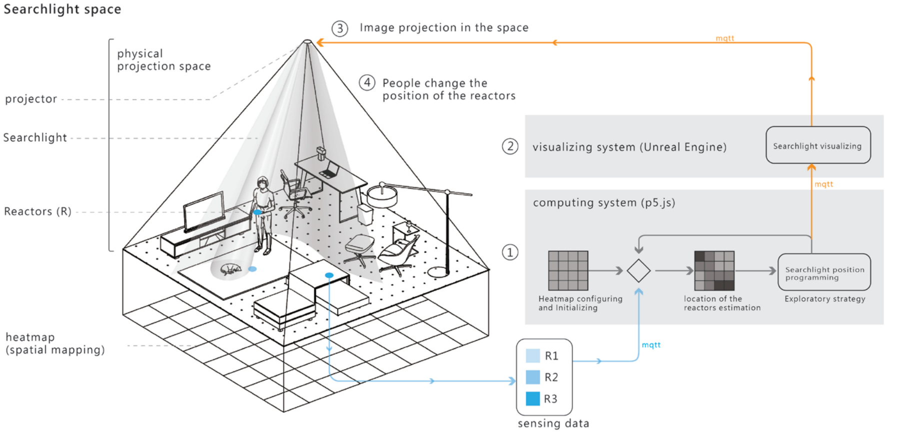


Figure 1. system diagram for the Searchlight space.

2.1.3. Cognitive map to heatmap

A cognitive map, which is a rich internal model, explains the connections between events and anticipates the effects of actions. The core aspect of the Searchlight Space is a cognitive map containing brightness and spatial features within its field of vision as the base to recognize and respond. This map is created from the heatmap, which represents interior space features as colors in a 2D matrix and guides the navigation of the Searchlight. Figure 2 shows the heatmap with the spatial dimension and other elements such as time, the location of the Searchlight, and the reactors to understand variations. Initially, all matrices have the same default weight, meaning that the

Searchlight has no distinguishable features as exploration targets but will become more intricate over time. The heatmap visualization gradually reveals the spatial features and the Searchlight's exploratory tendency.

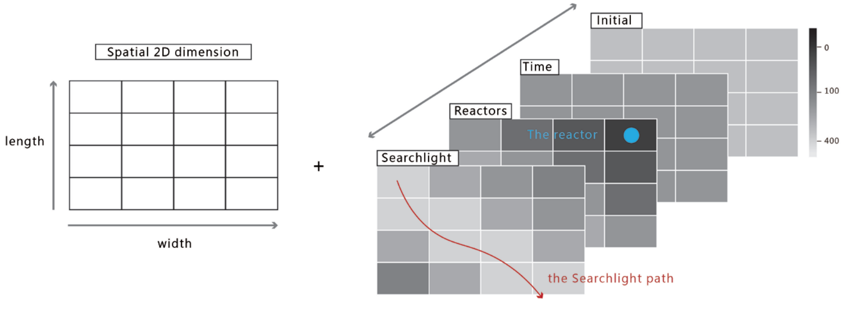


Figure 2. Spatial dimensions and other variations of the heatmap.

2.2. ADAPTIVE EXPLORATORY STRATEGY

2.2.1. Random scanning

The randomness of the Searchlight's navigation has been enhanced to facilitate extensive searches for reactor locations. Initially, a target unit is randomly selected, generating a gravitational force between the Searchlight and the unit. As depicted in the top row of Figure 3, the Searchlight moves towards the target unit, and once it reaches the unit, its weight increases. The lowest weight unit on the heatmap is then set as the new target. Moreover, when the Searchlight receives feedback from a reactor along the way, the system infers the approximate location and movement of the reactor by comparing changes in input data with the path coordinates passed by the Searchlight (bottom row of Figure 3). The location of the reactor significantly affects the weight changes of surrounding grids. With this random scanning strategy, the Searchlight can gradually locate and depict reactors on the heatmap.

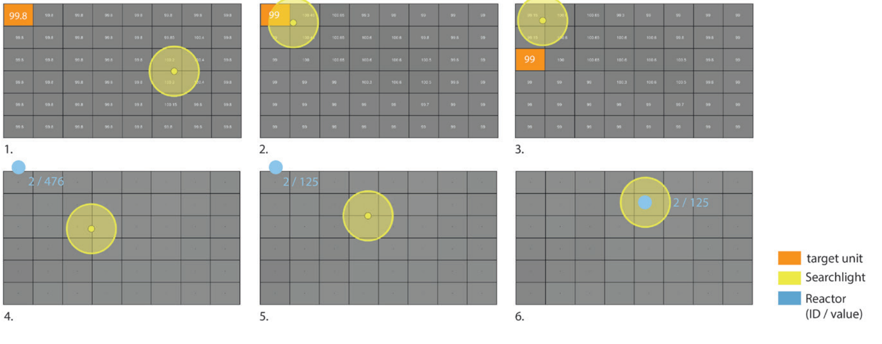


Figure 3. Random scanning: (top) Initial force for Searchlight, (bottom) System infers a corresponding physical location of a Reactor on the heatmap by variation of sensing data per time unit.

2.2.2. Focus area

Attention can gradually get information regarding target, while everything else is sort of becoming a fuzzy or rough. We have designed two focus modes: Fill and Negative Fill, allowing the space to use different modes to recognize the state of the environment. In the Fill mode, the Searchlight prefers to move towards dark areas, as shown in Figure 4 (top row). Because the reactor is brighter than the average, the Searchlight will avoid it and move toward other areas. As a result, the whole projected area gets the global brightness balanced. On the contrary, in the Negative Fill mode, the Searchlight prefers the light areas, see Figure 4 (bottom row), which will further enhance the lightened areas on the heatmap. The initial preference of Fill or Negative Fill would determine the overall reactions toward the reactors.

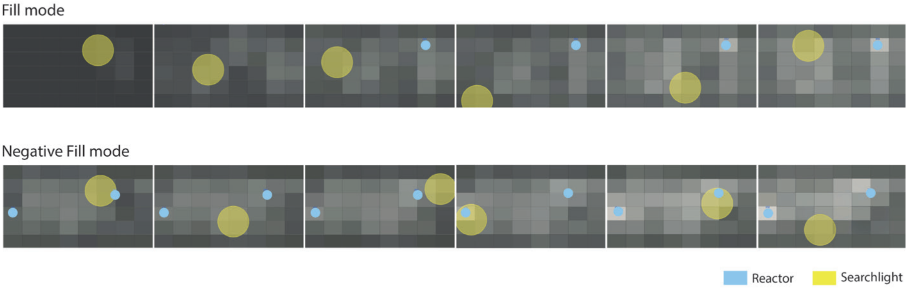


Figure 4. Focus area: Fill or Negative Fill mode results in different lighting patterns.

2.2.3. Overall coverage

While staying focused, the Searchlight occasionally roams to other areas to ensure no missing reactors, in case of been moved or newly added. This is similar to switching the sense from narrow and selective to broad and receptive. The switch is based on an increase in variance as a warning sign. Figure 5 shows that when the variance in the heatmap color increases, t3he Searchlight's navigation becomes repetitive and simplified. The system then switches automatically to the overall coverage strategy, changing the next navigation target from weight-based to history-based, to guide the Searchlight to less visited areas. After a period of time, the system switches back to weight-based focus mode. This self-regulated process ensures an effective and well-rounded perception of the space.

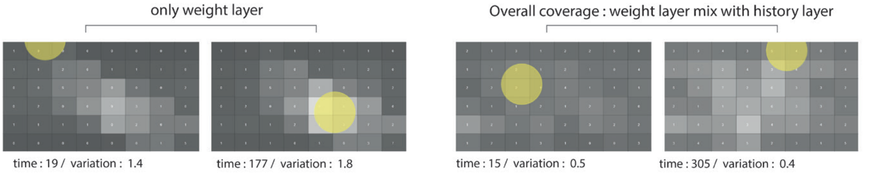


Figure 5. Overall coverage: 2 types of layers for the next target of the Searchlight.

3. Prototype and discussions

3.1. PROTOTYPE

3.1.1. Communication

MQTT protocol is designed as an extremely lightweight publish/subscribe messaging transport that is ideal for connecting remote devices with a small code footprint and minimal network bandwidth. In this prototype, we employ the MQTT protocol for communication among the projector, the reactors, the computing system, and the visualizing system. The reactor is controlled by an ESP8266 (wemos) and it transmits the values of the photoresistor in MQTT protocol to the cloud server over the Wi-Fi network.

3.1.2. From raw data to visualization interface

The prototype uses p5.js to configure and initialize the heatmap and generate the position of the Searchlight. The process is visualized in real-time in the web browser, allowing for efficient analysis and modification of the system.

For the next stage, the generated position data is sent to the visualizing system Unreal Engine (UE), a real-time 3D creation tool for visuals and immersive experiences. We use the Blueprint Visual Scripting system to use the input data to generate the navigation path for the animated light pattern. The pattern is created using the Niagara system for particle-based simulation. The integration of these tools allows us to efficiently establish the full workflow from raw data to visualization, creating an immersive interface for those interacting with the Searchlight Space (see Figure 6).

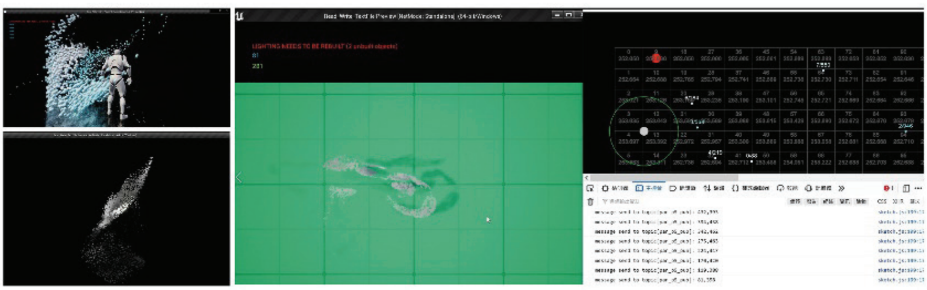


Figure 6. Visualization in UE. (Left) test the particle effect with Niagara system, (Right) receive and send data with MQTT, thus simulating Searchlight movement in UE and heatmap simultaneously.

3.2. TESTS AND DISCUSSION

3.2.1. Test for the prototype

This research planned a set of application scenarios for the prototype to test its feasibility and user experience. The preliminary work performed on-site includes the enclosure design of the reactors as shown in Figure 7 (Left), setting up a local database to record streaming data with the timestamp, and a camera records the interactions between participants and the sentient space. We invited five participants and ask them to conduct the following tests:

- Firstly, place the reactors randomly in the space to see the heatmap showing the relative positions of the reactors in the physical space.
- Secondly, through the observation of the movements of the Searchlight (Figure 8), trying to differentiate whether the Searchlight is in Fill or Negative Fill modes.
- Lastly, change the physical location of the reactors individually or in groups, to see how the system react to the change.

Figure 7. (Left) Enclosure design of Reactors and its interior, (Middle) In site scenario, (Right) A participant with portable Reactors is interacting with the Searchlight.

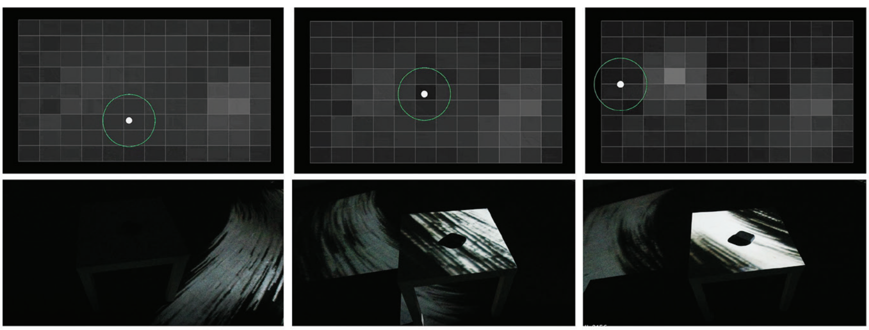


Figure 8. The Searchlight in the space gradually concentrative and purposeful (from left to right), and the heatmap is visualized simultaneously.

3.2.2. Results and Discussion Preliminary Feedback

The test results show that participants were able to quickly complete the goals in the early stage of test. The Searchlight provided a visually engaging experience that attracted participants' attention. The feedback and expectations from the participants include: (1) It is crucial to improve the clarity and comprehensibility of the intended actions and characteristics of the space for the installation to be implemented in an open environment. (2) The patterns of the searchlight can be arranged in various ways in UE, instead of just a single visual effect, since they are regarded as the language between people and space. (3) In addition to light, other scenarios that involve a different form of mediums could be explored. Overall, the Searchlight space has been tested and the following points have been outlined:

- The system creates a sensorimotor-like mechanism, where the reactors perceive constant changes from the physical environment initially, and the Searchlight is eventually projected back to the environment in response. This allows the space to adapt to a dynamic environment.
- To endow the space with basic biological cognitive ability, we compare the heatmap to its cognitive map and deploy a 2D matrix based on the spatial plane. The computing system converts the environmental state marked by the reactors into features on the map. The heatmap also considers factors such as time, past exploration paths, and focus contents to calculate the space's exploration goal based on preferences.
- Visualizing the exploratory behavior of the space through post-processed projected images – which showcase its intention of engaging with people, just like the magic house in Encanto.

4. Conclusion and Future Work

Nowadays, buildings are evolving into smart systems to promote urban development. This study explores the idea of making space a cognitive entity. On one hand, based on the biophilia theory, humans have a natural inclination to connect with nature or other biotic forms for both survival and personal fulfillment. If spaces were to possess sentient qualities, they could pique the interest and thoughts of occupants, strengthening their emotional connection to the space. On the other hand, "Cyburg" refers to a space filled with Information and Communication Technology (ICT) and acts as a cognitive actor, capable of processing information. Instead of actively asking the space to perform tasks like we do with computers, we need the space to have basic exploratory abilities, such as perception and behavior, to gradually reveal its conscious needs.

We propose a sentient space prototype designed to be an environment supported by autonomous algorithms capable of capturing, storing, analyzing, and acting on various data. Our contributions include a workflow for the prototype and experiments evaluating its feasibility and human-computer interaction. This prototype can be applied in various scenes such as art galleries, experimental theaters, or transition spaces like bars - public yet intimate at the same time. By reconfiguring it under different conditions, many completely new interactive experiences and environments would be created.

What are new thoughts on space, from the scale of a single room to an entire city? This research also expected to inspire people's imagination about a new relationship with space, and recognize the complex interplay among a diversity of cognitive actors (Lynch & del Casino, 2019). As human and AI learning becomes more advanced, a space could become a source of knowledge, information, and data for creating a blueprint for our future cities.

Acknowledgements

This research is partially supported by National Science and Technology Council, Taiwan, under Grant number NSTC 111-2420-H-A49-005.

References

Brefczynski, J. A., & DeYoe, E. A. (1999). A physiological correlate of the "spotlight" of visual attention. Nature Neuroscience, 2(4), 370–374. https://doi.org/10.1038/7280

Corbetta, D., DiMercurio, A., Wiener, R. F., Connell, J. P., & Clark, M. (2018). How Perception and Action Fosters Exploration and Selection in Infant Skill Acquisition. Advances in Child Development and Behavior, 55, 1–29. https://doi.org/10.1016/BS.ACDB.2018.04.001

Crang, M., & Graham, S. (2007). SENTIENT CITIES Ambient intelligence and the politics of urban space.10(6), 789–817. https://doi.org/10.1080/13691180701750991

Cuff, D. (2013). Immanent Domain, 57(1), 43–49. https://doi.org/10.1162/104648803322336575

Lee, J. H., Ostwald, M. J., & Kim, A. M. J. (2021). Characterizing Smart Environments as Interactive and Collective Platforms: A Review of the Key Behaviors of Responsive Architecture. Sensors 2021, Vol. 21, Page 3417, 21(10), 3417. https://doi.org/10.3390/S21103417

Lynch, C. R., & del Casino, V. J. (2019). Smart Spaces, Information Processing, and the Question of Intelligence, 110(2), 382–390. https://doi.org/10.1080/24694452.2019.1617103

Red'ko, V. G., Nepomnyashchikh, V. A., & Osipova, E. A. (2015). Models of fish exploratory behavior in mazes. Biologically Inspired Cognitive Architectures, 13, 9–16. https://doi.org/10.1016/J.BICA.2015.05.002

Schnädelbach, H., Slovák, P., Fitzpatrick, G., & Jäger, N. (2016). The immersive effect of adaptive architecture. Pervasive and Mobile Computing, 25, 143–152. https://doi.org/10.1016/J.PMCJ.2014.11.006

Thrift, N. (2014). The 'sentient' city and what it may portend, 1(1). https://doi.org/10.1177/2053951714532241

Zheltova, O. M., & Nepomnyashchikh, V. A. (2020). Organization of Exploratory Behavior in Danio rerio (Hamilton 1822, Cyprinidae) in a Maze. Biology Bulletin 2019 46:9, 46(9), 1059–1064. https://doi.org/10.1134/S1062359019090206

REMOSHARP

Exploring Solutions for Remote Real-Time Collaboration in Computational Design

ARASTOO KHAJEHEE[1], YASUSHI IKEDA[2] and JOSE LUIS GARCÍA DEL CASTILLO Y LÓPEZ[3]
[1] *Keio University, Japan; The University of Tokyo, Japan (current).*
[2] *The University of Tokyo, Japan.*
[3] *Harvard University Graduate School of Design, USA.*
[1] *a.khajehee@outlook.com, 0000-0001-5381-7784*
[2] *yasushi@arch1.t.u-tokyo.ac.jp, 0000-0002-2016-4083*
[3] *jgarciadelcasti@gsd.harvard.edu, 0000-0001-6117-1602*

Abstract. Parametric Design and Fabrication are highly collaborative fields. In most projects, many people are involved with different backgrounds, skill levels, and scripting methods. The collaboration platforms currently available for these fields are primarily cross-software frameworks. They seem to miss out on the nuances of parametric design thinking, namely, focusing on the "how" rather than the "what" in creative projects. In addition, this research identifies "lack of real-time communication" as the most problematic issue with collaborative projects in Parametric Design and Digital Fabrication. This paper introduces "RemoSharp" as a platform for remote interactive collaboration in parametric modeling. This research provides qualitative and quantitative case studies that showcase the capabilities of RemoSharp as a tool for remote interactive collaboration. We demonstrate how users with different skill sets can interactively participate in projects, providing evidence ofhow RemoSharp may encourage modes of communication that improve group work in parametric design workflows. Altogether, this research showcases how RemoSharp provides a generalizable solution for remote collaboration in computational design and digital fabrication.

Keywords. Remote Collaboration, Human-Human Interaction, Real-Time Communication, Computational Design, Communication Tools

1. Introduction

Today, the AEC industry has gotten increasingly digitized and data-driven. Almost all major projects in the industry are collaborations involving the exchange of digital data in high volume and numerous forms. Digital tools are essential in creating the necessary means of communication between people in this field. At their core, digital tools are developed to facilitate the flow of information among people. These tools

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 443-452. © 2023 and published by the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA), Hong Kong.

have enabled us to exchange information and collaborate with much better efficiency and fewer chances of misunderstanding and error.

In such a data-driven environment, the speed of the flow of information is a primary concern, with great room for exploration. Synchronicity, or the lack of it, may highly affect how we approach collaborative projects. With synchronous Real-Time (RT) data flow, people involved in a project will immediately see their colleague's intentions and work. In non-synchronous communication systems (non-RT), there would be a time delay between a message from one person and the received response from the other. A simple example of RT communication is people talking in person; email is a well-known non-RT. Both RT and non-RT have their pros and cons. This paper focuses on exploring the interactivity of RT communication in the computational and parametric design of Architecture.

Many studies have explored how RT fabrication may transform digital fabrication into an interactive design study process. RT alludes to an interactive kind of communication that involves a simultaneous flow of data and information between multiple agents. In other words, RT systems include people or devices with instant feedback or negligible time difference among them. In contrast with existing approaches, interactive fabrication allows RT input to digital fabrication tools, giving it an interactive aspect (Willis et al., 2010).

A robot adjusting for the natural deformation of clay after 3D printing is a kind of RT interaction between a robotic fabrication agent, a human designer, and computer vision (García del Castillo y López, 2019). RT design to fabrication Computer Numerical Control CNC milling is another example of RT interaction between humans and machines (Khajehee, 2022). Sukegawa et al. showcased how a human and a robot equipped with computer vision technologies can gradually arrive at a mutual understanding of the characteristics of a timber block (2022). We can note different RT human-machine interaction examples like "MIMUS: The Curious Industrial Robot" that interacts with spectators in a playful manner (Gannon et al., 2016).

Currently, most RT interaction studies focus on human-machine interaction, in many cases with fabrication in mind. This paper tries to find solutions to bring this RT interactivity to the design stage and between human designers. In this research, RemoSharp was developed as a tool for RT collaboration in parametric design. It is a plugin for Grasshopper3D (GH). What makes RemoSharp special is its compatibility with most GH plugins. It can potentially be used to share GH definitions developed for design, fabrication, computer vision, etc.

2. Background

2.1. EXISTING PLATFORMS FOR COLLABORATION

Speckle is an open-source digital infrastructure for exchanging data, including 3D models, between popular architecture, engineering, and construction software packages. Speckle creates data streams that sync data from a source application like Rhino to its internet-based servers. Other applications, such as Revit, can read the data stream from those servers and recreate the geometry in their native environment (Speckle, 2021). The main focus of Speckle seems to be on syncing geometrical data.

While it has excellent potential for teams to share their work at a higher pace, its focus on streaming geometry can make it less intuitive for collaborative design.

Grasshopper Live is a project developed as a proof of concept for synchronized GH documents for collaborative work. The project is inspired by web applications like Google Docs that allow multiple users from different computers to change a single document in RT. Like Google Docs, both users can change the document, and their changes will be synced for the other users. The functionality is relatively limited, supporting only the creation of standard components, moving, and deleting them. At the time of publishing this paper, the latest version of this project's repository is unfortunately dysfunctional; the repository's files cannot be built into a program solution (Portelli et al., 2019). While this project was not the source of inspiration at the beginning of this research, it became a driving force for further development.

Miro is a popular online collaborative whiteboard platform. It is an infinite whiteboard-like canvas for notes, pictures, and other visual or text-based items. It is a platform for work organization, idea and brainstorming, and planning. The platform's user interface is node-based, in which users create objects and connect them with wires to indicate relations. Miro instantly synchronizes changes for all users in RT. While Miro is not a CAD tool, its user interface resembles parametric visual programming software packages like GH. Miro's interactivity and RT communication level inspired this research (The Visual Collaboration Platform for Every Team | Miro, n.d.).

2.2. SERIALIZATION OF GEOMETRICAL DATA

A common way to share geometrical data between software programs is to "serialize" or encode the information into text and send it. On the receiving side, this text is "deserialized" or decoded to the original geometrical shape. Because of the complex nature of geometrical data, this text can become a considerably long string of text. Speckle uses the serialization method to share data between peers (Serializing & Deserializing Data, 2022). While serialization of geometry can sync data across software platforms, it can become heavy and slow. As a result, the RT aspect depends on the type of geometry that is serialized, sent, received, and finally deserialized. The more complex the geometry, the heavier the final data will get (Figure 1).

Figure 1: Geometry text representation: the length of the serialized text for the more complex geometry is six times longer (heavier) than that of the simple geometry

3. RemoSharp: a tool for remote real-time collaboration

RemoSharp is a solution for interactive collaboration in computational design in GH's visual programming environment. It is a plugin software that enables other GH plugins to be used collaboratively. The name "RemoSharp" is an abbreviated combination of the words Remote (Remo) and CSharp Programming (Sharp). Remote C# scripting was the first idea that gave birth to RemoSharp. The main idea behind the system was to send instructions on "how geometries are made" instead of the geometries themselves. This approach was inspired by the concept behind parametric modeling: "defining the process that creates the form, not the form itself." Likewise, RemoSharp sends short instructions about geometry generation functions. The advantage of this approach is that the amount of data sent and received does not change based on the size and complexity of a model. Thus, the latency of the collaboration setting stays constant throughout the modeling and design process. Unlike geometry serialization, C# scripts are very lightweight. Highly complicated shapes can be scripted in less than a few thousand text characters. As a trial, the first iteration of RemoSharp could run C# scripts on other computers (Figure 2).

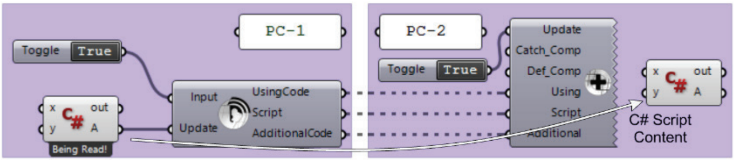


Figure 2: (left-gray) Eavesdrop extracts the code from a C# component. (right-gray) creates a C# component on another computer's GH canvas running the incoming script

RemoSharp is meant to be used in GH, a node-based visual programming language. While GH allows advanced users to script parametric geometry using C#, the main environment does not require any coding abilities. After the remote C# scripting implementation, RemoSharp's functionality was extended to allow coding and regular visual programming. This change enabled non-programmers to use it for remote interactive collaboration in GH (Khajehee, 2021). The current version uses automatically-detected command-based interactions between documents. This approach keeps the content of connected GH definitions the same while allowing simple input values, like number sliders and toggles, to differ or synced. (Figure 3).

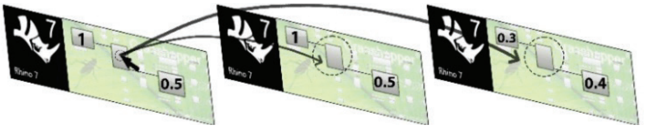


Figure 3: Synced GH documents: RemoSharp connects multiple GH documents to create a RT collaboration pipeline: while the definition is the same, the input data may be the same or different

RemoSharp is built on WebSocket network communication protocol. It can have different network scales: from two computers in a room, multiple computers on a university campus, or multiple computers across the internet. It supports most GH interactions, such as component creation/deletion, wire connection, move, lock/unlock, and hide/unhide. Syncing geometry is also implemented as a feature, not the main focus. Figure 4 illustrates an example RemoSharp command.

```
// command type        component Type Full String              X      Y
> RemoCreate, Grasshopper.Kernel.Parameters.Param_Point, 1000, 738
```

Figure 4: Command to create a "Construct Point" component at the X:1000, Y:738 Coordinates

4. Case Studies

4.1. INTERNATIONAL REMOTE C# SCRIPTING

The following experiment was conducted to demonstrate the remote capabilitites of RemoSharp and how advanced users can use it for collaborative remote C# scripting. RemoSharp's ability to execute C# scripts from text input was showcased in this case study. The demonstration was created as a YouTube live stream on the ParametricCamp channel, showcasing RemoSharp as an example of advanced development in GH (ParametricCamp, 2022). The participants of the case study were the authors of this paper, connected from the United States and Japan. The script was written in a way to allow inputs from both sides. One person could simultaneously change the number of points as the other person was moving them around (Figure 5).

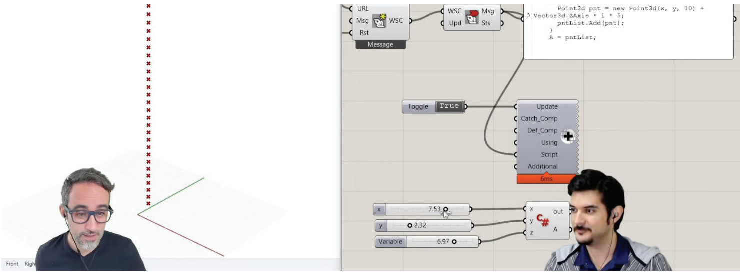


Figure 5: Remote collaborative C# scripting: Connected from Japan and USA, the users create the array of points together in an international RT collaborative computational design session

4.2. COLLABORATIVE PARAMETRIC MODELLING

This case study shows how non-programmers can use RemoSharp for interactive collaboration in GH. RemoSharp is used to allow all participants to communicate and work on the project simultaneously using their own computers. This case study combines RemoSharp's network-building features: RT remote connection and geometry streaming. The networking system involved two sets of one-to-one connections and a geometry streaming pipeline between the two sets. The geometrical results of each one-to-one set were shared via the geometry stream to achieve the final result. This study involved 4 participants with different skill levels, three advanced GH users and one with moderate skills. The participants were asked to model the "True Talker Pavilion" in Amsterdam (StoneCycling, 2016) (Figure 6).

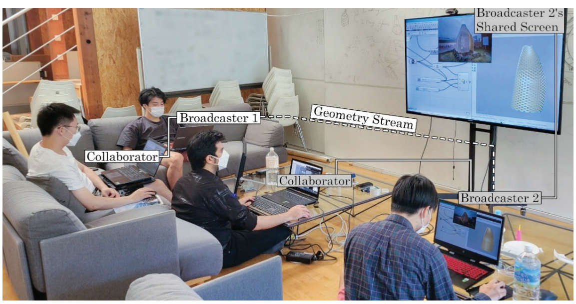


Figure 6: Two sets of one-to-one collaboration systems combined using a Geometry Stream

After setting up the network and establishing the interactive collaboration platform, the participants started discussing how to model the building together. Since there were two one-to-one networks available, the participants decided to divide the task into two parts. One group worked on the surface building, and the other group worked on the bricklaying layout (Figure 7).

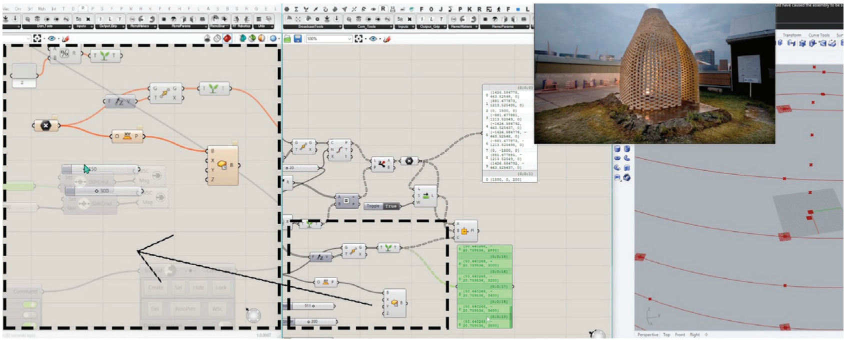


Figure 7: Screenshot of one of the one-to-one setups: (left) first user's connected GH canvas, (middle) second user's connected GH canvas, (right) second user's Rhino Viewport. Semi-transparent components belong to previous versions of RemoSharp (currently deprecated)

The following are a few of the questions and suggestions that came up during the first collaborative modeling stage:

"How do you want the bricks to be oriented? Plane-based or vector-based?"

"Will you use orient to transform the brick onto the curves?"

"How are you dividing the curves?"

"How should we make our file consistent with the other group's work?"

"Do you want to use circles and loft them? Or to use the pipe tool?"

At the final stage of the collaboration, one group asked the following question:

"What type of geometry do you want me to send? Surface or curves?"

This case study combined different types of collaboration network configurations. Users divided their tasks based on how they wanted to configure their collaboration networking systems. Interestingly, even though the tasks were divided, the participants were aware of the other group's work and processes. The awareness was interactive, adjusting to the flow and thinking process of the other group. The degree of engagement between the participants varied according to how co-dependent their work was at a given time. For example, the two members of the first group had more communication between themselves compared to their communication with the members of the other group. This finding shows that having a RT parametric modeling-oriented interactive system encourages more idea exchange and interaction than just streaming the final geometry. The details of the questions were also in line with computational design thinking. The participants discussed which GH components to use and how to create the GH definition.

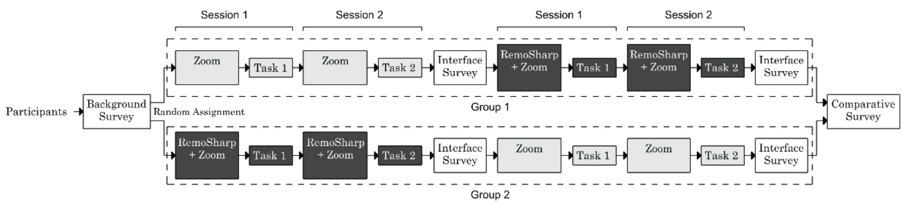
4.3. CONTROLLED USER STUDY: REMOSHARP FOR EDUCATION

This controlled user case study focuses on how people's interaction with GH changes when RemoSharp is used. The hypothesis was that RemoSharp enables the instructor to be more engaged with the students' GH canvas, thus making him/her more aware of his/her teaching. At the same time, it enables the students to interact with the instructor's teachings in a more hands-on manner. As hands-on interaction with a software package is essential for learning, this research also argues that RemoSharp can enhance the learning experience of GH in an educational setting as well.

Can RemoSharp make the learning experience more enjoyable and intuitive? The case study was designed to provide qualitative and quantitative evidence to answer these questions. The goal was to teach GH to students with minimal or no exposure to the software and compare their impressions with conventional teaching methods.

It is worth noting the reason behind the small number of participants involved in this study. Unfortunately, conducting the test became possible after COVID-19 restrictions were partially eased up. As a result, increasing the number of participants in the study became quite challenging, especially given Japan's restrictive health measures.

4.3.1. METHODOLOGY AND USER GROUP



The proposed study was conducted according to the cross-over model. In a cross-over study, participants are exposed to different environments sequentially. The results are compared between groups while the sequence is randomly changed. The primary affordance of this experiment model is to compare the effect of all the environmental variables on the same subject, at the same time as identifying and isolating carryover or residual effects of the previous exposure to the next (Jones & Kenward, 2014). The test was conducted as a "Remote Online Grasshopper workshop series," including two rounds of 2 sessions over a month. For the first round, the first group was taught only via Zoom screen sharing, while the second group was taught using RemoSharp and Zoom. The same material was taught for the second round, but the teaching method was switched; the first group was taught using RemoSharp and Zoom, and the second was taught using only Zoom. (Figure 8).

Figure 8: Diagram of the cross-over structure of the controlled user study.

Each session included GH instructions followed by a hands-on task for the students. All the steps required for finishing the exercises were explained independently during the explanation part. During the exercise part, hints were given to the students as a teaching method. It was ensured that the hints were not giving away the answer directly, but they would help the students to find the solution themselves.

The students' interactions with GH were logged with Remosharp's GH logging tool, "GHL." The resulting log includes creating/deleting components, wire connections, moving, and disabling them. In addition, the GHL component detects canvas viewport movement, including zoom in/outs and view change.

Two groups of 4 volunteer students, 8 in total, were created. The group assignment criteria were solely based on schedule considerations. The students were asked to participate in at least four weekly sessions together. It was emphasized that they should attend all four sessions. All students were in their bachelor's degree studies. The study followed Laurel et al. guidelines for user study observation (1990).

4.3.2. Results

Of the 8 participants, all finished their tasks with some help and hint from the instructor. Sessions, on average, took about 1:50 hours, depending on how long the students took to finish the exercises. Figure 9 shows the students' modeling exercises results. Following the second session of each round, the students were asked to evaluate their experience in the workshop. Figure 10 shows the interface survey comparative results.

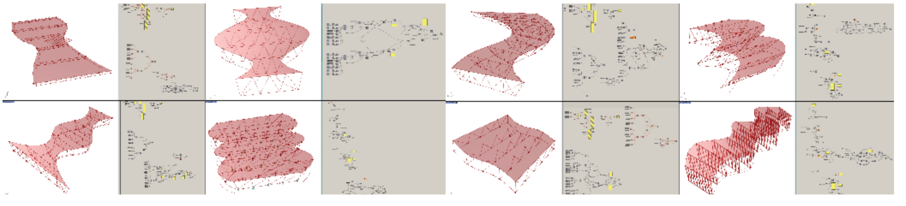


Figure 9: Finished exercises by the students: Students were asked to model a 3D truss from points. With some hints from the instructor, all eight students successfully finished their exercises.

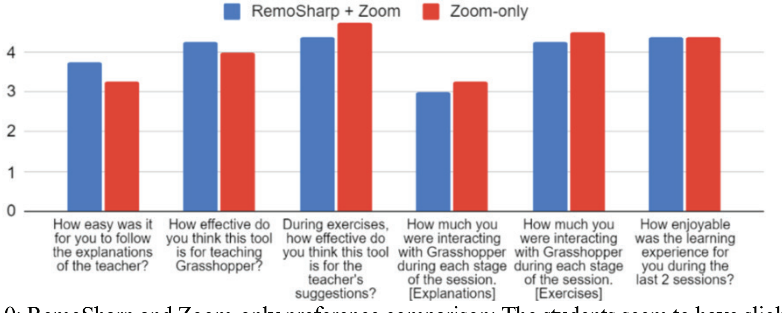
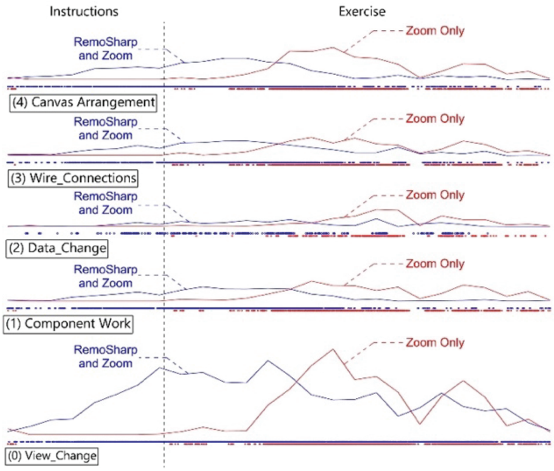


Figure 10: RemoSharp and Zoom-only preference comparison: The students seem to have slightly enjoyed having RemoSharp rather than Zoom-only. However, the difference was not so significant.

he students' interactions with GH were logged during all sessions. The log in this study was recorded as a function of interaction over time. Interactions were recorded from the start of the session until the students finished their exercise. The log's data was divided into five categories: view change, component work, data change, wire connections, and canvas arrangement. The log data shows increased interaction during the explanation stage in all categories. While the overall interaction amount remains the same, when RemoSharp is used, it shifts toward the beginning of the session. It shows that RemoSharp may encourage the students to interact with GH earlier. In Zoom-only sessions, the students seemed to prefer to listen solely. This finding seems consistent with their comments about their inability to keep up with the explanations if they try to build the GH definition during the instructions. With just Zoom, students' interaction with GH seemed to have begun after the explanations (Figure 11).

Figure 11: Student's GH Interaction Log: Interactions are plotted as an action/time function. (blue) RemoSharp and Zoom, (red) Zoom Screen Sharing only.



As a final impression, all eight students preferred to have RemoSharp used in GH workshops when the following question was asked of them:

"If the Grasshopper workshops continue, which system would you prefer?"

5. Discussion and Conclusion

One of the fundamental ideas behind the parametric design paradigm is understanding the processes that shape a design. The collaboration platforms currently available for parametric modeling are mostly general-purpose frameworks that try to connect multiple software packages. This approach leads to viable solutions. Nevertheless, they miss out on some of the nuances and ideas behind parametric design thinking. This research developed "RemoSharp" as a solution for interactive collaborations that retains the underlying principles of parametric design and modeling. Using RemoSharp, "how models are made" is shared among collaborators. RemoSharp enables users with varying skill levels and specialties to collaborate in GH. It helps users grasp a deeper understanding of the model together. It shares the mechanism in RT for all users. People can share the parametric change and their understanding of the system interactively. It has tools with the potential to incorporate collaborative parametric design with RT fabrication and other related fields. RemoSharp tries to be a generalizable solution that makes parametric design collaborative and interactive.

It is worth noting that RemoSharp is still under development at the time of publishing this paper. During the controlled user study, there were instances of RemoSharp disconnecting or being too heavy for the students' less powerful laptops. The technical problems may have affected the impression of the students. In later version updates, this issue was improved, but there is still room for improvement. The issue of security was not in the scope of this research. Exploring how we can secure RT collaborative computational design environments may be worth exploring. Assignment of roles and level of access can also present grounds for development. RemoSharp can potentially enable most GH plugins to be used collaboratively. This calls for further research to see how different operations can benefit RT collaboration.

References

Gannon, M., Sandoval, J., McPhail, K., & Snell, B. (2016). Mimus, The Curious Industrial Robot. ATONATON. https://atonaton.com/mimus

García del Castillo y López, J. L. (2019). Enactive Robotics: An Action-State Model for Concurrent Machine Control. Harvard Graduate School of Design. Retrieved from: http://nrs.harvard.edu/urn-3:HUL.InstRepos:41021631

Jones, B., & Kenward, M. G. (2014). Design and Analysis of Cross-Over Trials (0 ed.). Chapman and Hall/CRC. https://doi.org/10.1201/b17537

Khajehee, A. (2022). Development of an Affordable On-Site Wood Craft System: Interactive Fabrication via Digital Tools. POST-CARBON - Proceedings of the 27th CAADRIA Conference, Sydney, 9-15 April 2022, Pp. 31-40. https://doi.org/10.52842/conf.caadria.2022.2.031

Khajehee, A. (2021). RemoSharp: Remote Real-Time Collaboration in Rhino and Grasshopper. Food4Rhino. https://www.food4rhino.com/en/app/remosharp

Laurel, Brenda., Mountford, S. Joy (1990). The Art of human-computer interface design. Addison-Wesley Pub. Co., Inc., USA.

Portelli, C., Ertugrul, E., Poulsen, E., & Poinet, P. (2019). Grasshopper Live [C#]. CORE studio. https://github.com/tt-acm/GrasshopperLive (Original work published 2019)

ParametricCamp (Director). (2022, February 2). 4.12 Remote Scripting via RemoSharp - Advanced Development in Grasshopper. https://www.youtube.com/watch?v=RsuDryv8slM

Serializing & Deserializing Data. (2022, August 11). https://speckle.guide/dev/dotnet.html#serializing-deserializing-data

Speckle (Director). (2021, April 28). Speckle in One Minute. Retrieved from: https://www.youtube.com/watch?v=B9humiSpHzM

StoneCycling. (2016). Storytelling at TrueTalker Pavilion in Amsterdam |. StoneCycling. Retrieved from: https://www.stonecycling.com/projects/storytelling-at-truetalker-pavilion-amsterdam/

Sukegawa, C. (2022). Smart Hand for Digital Twin Timber Work -The interactive procedural scanning by industrial arm robot. Jeroen van Ameijde, Nicole Gardner, Kyung Hoon Hyun, Dan Luo, Urvi Sheth (Eds.), POST-CARBON - Proceedings of the 27th CAADRIA Conference, Sydney, 9-15 April 2022, Pp. 131-140. https://doi.org/10.52842/conf.caadria.2022.2.131

The Visual Collaboration Platform for Every Team | Miro. (n.d.). Https://Miro.Com/. Retrieved from: https://miro.com/

Willis, K. D. D., Xu, C., Wu, K.-J., Levin, G., & Gross, M. D. (2010). Interactive fabrication: New interfaces for digital fabrication. Proceedings of the Fifth International Conference on Tangible, Embedded, and Embodied Interaction, 69–72. https://doi.org/10.1145/1935701.1935716

IMPLEMENTATION OF A HUMAN-CENTRIC INTERACTIVE SMART SPACE USING EEG AND FACE EMOTION AI

SEUNG YEUL JI[1], Mi KYOUNG KIM[2], JIN CHEOL BAE[3], SE YEON KANG[4] and HAN JONG JUN[5]

[1,2,3,4,5]*Hanyang University, Korea.*
[1]*musicji83@hanyang.ac.kr, 0000-0003-1268-0384*
[2]*nnyang2mk@hanyang.ac.kr, 0000-0002-6692-405X*
[3]*zbjc95@hanyang.ac.kr, 0000-0001-6392-1550*
[4]*seyeon95@hanyang.ac.kr, 0000-0002-0780-5042*
[5]*hanjong@hanyang.ac.kr, 0000-0003-1798-183X*

Abstract. Various approaches have been proposed to create a smart space using a variety of sensors and instil positive emotions in a user-centred space. The core feature of building human-centric smart space is the analysis of user space using various sensors. In particular, as biometric sensors used in medical applications involving humans are simple and fused, research to use them in everyday environments is being actively conducted. In this study, digital twins were used to monitor humans in the building space using EEG data, analyse their situational characteristics, and create a smart environment. However, to compensate for the limitations associated with accurate expression of emotional information using only brain waves, the emotional states users were supplemented through the image-based Face Emotion AI, and a smart space was established to improve the emotional states of users.

Keywords. Human-centric, Smart Space, Face Emotion AI, Face Emotion AI using Vision, EEG.

1. Introduction

Smart systems are being introduced into existing architectural spaces in various ways to create a positive space using sensors. These systems can positively impact social issues those afflicted by the rapidly changing climate, aging-oriented social structure, or architectural welfare. The core technology underlying building human-centric smart spaces is the analysis of space for users using various sensors. Sensors for analysing the architectural spaces inhabited by humans are divided into the fixed method that is attached to the space and the moving method that is linked to humans, such as in a smartphone. In particular, as biometric sensors that are primarily used in the medical field have been simplified and fused, research focusing on their use in everyday environments has been actively conducted, and the moving method has been fused to ensure that experiments can be conducted in various environments. Among biometric

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 453-462.

sensors, EEG sensors used to monitor the brain have made significant progress; however, their use in other fields remains limited. In this study, we propose an interactive space implementation method based on the relationship between architectural spaces and humans using deep-learning-based recommendation algorithms.

In this study, the primary task of the human-centred smart space is to collect and analyse brain wave data and analyse the human mind, such that it is reflected in the actual space through patterns of specific emotions among more than 60 data points per second and induces positive emotions. Data analysis techniques have been used in previous studies to read people's minds in space using interviews, video recording, and questionnaires. Because such information can be subjective and arbitrary, depending on the processing of cognitive information, the results may be distorted. To compensate for such errors, the response of biometric information is collected in its raw form without distortion, and patterns are organised to utilise the flow of emotional information. Even in the case of EEG sensors, the miniaturisation of sensors and the resolution of EEG collection have been rapidly developing; moreover, with the advancements in software tools for collecting and processing EEG data, an automation function for determining wear errors or nonconformity of the collection environment has been developed. The applicability of this tool is expanding from controlled measurement environments in medical applications to a general living environment.

Therefore, this study aims to understand and utilise the psychological information of users by collecting EEG data in architectural spaces and using it as a data indicator. In contrast to previous studies, this study supplements the task of promoting psychological stability of users by interactively controlling the space based only on EEG information. Additionally, the superficial image information of users is employed to synthesise the emotional information obtained via EEG data and use it for smart space operation.

1.1. METHODOLOGY

This study focuses on driving a linear actuator motor for the convergence mechanism of brain waves and image-based Face Emotion AI as well as the interactivity of space to develop a smart space that induces a stable psychological state. For the safety of users, a sensor is used as the physical engine underlying the game engine to create a change model for the space size control system using deep learning algorithms and control the space using emotional information obtained via EEG indicators and facial expressions. Moreover, this study establishes a personalised smart space implementation model in which two pieces of emotional information data are fused based on the user's EEG and Face Emotion AI considering that the machine is affected by the human mind. The configuration of the smart space system limited the scope of the study to space optimisation based on space expansion and reduction by indicators in the EEG sensor data corresponding to stress and concentration of the users. (Figure 1).

Figure 1. Smart space operation using EEG

In this study, an interactive space control model of a smart space was established to create a psychologically optimal space for users, and an interactive motor control process driven by the fusion of brain waves and Face Emotion AI was developed to perform optimisation analysis of the data reflecting the emotional information of users. The core features of the system that reflects EEG and Face Emotion AI are smart space control that considers emotional information, and the interactive system for the smart space that continuously improves the analysis of the stress index loss before and after the change of space size using the DeepFM algorithm. Moreover, the trend of emotions is analysed based on the feedback of EEG data and Face Emotion AI for optimising the mechanical motor control.

2. EEG and Face Emotion AI Convergence Mechanisms

The essential feature of building a smart space driven by the convergence of EEG and Face Emotion AI is a virtual space model, and the primary purpose is to derive the variables of the optimal space via simulation considering the emotional state using a game engine. To this end, the interactive space element is visualised and digitised via a simulation of the pseudo model using the game engine. Users' EEG data and Face Emotion AI information in the physical space is controlled through the data feedback process using the virtual model in the game engine. Specifically, through the integration of AI-based deep learning and the biometric information of users, it is possible to intelligently diagnose the operation and control of the interactive space system and predict the psychological stability of users. Ultimately, an interactive smart space can be established by utilising the physical-level space transformation system and data feedback from the emotional information of the six brainwave channels and the facial image information. The mechanism that utilises EEG and Face Emotion AI for smart spaces is shown in Figure 2. This mechanism collects environmental information about a space for driving an actual spatial structure and recognising the dynamic state of the space using sensors. In this study, EEG and Face Emotion AI information, which represent the emotional state of the user, were used as environmental variables for the space, and the structure of the space was constructed using a three-dimensional model and a 1/100 scale model built based on the game engine to set up a process that drives the space of the actual building and sets the foundation for AI-based analysis. For the efficient operation of the system in the smart space using AI, intelligent diagnosis of the variable space system and prediction of emotional information before and after spatial transformation is necessary for the management and control of biometric information.

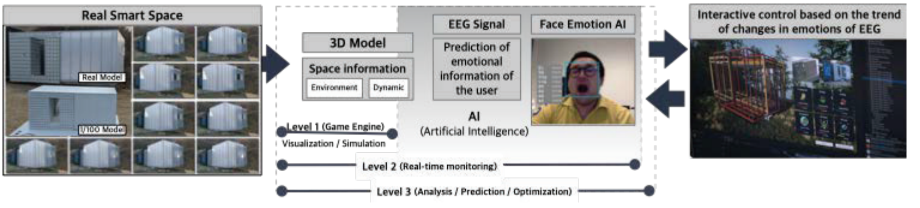


Figure 2. Fusion mechanism for EEG and Face Emotion AI

Brainwave data and video image information on one time series axis are used as insertion values in an intelligent algorithm for realising a smart space. Emotional information is extracted from existing EEG data. Emotional information obtained using Vision AI is labelled with KeyStroke for the time series of EEG data and used as additional supplementary data, and the emotion data obtained using EEG is strengthened. Thus, it plays a key role in data processing as it controls the space with the pre-processed emotional state data while grasping the emotional state of the user before and after the system operation and transformation. In this study, based on the deep learning algorithm, a smart space control process structure was used, considering the relevance of changes in the EEG-based emotional information according to the environmental factors of the space, to build a training set and train the deep learning model. EEG-based biometric data composed of training sets were used, and an analysis algorithm architecture for the changes in the emotional information in EEG according to the space before and after the transformation was constructed. In particular, a simulation of the user's emotional information was performed in real time in a game engine environment using the 3D model and 1/100 mockup model. The actual physical environment conditions were taken into account in the simulation and visualisation (Figure 3).

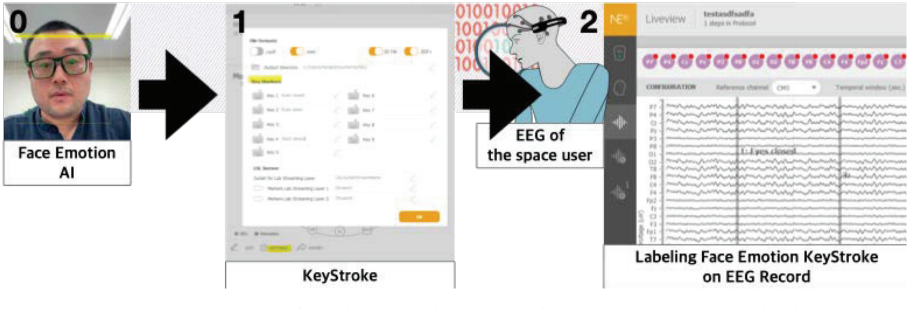


Figure 3. Game-engine-based EEG and Face Emotion AI convergence mechanism

3. EEG and Face Emotion AI for Intelligent Control of Smart Space and Artificial Intelligence Convergence

To effectively operate the smart space motor system using the data from the user's EEG and Face Emotion AI, a DeepFM-based artificial intelligence model was constructed and divided into reality, semi-reality, and virtual reality for the same task. The data were collected through DeepFM by moving the space using a linear-actuator-type motor and monitoring the stability using a sensor that transmitted the state information of the building while moving. The processing of the sensor data for motor driving and stability using the semi-sensor reality and of the motor information using the 1/100 model and the physical engine of the game engine was implicit in the data collection from the sensors, and the safety of the real environment was accounted for. In particular, the virtual reality sensor for which the game engine was used accounted for the collision event using the collider and trigger functions employed to implement a realistic game based on the information collected by the reality and semi-reality sensors. This information was used in connection with DeepFM (Figure 4).

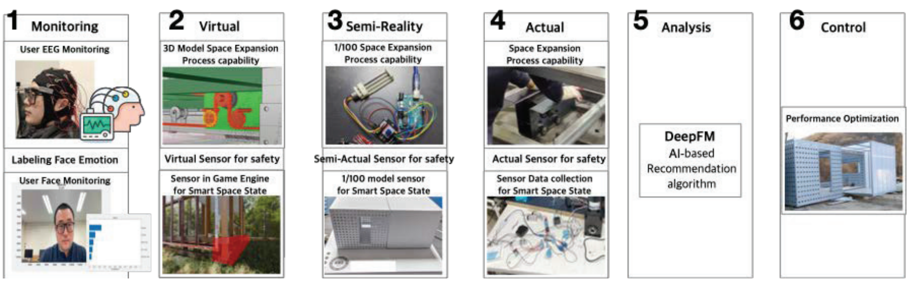


Figure 4. Optimising smart space control with DeepFM

Explicit capability refers to a smart space that moves, and implicit capability is based on a simulation based on EEG to induce the desired operation of the smart space. Implicit capability is carried out based on a game engine and verified through a model. DeepFM collects the state information of the three situations between game engines, 1/100 models, and real models and performs calculations and provides judgment results. In this study, DeepFM recommends the distance by which the space moves according to the size of the space expansion. In particular, it is used as an information resource for emotional information and an artificial intelligence model that converts EEG data depending on the channel and the EEG data from Emotiv into the emotion index, according to the environment conditions of the smart space. Figure 5 presents a data structure for operating the DeepFM algorithm while converging EEG and Face Emotion AI. In DeepFM, emotional information in a smart space is organised by classifying properties with low-order features and high-order features based on time information. Columns based on the number of horizontal arrays collect information such as User, EEG, Face Emotion, physics simulation information of game engines, gyro sensors, time, and EEG-based Emotional Metric information in seconds. The collected data are assigned a motor control numerical value of Target y, which controls the smart space for optimal space expansion through interaction mechanisms and training around the game engine. CTR is predicted by DeepFM. Subsequently, the smart space of the game engine is operated based on the recommended Target y value. DeepFM analyses the user's emotional information through the physical operation of the game engine and remaps variable values, ranging from one to five, within which the space moves, wherein the variables are limited to the Target y value. The Remap function is used for remapping. In this study, the remap function was the NumPy library of Python. The different number patterns accumulated in each sensor were remapped to the specified number range. For example, if the EEG has a distribution of 1 to 100, the range of the number is remapped between 1 and 5 while maintaining the ratio pattern. After applying the data verified through the game engine to the small motor of the 1/100 reduction model, the data are applied to the actual building if there is no abnormality in the action. The next work process is illustrated in Figure 5.

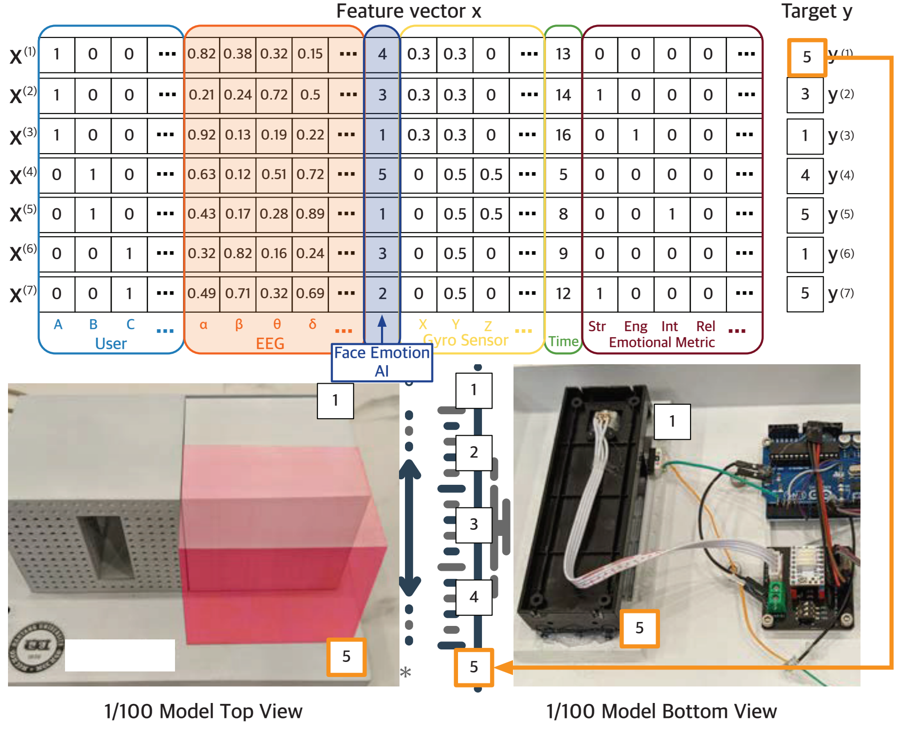


Figure 5. Data structure for smart space operations with the DeepFM algorithm

4. Control of the Size of a Smart Space Driven by the Integration of EEG, Face Emotion AI, and Artificial Intelligence in a Smart Space of a Single User

Spatial control of smart space involves patterning, environmental factors, a combination of technologies, and the user's emotional information. Emotional information is extracted from the image information of facial expressions and reflected in real time to compensate for various uncertain factors generated from the user's brain waves. The linear-actuator-type interactive motor control system is driven by an analog-type DC motor, and it has dynamic and non-linear characteristics. This type of motor provides a high power for the operation of heavy-weight objects. Nevertheless, instability in the smart control of the motor and in the control of biometric information is likely to occur. Since EEG data for this research has been collected since 2012, optimisation of the operation of the smart space for only one person was performed considering that the EEG model for deriving emotional information is highly safe and that a building with a weight of nearly 1 ton was operated. Using the proposed system, an interaction experiment was conducted with one experimenter by utilising emotional information in smart spaces using the EEG sensors and Face Emotion AI.

Smart space recommendation item by CTR prediction	CTR (c)	Space size (m2)	Real space transformation count	Degree of concentration upward derivation number (Recommended)	Recommended contrast concentration upward frequency accuracy	Number of induced stress levels (Recommended)	Recommended stress down frequency accuracy
1	$0 \leq c \leq 0.2$	9	8	8(11)	72.7%	2(11)	18.1%
2	$0.3 \leq c \leq 0.4$	10.8	6	1(9)	11.1%	2(9)	22.2%
3	$0.5 \leq c \leq 0.6$	12.6	3	1(10)	10%	1(10)	10%
4	$0.7 \leq c \leq 0.8$	14.4	2	1(6)	16.6%	1(6)	16.6%
5	$0.9 \leq c \leq 1$	16.2	7	1(10)	10%	4(10)	40%

Table 1. Number and ratio of accuracy according to higher concentration and lower stress induction compared to the recommended number of times in AI

First, as shown in Table 1, the CTR (c) is set as a predetermined value. Its range is set to converge to 0 as the concentration increases and to 1 as the stress decreases. The CTR was predicted using DeepFM. Based on the range of the predicted values, the points where the smart space moved were designated between 1 and 5 and recommended. Ten experiments were conducted for 4 min each, and the user's EEG data according to the space situation were entered into to DeepFM. These data were sequentially applied to the three situations: virtual reality, semi-reality, and reality. There were 25 modifications of the space for a total of 50 min. The accuracy of the recommendation of DeepFM for the spatial deformation, which induced a high concentration or low stress, was 50% of the recommended accuracy of the case in which the concentration increased. The accuracy was the highest when the size of space 1 was recommended, and the accuracy for low stress compared to the recommended number for the case of stress was the highest when the size of space 5 was recommended (Table 1).

5. Conclusion

This study expanded and contracted the space of a building using the user's EEG information (Figure 1). To move a space of approximately 1 ton, a smart space needed to be operated using the user's biometric data-based emotional information. This is because, unlike when receiving and operating the user's electronic electrode signal, the EEG information cannot guarantee safety when applying signal information to the building because of the generation of noise or the range of change in emotion of the residents. In addition, there is a limit to determining the user's emotional information solely based on the EEG data. Therefore, the reliability of deriving emotions from biometric information was supplemented using an algorithm that photographs the user's face and analyses emotions through facial expressions using Face Emotion AI. Two AI algorithms were sequentially used to determine the virtual environment of the game engine created under the same conditions as the actual building and the 1/100 reduction model. The first AI algorithm helped derive emotions from the user's facial information while integrating the EEG data and the biometric data obtained from the Face Emotion AI (Figure 2). The second AI algorithm provided recommendations for inducing a positive emotional state by integrating biometric data with DeepFM, and the recommended target was limited to five motor control positions. An experiment was conducted on the relationship between the user's emotional information and their smart space using the accumulated data one user from 2012 in the user's biometric information. When stress is increased and the engagement of space expansion is decreased, the experiment was conducted using the DeepFM algorithm to recommend reductions in space. The following relationship was applied using the traditional Chombard de Lauwe criteria: To secure the safety of the space where the user is located, the game engine incorporates a safety device through primary simulation, the model is used as a secondary stabiliser, and the methodology of the mechanism in which the final building is operated is applied (Figure 4). A single user lived in the building for 50 min and an experiment was conducted. A total of 25 changes in the space occurred. The number of moves is not significant because it can change according to the setting of the bias weight when training the AI algorithm model. However, to induce a positive psychological state, the process of linking biometric data and motors was significant. In particular, the alternatives to the space recommended by DeepFM were limited to five in this experiment to minimise the motor load. However, experiments such as fine adjustment through various equal parts will be conducted in future studies after mounting a physical safety device switch or sensor on the building drive. The experimental results demonstrate that, after the change in the space based on the user's EEG data, the alternative to the space in which engagement was induced with the highest efficiency was 1, and the alternative to the space induced to reduce stress efficiently was 5. However, such a basic experiment was conducted to determine the feasibility of implementing a linkage work that reflects the psychological state of the user in the building space based on the biometric data. In future studies, we will consider and apply an AI algorithm that is effective for changing analogue and digital data into feedback, such as biometric data or motor control for digital data space operation. An AI model optimised for EEG and smart space should be selected after setting an optimal algorithm through the setting process of an effective DeepFM algorithm through the fine-tuning of the bias or weight of a deep learning algorithm not considered in this study. In addition, further research on the intelligent construction and management of all elements of the entire process should be conducted by improving information collection and communication technology through sensors to analyse spatial conditions, performing real-time analysis of users' emotions for space, and realising in-depth integration of digital twins, AI, and the latest edge technologies. Future research will focus on addressing the limitations of this study and extensively exploring the relationship between architectural space and biological data.

Acknowledgements

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government(MSIT).(No. 2022R1A2C3011796)

References

Bleichner, M.G., Lundbeck, M., Selisky, M., Minow, F., Jäger, M., Emkes, R., Debener, S., & De Vos, M. (2015). Exploring miniaturized EEG electrodes for brain-computer interfaces. An EEG you do not see? Physiological Reports, 3(4), e12362–e12362. https://doi.org/10.14814/phy2.12362

Broom, A. (2005). Using qualitative interviews in CAM research: A guide to study design, data collection and data analysis. Complementary Therapies in Medicine, 13(1), 65–73. https://doi.org/10.1016/j.ctim.2005.01.001

Dong, B., Kjærgaard, M.B., De Simone, M., Gunay, H.B., O'Brien, W., Mora, D., Dziedzic, J. & Zhao, J. (2018). Sensing and data acquisition, in Wagner, A., O'Brien, W., Dong, B. (eds), Exploring Occupant Behavior in Buildings, Springer, Cham, Switzerland (pp. 77–105). https://doi.org/10.1007/978-3-319-61464-9_4

Eberhard, J. (2009). Brain Landscape: The Coexistence Neuroscience and Architecture. Oxford University Press, New York.

El Saddik, A. (2018). Digital twins: The convergence of multimedia technologies. IEEE Multimedia, 25(2), 87–92.

Familiar, M.S., Martinez, J.F., & Lopez, L. (2012). Pervasive smart spaces and environments: A service-oriented middleware architecture for wireless Ad Hoc and sensor networks. International Journal of Distributed Sensor Networks. https://doi.org/10.1155/2012/725190

Filipponi, L., Vitaletti, A., Landi, G., Memeo, V., Laura, G., & Pucci, P. (2010). Smart city: An event driven architecture for monitoring public spaces with heterogeneous sensors. In Proceedings of Fourth International Conference on Computer and Communication Technologies. pp. 281–286. https://doi.org/10.1109/SENSORCOMM.2010.50
Guedes, A., Macedo, R., Jaramillo, G., Cardoso, S., Freitas, P.P., & Horsley, D.A. (2018). Hybrid GMR sensor detecting 950 pT/sqrt (Hz) at 1 Hz and room temperature. Sensors, 18(3), 790. https://doi.org/10.3390/s18030790
Ibrion, M., Paltrinieri, N., & Nejad, A.R. (2019). On risk of digital twin implementation in marine industry: Learning from aviation industry. Journal of Physics: Conference Series, 1357(1). https://doi.org/10.1088/1742-6596/1357/1/012009
Johnson-Laird, P.N. (1983). Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness. Harvard University Press.
Kwon, D.Y., & Gross, M. (2005). Combining body sensors and visual sensors for motion training. in Proceedings of the 2005 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology, Valencia, Spain (pp. 94–101).
Lee, J. (2018), Smart space based on platform using big data for efficient decision-making. Informatization Policy, 25(4), 108–120. https://doi.org/10.22693/NIAIP.2018.25.4.108
Liu, Z., Meyendorf, N. and Mrad, N. (2018). The role of data fusion in predictive maintenance using digital twin. AIP Conference Proceedings, Annual Review of Progress in Quantitative Nondestructive Evaluation.
Ni, Q., García Hernando, A.B.G., & de la Cruz, I.P. (2015). The Elderly's independent living in smart homes: A characterization of activities and sensing infrastructure survey to facilitate services development. Sensors, 15(5), 11312–11362. https://doi.org/10.3390/s150511312
Sheng, J., Amankwah-Amoah, J. and Wang, X. (2018), Technology in the 21st century: new challenges and opportunities. Technological Forecasting and Social Change. https://10.1016/j.techfore.2018.06.009
SternbergRJInside Intelligence (1986).
Tao, F., Zhang, M., & Nee, A.Y.C. (2019). Digital Twin Driven Smart Manufacturing. Academic Press.
Teplan, M. (2002). Fundamentals of EEG measurement. Measurement Science Review, 2, 1–11.
Trnka, A. (2014). Big data analysis. European Journal of Science and Theology, 10(1), 143–148.
Vieira, S., Gero, J.S., Delmoral, J., Li, S., Cascini, G., & Fernandes, A. (2020). Brain activity in constrained and open design spaces: An EEG study. In Boujet, J.-F. (eds), The Sixth International Conference on Design Creativity-ICDC2020, Design Society (pp. 68–75).
Villanueva, F.J., Santofimia, M.J., Villa, D., Barba, J., & Lopez, J.C. (2013). Civitas: The smart city middleware, from sensors to big data. In Proceedings of the 2013 7th International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS'13), IEEE (pp. 445–450).
Winograd, T. (2000). Towards a human-centered interaction architecture. Human-Computer Interaction in the New Millennium.
Xu, L.D., Xu, E.L. and Li, L. (2018). Industry 4.0: State of the art and future trends. International Journal of Production Research, 56(8), 2941–2962. https://doi.org/10.1080/00207543.2018.1444806
Zhuang, Y., Lin, F., Yoo, E.-H. and Xu, W. (2015). AirSense: A portable context-sensing device for personal air quality monitoring. In Proceedings of the ACM Mobile Health (pp. 17–22). https://doi.org/10.1145/2757290.2757293

(IN)VISIBLE CITIES

What Generative Algorithms Tell Us about Our Collective Memory Schema

EDUARDO RICO CARRANZA[1] and SHENG-YANG HUANG[2] and JULIAN BESEMS[3] and WANQI GAO[4]
[1,2,4]*Architectural Association School of Architecture*
[1,2,3]*Bartlett School of Architecture, University College London*
[1]*eduardoorico@aaschool.ac.uk, 0000-0002-1659-6166*
[2]*huang.sheng-yang@aaschool.ac.uk, 0000-0002-2748-5031*
[3]*j.besems.18@ucl.ac.uk, 0000-0003-2601-1173*
[4]*wanqi.gao@aaschool.ac.uk, 0000-0003-0717-0726*

Abstract. The last decade has witnessed a turn in AI technologies working with differentiable neural network architectures learning the embedded functions between data points and performing generative operations synthesising unseen data. The move to a continuous and generative AI paradigm aligns with ideas in the field of cognition and psychology, where a growing body of authors are beginning to conceptualise memory and our representation of the past as a dynamic, malleable and ultimately generative field. So, how effective are generative algorithms in supporting and enabling this creative process of remembrance? To answer this research question, we propose an experiment on how the spatial movement and exploration of maps of real and imagined images can help our brain reconstruct its memories in a dynamic yet accurate manner. We develop an application allowing visitors to dynamically explore real and AI-generated images of a given site clustered by similarity in a virtual 3D space. Analysing visitor paths and observed images helps us understand visitors' perspectives on real and AI-generated data such as an increased preference for synthetic images by visitors familiarised with the site. We conclude with recommendations on how to approach visitor experience in generative AI-powered applications for engagement with historical and archival data.

Keywords. Collective Memory, Embedded Differentiable Functions, Latent Space, Spatial Cognition, StyleGAN2, Schema, Visitor Paths

1. Introduction

The past decade has seen a shift in AI technology from pattern recognition to creative and generative capabilities due to growing datasets and a focus on neural network architectures that learn differentiable functions. This allows for not only pattern detection and identification but also generation of new data "in-between" the information provided during training, providing a smooth flow through a continuously defined data field. The generative turn in AI has had a significant impact on creative work that involves producing novel data, such as synthesized images, 3D, texts, music, and short videos. This paradigm shift is reflected in other fields of cognitive science, where the concept of memory has moved from a fixed object/pattern to a differentiable/generative approach. Citizens are now seen as continuously constructing a collective schema of the city through their individual memories.

These parallel shifts in AI and memory can lead to the creation of new technologies for visualizing and describing memory. While AI has mainly been used for synthesizing new data, few studies have applied AI to exploring our perceptions of the past. To address this gap, we propose an experiment that creates an app for visitors to explore an image catalogue combining real and imagined data. We develop a framework for analysing visitor experience and test it with our tool. Our results show the relevance of this form of work and potential for further development in engaging with memory and data archiving. This article provides background and methods used, as well as results and discussion of main outcomes.

2. Generative Paradigm of Collective Memory

Collective memory is a term that bears a strong currency in design and planning disciplines since it is meant to encapsulate a common representation of the past of our environment that finds its way to planning and design in various manners. The understanding of the collective origin of our memory as well as the methods to conceptualise and “generate” it is, therefore an important task that practitioners have traditionally given attention to, which is currently under strong scrutiny as part of a wider debate on cities.

Collective memory is a widely used term derived from the concept of social hallucination conceived by French sociologist Maurice Halbwachs as distinct from individual memory (Halbwachs, 1992). As it needs to embody its ongoing connection to contemporary discourse and identity, it also differentiates itself from history and collective remembering for its constructive nature in cognition (Wertsch & Roediger III, 2008). The creation of urban space and the exploration of cognition by collective memory began in the era of neo-rationalism. Typology was used to explain the methodology between collective memory and the representation of urban space, which is the so-called architectural rationality. Aldo Rossi argues that architecture can be generalised in the accumulation of history into a variety of typologies with certain definite characteristics. Accordingly, he advocates the 'analogical thought' retrieving the "archaic, unexpressed, and practically inexpressible" thought in memory. Therefore, Rossi argues that the goal of architectural design is not self-expression, but that it should fit by creating that similarity The collective memory of the residents of the city.

Christine Boyer argues that urban representations are always mediated in perceived reality; they substitute for objective reality and do not imitate it (Boyer, 1996). So how is the generative force of urban representation forms formed? How does it work? The collective memory of the city is an ongoing process of construction.

Traditional paradigms of the definition of memories of cities were conceptualised as records of objective facts. Perceptual data were stored in the brain in the form of sequences of precepted representations. Memory was a set of fixed items, stored in a chronological manner that could be retrieved and represented accurately provided we had enough evidence stored. By the mid-twentieth century, however, there was ample scientific evidence that the form of memory was not a sequence of precepted representations but its characteristic patterns. Such beliefs gradually became widely accepted scientific beliefs in the ensuing half-century. On the other hand, Piaget's generative epistemology and its related schema construction theory established the explanatory foundation of collective memory in the rational practice of urban design. Schema is a mental model, a pattern of knowledge and experience, which is not only a cognitive structure existing in memory but also a construction scaffold on which memory and knowledge depend. According to Jean Piaget, a schema can be represented as a classification system that can organise, generalise, modify, and create object information. The object can be recognised by the subject only after the transformative processing by the subject's mental structure, and the degree of the subject's knowledge of the object depends entirely on what kind of cognitive schema the subject has. In this sense, the object structure is established by the subject. The cognition of the object also evolves with the development of the subject cognition schema, which becomes what Piaget calls the construction of the object. Therefore, the development of cognition relies on individual activities to trigger the interactions between the subject and the object, in which the dual construction of the subject and the object is carried out (Chelstowski, 1971; Piaget, 1970, 1971, 2003).

Collective memory is thus the fabricated output of a schema of urban representations based on urban objects. Such a simulacrum feature responds to Boyer's claimed differentiation between perceived representations and reality and rejects the Platonian form of cities.

In parallel to this paradigm shift, a similar transformation has been happening in computational urban studies, from symbolism based on a given schema toward connectionism based on data-driven, ever-developing schema construction. The field of AI has seen a strong development of generative algorithms, particularly in the field of computer vision and image synthesis. Many of these techniques compress the information contained in the image into a numeric representation in the shape of vectors that can be manipulated to give birth to new images by asking the algorithm to perform a reverse journey of data synthesis. The structure of these abstract representations, also known as latent spaces, is commonly referred to as differentiable, meaning that the computer is programmed to learn the smooth, continuous functions that connect, interpolate and hopefully extrapolate between the training data. Understanding the nature of this latent space, and how it relates to the final images produced, has been the subject of a substantial amount of work in research and visual arts, which has lately produced a wealth of image-generation algorithms which are now finding their way into commercial products.

Working more specifically with historical images or typological studies of architecture, projects such as Brutal Nature (Moullinex, 2020) allow the exploration of imagined brutalist architecture while for the generation of novel architectures represented in a 2D scattergraph produced by a dimensionality reduction technique (Chen & Stouffs, 2021; Meng, 2021). These methodologies, such as t-SNE or other algorithms, help the representation and navigation of large image datasets, with implementations such as the PixPlot collection focusing on historical images (Duhaime et al., 2021). Working on a similar idea of spatialising data for navigation, Refik Anadol uses a t-SNE representation of the latent of GAN generated images of Gothic architectures (Anadol, 2019) to develop an explorer which brings up the image corresponding to the latent space linked to a particular 3D location of the camera. Similar attempts can also be seen in Immanuel Koh's 3D GAN Housing project. These cases all reveal the new, sub-symbolic, paradigm of spatial cognition and its construction (Koh, 2022).

3. Research Question

The exercises mentioned in the previous section focus strongly on the showcase of the creative capacity of these AI-powered algorithms, making an emphasis on the spatial qualities of the represented data and the emerging aesthetics in the images. There is less attention on the visitors themselves and the understanding of how these tools and the data produced are perceived as a form of documenting history or memory. There seems to be a gap in the research when studying how we can engage with this type of algorithm and how we react to the nature of the interpolated data in relationship with the real data. As a result, the following research question is formulated:

How do people observe datasets that combine AI-generated and real images of a given space? and how does their prior knowledge influence this experience?

Studying the engagement with these algorithms applied to historical datasets may help us understand how we approach a connectionist definition of the construction of heritage, memory and to archival material in general. The following sections describe an experiment where we begin setting up the methods of navigation of memory as well as forms of analysing our behaviour. We try to understand what the images that would typically call our attention are and what are how different visitors react to genuine and imagined AI-generated images.

4. Experiment Definition

Our experiment consists of the deployment of a 3D interactive app where visitors freely navigate and explore real and imagined pictures of a given place. By deploying the same app with different groups of participants who have different levels of familiarity with the site, we can develop comparisons that can help address the questions on the weight of prior knowledge on how we look at imagined and real images.

The space chosen is the Barrel Vault, in the Architectural Association School of Architecture, London. This is a narrow, elongated room with a distinct window system and a ceiling formed by a vault-like concrete and wood roof system (see Figure 1, left). The space has been traditionally used for pinups, exhibitions as well as teaching and lecturing and has a long tradition as a defining space within the school. The exploration of the images of the Barrel Vault, both real and imagined, is done through an app that is accessible through a link and is installed on the user's PC. The usage of the app is done on the computer screen, without the need for any particular tool and is carried out independently, without support or guidance from the research team. User instructions are briefly given via an introductory video supplied with the app link altogether with an explanation of the experiment as well as how data would be recorded and used. The interface records the movement of the visitor in this space as well as the index images being observed. This data is later processed to understand the visitor experience and, to an extent, their engagement with the images of the past.

An initial set of images of the Barrel Vault, which we shall denominate "real" images, come from several sources, both archival and current. We obtained 131 archive images from the AA historic database of school photographs, which contained photographs of events, presentations, performances and general school life dating back three decades. We then complemented them with 569 images obtained from 2 video shootings of the space carried out by the research team. The archive images are more likely to include a diversity of angles, themes and textures, while the ones coming from the video shootings provided a stronger quantity of information on spatial textures and overall structure. These "real" images consist of what we will call the "anchors" dataset.

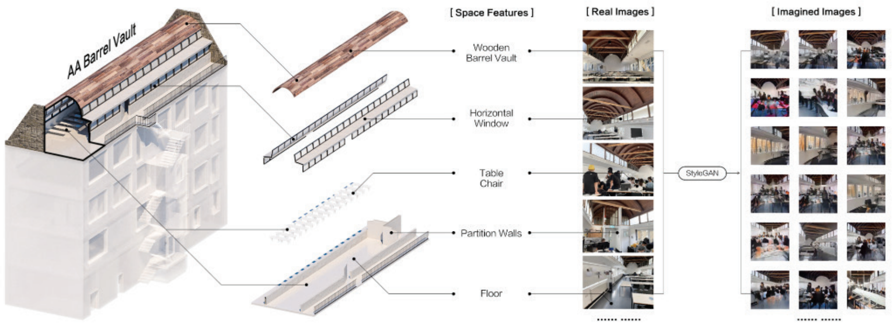


Figure 1. Site location in Barrel Vault, Architectural Association, London and real and generated images

The anchors are then fed to a StyleGAN2 algorithm (Karras et al., 2020) that is trained over 6000 Epochs and used to produce 20,000 "imaginary" pictures of the same space. Given the nature of the training, which in this case took place departing from an empty network, StyleGAN2 is likely to learn, and therefore replicate during inception, details and structures that are provided more often in the training dataset. These are likely to be some structural characteristics of the space, such as the ceiling, the repetitive nature of the roof beams and elongated windows. These appear in the majority of the images and are a distinctive feature of the room. The representation of people and activities as well as other details (objects, furniture or similar) are not so well detailed by the algorithm and would typically appear more blurred or unclear (Figure 1 top right). As a result, this second set of imaginary pictures is likely to have a distinct character when explored in detail, with a higher degree of vagueness or lack of definition.

The dataset composed of both “real images” (or anchors) and “imagined pictures” is then located in the virtual 3D space of the app that the visitor then can navigate freely (Figure 2 above) on their computer. The exact position of the images is estimated according to their visual similarity with other images forming clusters or coherent groups. This is done by feeding the images into an image classification neural network (VGG19) and extracting an intermediate abstract representation of these images as features. A dimensionality reduction algorithm is then run on these features with a t-SNE algorithm, which produces one 3D point in space per image. The algorithm generates the position of the points preserving similarity between images, hence clustering together images with common visual characteristics. When using the app, The visitor always enters the 3D virtual space in the same location within the 3D that contains all images (anchors and imagined). In the first instance, only the anchors (real images) are visible to the user. The user is led to focus more on the anchors close to the location of the camera thanks to a “mist” effect that fades distant objects in the background. At all moments, all images are rotated towards the camera, allowing the user to focus more on the immediate environment. The visitor can move through the environment using the mouse and keyboard and click on the existing anchors. By doing so, the app reveals all images in a nearing radius that “spawn” radially from the anchor and appear in their respective positions. The process can go on so that when clicking on any further image (anchor or imagined), more imaginary images are revealed (Figure 2 below). Clicking repeatedly on an image enlarges the radius until a certain maximum. The app records the path that is navigated by the visitor as well as the images clicked.

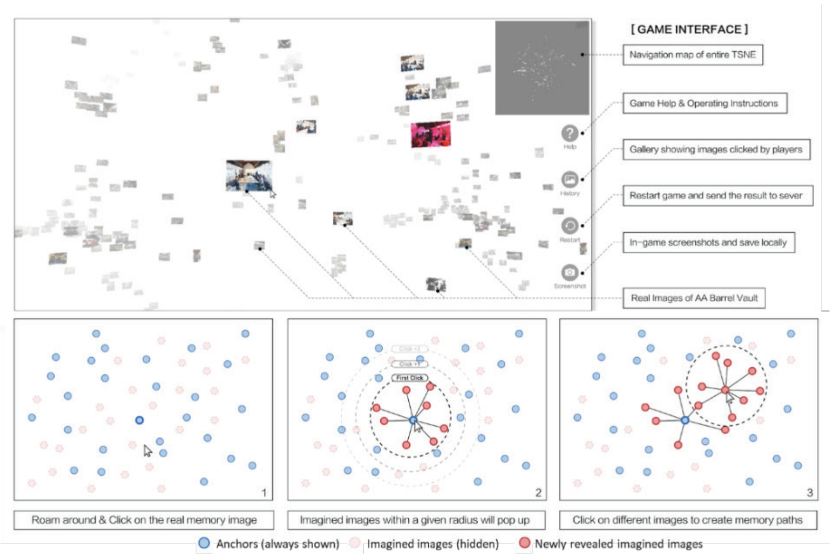


Figure 2. Site location in Barrel Vault, Architectural Association, London and real and generated images

The underlying assumption is that different knowledge and familiarity with the site being represented (Barrel Vault) may yield different forms of exploration. It is to be expected that people that already know the place may focus their attention on different types of images (real and imaginary) and that this can influence the ratio of the images visited. To evaluate these hypotheses data for each group previously mentioned is recorded separately and the results are used in a comparative study.

5. Experiment Results

The model was run a total of 59 times with 19 visitors familiar with the Barrel Vault site and 40 visitors that were not familiar. On each occasion, we gather data on the use of the tool, such as the time spent on the application, movement path and speed and images clicked for spawning. This last is used as a proxy for the attention given to a specific image. On average, the uses tended to be between one and 5 minutes, with some visitors taking much longer periods to evaluate the tool (Figure 3). We found cases where the tool was open for more than 2 hours, suggesting that several visitors were going through the exercise or that the tool was inactive for a large proportion of the time. These cases (three in total) were removed from the average analysis since they proved to be disorienting for the final results. These results seem to suggest a relatively good amount of time, especially considering that some of the visitors had no direct connection with the school or the research team.

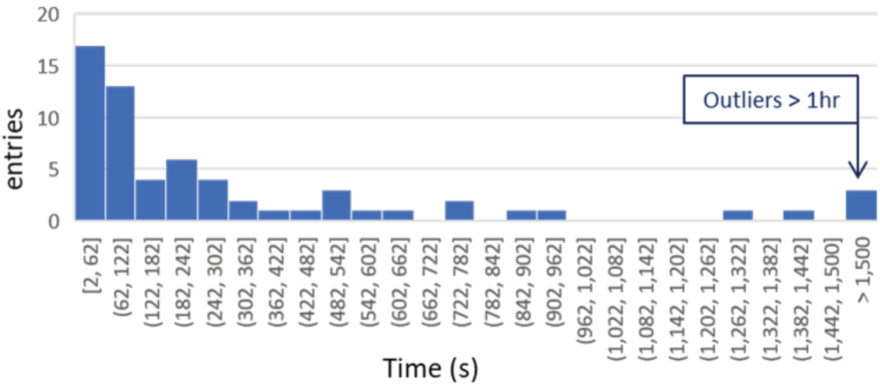


Figure 3. Distribution of total entries according to the time spent on the tool

Looking at a comparative analysis between visitor groups (Table 1 and 2), we saw that the average usage times for both groups are remarkably similar around 250 seconds, indicating a similar visitor engagement. An equal pattern can be observed with the number of images clicked, with both groups close to the aggregate average of 30 images clicked. On average, visitors click generated images more than twice as many as anchors (2.3 overall generated/real ratio). While there are many more generated images than anchors, only the last group is effectively shown when entering the interface. This means that the visitors kept unveiling newly synthesized images and explored them twice as often as they did with the real ones, hence suggesting a good degree of curiosity. The differences begin to appear when we look at the types of images clicked by these groups. The visitors that were familiar with the site did, on average 3.1 imaginary queries per real query, and those non-familiar with the site did substantially less (2.1). This is to be expected since people unfamiliar with the site are more likely to seek to familiarize themselves with the Barrel Vault before looking at the imaginary pictures that may appear harder to interpret since they have blurry or sketchy aspects to them due to their generative process. On the contrary, people that already know the place, entertain themselves longer by looking at interpolated or generated pictures of what they already know. Those familiar with the Barrel Vault which, by definition, will have more specific memory of the place, are likely to have a stronger engagement with abstract representations of the site. This could be linked to Piaget's schema construction mechanism. Strong prior knowledge of a location can lead to the enhancement of existing schemas or prompt the accommodation process to form updated spatial cognition through the introduction of heterogeneous visual stimuli from the generated images. If valid, Aldo Rossi's mental similarity hypothesis allows for this possibility to impact the collective memory.

		Average values across the group				
Metric	Valid entries	Total duration	Images clicked	Real images	Generated images	G/R
Unit	Number	Second	Number	Number	Number	Ratio
Total	59	250.7	30.1	9.0	21.0	2.3
Familiar	19	248.5	28.9	7.1	21.8	3.1
Unfamiliar	40	251.8	30.6	10.0	20.7	2.1

Table 1. Average visitor indicators per group.

		Variance across the group			
Metric	Valid entries	Total duration	Images clicked	Real images	Generated images
Unit	Number	N/A	N/A	N/A	N/A
Total	59	95574.0	3600.1	216.3	2227.4
Familiar	19	126582.5	1180.8	42.1	944.0
Unfamiliar	40	83709.5	4808.1	299.6	2876.4

Table 2. The variance of each indicator per group. Higher variance indicates less homogeneity in the observation.

We then turn to look at the images themselves and we produce an analysis of the images clicked by each visitor (Figure 4). We can see how some visitors click several times on a given anchor. Each time this happens, the app reveals a growing number of hidden images until these are exhausted. We can see how most of the clicks are repetitive as if the visitor was trying to “squeeze” as many images as possible around a given anchor before moving to the next. We equally carry out a study of the popularity of images by aggregating the number of clicks they attracted (Figure 5). Results from the exploration indicated that a larger proportion of the time spent by participants was exploring images with people and activities, rather than static objects. Equally, images with strong colours (exhibition or similar) disproportionally call the attention of the visitor when navigating through the t-SNE and are more frequently clicked. This also happens with images of people performing activities (pinups, designs or crits) which attract more attention than images of space.

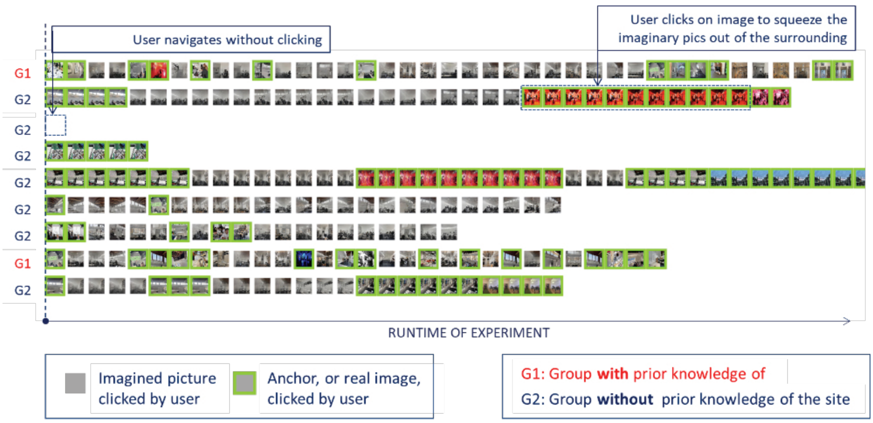


Figure 4. Examples of visitor exploration

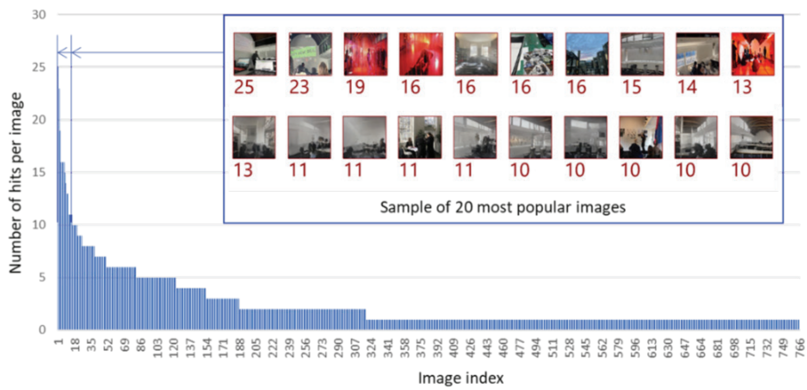


Figure 5. Popularity analysis and clicked figures

6. Conclusion

We have proposed an innovative form of using AI to study image-based archival material which stems from a continuist approach to coding as well as memory. We have tried to use algorithms that both organise images (t-SNE) as well as interpolate between them (StyleGAN2) and tried to study their deployment with visitors.

Moreover, we have proposed methods of measuring the engagement of visitors with these tools and tried to extract conclusions on how the analytical data can relate to visitors' preferences and backgrounds. Our results indicate that those familiar with the site were more interested in the imagined pictures, but were more homogeneous in their use of the tool. This tendency may be linked to the behaviours of assimilation and accommodation in the collective schema-construction process of the group. This would suggest a different approach to questions of memory and representation, that, could be argued, relate to the terms of Piaget's schema. It reveals new tools for analysing the mechanisms by which the spatial cognitions of cities are collectively formed. While it is too soon to extract definitive conclusions, this experiment provides insight into how our relationship to images, memory and cities can be studied and approached via generative algorithms.

References

Anadol, R. (2019). *Refik Anadol | Architectural Photographic Memories. Parametric Architecture.* Retrieved December 6, 2022, from https://www.youtube.com/watch?v=yCC5GgakAeE.

Boyer, M. C. (1996). *The City of Collective Memory: Its Historical Imagery and Architectural Entertainments.* MIT Press.

Chelstowski, B. (1971). The Constructivistic Structuralism of J. Piaget. *Science and Society,* 35(4), 481-489.

Chen, J., & Stouffs, R. (2021). *From Exploration to Interpretation-Adopting Deep Representation Learning Models to Latent Space Interpretation of Architectural Design Alternatives,* CAADRIA 2021 (pp. 131-140). The Association for Computer-Aided Architectural Design Research in Asia (CAADRIA), https://doi.org/10.52842/conf.caadria.2021.1.131.

Duhaime, D., Leonard, P., Eskildsen, T., Choudhary, S., DeRose, C., Sanger, W., Arneja, M., Fish, E., Sorba, O., & Reagan, D. (2021). *PixPlot*. Yale Digital Humanities Lab. Retrieved 10 November from https://github.com/YaleDHLab/pix-plot

Halbwachs, M. (1992). *On Collective Memory.* University of Chicago Press.

Karras, T., Laine, S., Aittala, M., Hellsten, J., Lehtinen, J., & Aila, T. (2020). *Analyzing and Improving the Image Quality of StyleGAN, The IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 8110-8119). Institute of Electrical and Electronics Engineers, https://doi.org/10.48550/arXiv.1912.04958.

Koh, I. (2022). Architectural Plasticity: The Aesthetics of Neural Sampling. *Architectural Design,* 92(3), 86-93, https://doi.org/10.1002/ad.2818.

Meng, S. (2021). Exploring in the Latent Space of Design: A Method of Plausible Building Facades Images Generation, Properties Control and Model Explanation Base on Stylegan2. The International Conference on Computational Design and Robotic Fabrication (pp. 55-66). DigitalFUTURES.

Moullinex. (2020, 6 October 2020). *Brutal Nature.* https://www.facebook.com/moullinex/videos/342519493629041/?redirect=false

Piaget, J. (1970). *Genetic Epistemology.* Columbia University Press.

Piaget, J. (1971). *Biology and Knowledge: An Essay on the Relations between Organic Regulations and Cognitive Processes.* Edinburgh University Press.

Piaget, J. (2003). *The Psychology of Intelligence.* Routledge.

Wertsch, J. V., & Roediger III, H. L. (2008). Collective Memory: Conceptual Foundations and Theoretical Approaches. Memory, 16(3), 318-326, https://doi.org/10.1080/09658210701801434.

URBAN CULTURAL INHERITANCE

Generative Adversarial Networks (Gans) Assisted Street Façade Design In Virtual Reality (Vr) Environments Based On Hakka Settlements In Hong Kong

JIAHUA DONG[1], QINGRUI JIANG[2], ANQI WANG[3] and YUANKAI WANG[4]
[1]*Chinese University of Hong Kong*
[2,4]*University College London*
[3]*Hong Kong University of Science and Technology*
[1]*jiahuadong@cuhk.edu.hk, 0000-0003-2400-8819*
[2]*qingruimi@outlook.com, 0000-0003-0740-7383*
[3]*awangan@connect.ust.hk, 0000-0003-4238-6581*
[4]*ucbqy552@ucl.ac.uk, 0000-0001-5107-5932*

Abstract. In Hong Kong, the Hakka settlements are the home of indigenous people who have been involved in agriculture and fishing for over 200 years, which has a special place in Hong Kong's history. However, these settlements are gradually being abandoned as ghost towns due to rapid urbanisation, where the city is progressively constructing high-density habitats to accommodate the exponentially increased population since the 1950s. This challenges designers to rethink means of preserving urban cultural heritage, while engaging in continuous urban regeneration processes. This study investigates workflows to detect historical building styles in one of the most densely-populated cities in the world - Hong Kong - that further deployed in human-computer interfaces in the virtual reality (VR) environment as a collaborative and suggestive design decision-making tool and co-design engagement tool for style infilling on urban regeneration to maintain urban culture inheritance. The workflow aims to discuss the preservation of historical buildings in the context of urban regeneration, urban cultural inheritance through digital archiving, bottom-up community engagement and education in urban design. This study discusses the impact of emerging technologies on the making of communities in human-centric discourse, rethinking social power and the right to the city in participatory mechanisms.

Keywords. Urban Culture Inheritance, Hakka Settlements, Facade Generation, Human-Computer Interaction (HCI), Virtual Reality (VR)

1. Introduction

Urban regeneration has generally become an irreversible process in modernisation.

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 473-482.

Designers and researchers are being challenged to consider more complex socio-technical mechanisms in the post-rational moment. Urban culture, as a reflection of a sense of belonging in community making, has raised the attention of planners. Cultural inheritance refers to the preservation and dissemination of knowledge on cultural information through communication, imitation, teaching and learning in the long term (Peedicayil, 2001). Thus, in urban cultural inheritance is the preservation and heritage of physical buildings, objects, and landscapes, but also the intangibility of how spaces are being inhabited, methods to design production and construction, and public education.

In Hong Kong, Hakka settlements are habitats of the indigenous populations who have been engaging in agricultural and fishing activities for more than 200 years; especially in the New Territories, which has a special place in Hong Kong's colonial and policy history (Kong, 2018). Lai Chi Wo, Hong Kong's oldest, largest and best-preserved rural settlement, has won UNESCO's prestigious 2020 Special Recognition for Sustainable Development award for promoting cultural heritage preservation. The social capital of Hakka villages is the relationship inside the family's settlements - goodwill, fellowship and social intercourse, which can be manifested through their settlements as the symbol foundation of a sense of belonging (Lin, 2022). However, these settlements are gradually being abandoned as ghost towns due to rapid urbanisation, where the city is progressively constructing high-density habitats to accommodate the exponentially increased population since the 1950s (Wong, 2013).

Urban regeneration may help in modernising old urban fabrics and standardisation lowers the marginal costs of construction. It can sometimes produce homogeneous environments that are not inclusive of varying community needs (Jacobs, 1961). While the visual qualities of historic street spaces may generate a sense of cultural identity and belonging, it may be difficult to accommodate typologies and other contemporary needs (Raman, 2014). Urban planning is an iterative, collaborative process that frequently only involves designers and civic stakeholders (Kim et al., 2022). Andre Leroi-Gourhan (2002) argued that "innovation happens only when technical supply matches culture demand." The digitalisation can be a backcasting to tradition and bring urban design back to everyone through the interactions between humans and machines (Negroponte, 1969). Generative adversarial networks (GANs), firstly introduced by Ian Goodfellow and his colleagues (2014), as an artificial intelligence framework mainly used to automatically generate images, are capable of this reproduction of traditional cultural embodiments in digitalisation by fusing two styles of datasets. However, such research lacks data-based metrics in terms of evaluation of user feedback and consideration of data collection, which refers to reflection on the demands of citizens. The research on Human-Computer Interaction (HCI) concerning the design and evaluation of user interfaces can address this problem by utilising the participants' responses. A digital twin is a digital replica of a physical asset, system, or process that can be used for documentation, simulation and evaluation (Ng et al., 2022). With the burgeoning VR technology, virtual environments as a representation of the digital twin allow citizens to truly experience the automated generation of urban regeneration approach, which can take advantage of the HCI approaches as well.

In this paper, we implemented a framework for participatory design in respect of replacing the city facades in an immersive VR, mainly through an online VR platform and the image generated by CycleGAN automatically. This research aims to generate design options to guide collective decision-making processes with the use of neural networks and use immersive digital twins to best utilise community engagement in VR environments with user-friendly interfaces. This study also aims to look at the transformation from historical settlements to modern society in the urban regeneration process in the revolution of digitalisation. The final goal of this research is to build participatory mechanisms and shared visions for designers and citizens in historical heritage preservation, and co-design for future cities using meaningful intercultural interactions to inherit urban culture.

2. Related Works

GANs can be used to suggest design options that act as a communicative interface between human and computer. GANs have a few variants, one of the types is CycleGAN (Zhu et al., 2017). It enables the adding of more generators and discriminators to train unpaired images and facilitates more flexibility in training processes for design and decision-making (Avinash, 2022). Since the CycleGAN is trained by having the forward model generate an output image from the input image, and then having the reversal model attempt to recover the original input image from the output image. This process helps to ensure that the output image retains the content of the original image, while taking on the style of the target domain. CycleGAN has shown to be very useful in facade design by using labelled images to customise the design of facades. For example, Sun et al. (2022) worked on fine-tuning the algorithms and optimising the dataset in CycleGAN-assisted facade generation in Renaissance, Baroque-style buildings, which has successfully been tested and applied to other urban blocks. Also, CycleGAN has been used to mix two types of facade styles that has its own characteristics for urban infill (Khairadeen Ali & Lee, 2021). In design practices, David Ruy at SciArc explored the GAN-assisted style transfer of how architectural form and design are reproduced in the digital world in mixed approaches, where the machine learning tool is acting as a speculative designer that provides design options beyond human intelligence.

Digital twins can be used for a wide range of purposes, for example, in urban archives. It is a concept that has emerged in the field of the Internet of Things (IoT), where sensors and other devices are used to gather data from the physical world and create a virtual model of it (Ng et al., 2022). In the case of Hakka settlements regeneration, digital twins can be created by capturing detailed information about the physical layout and characteristics of the architecture. The Hong Kong government has been using drone scanning to document the high-resolution photogrammetry model in all built areas since 2017, encouraging more designers and researchers to use digital twins in urban design and planning.

Participatory design has been discussed since the 1960s (Lynch, 1960; Arnstein, 1969). With the development of HCI, VR technology has the potential to be a valuable tool in participatory urban design by engaging citizens in the design process since HCI concerns the design and evaluation of user interfaces, the development of new interaction techniques and technologies, and the understanding of how people use and respond to technology. Kim et al. (2022) deployed pix2pix and CycleGAN into real-time desirable street design tools, which allows citizens to stylise their own city streets.

Leong and Janssen (2022) developed an interactive platform for citizens to preserve a heritage site in Singapore. Loyola et al. (2019) used VR immersive scenes to collect data on the human sense of the quality of space. Also, VR-generated scenarios can be used for treating psychological issues in humans using participatory prototyping and storyboarding (Flobak et al., 2019). By using VR technology, citizens and different community members can participate in the design process in a more interactive and immersive way as a collaborative intelligence.

From previous research, we found there is a lack of research on the preservation of urban cultural heritage. Also, the research on participatory design using VR is limited.

Hence, our research questions were formulated as follows:

- How to automatically generate design options with the use of GANs to inherit urban cultural heritage in the discourse of urban regeneration?
- How can immersive digital twins in VR be best-utilised co-design processes and used in community engagement?

3. Methodology

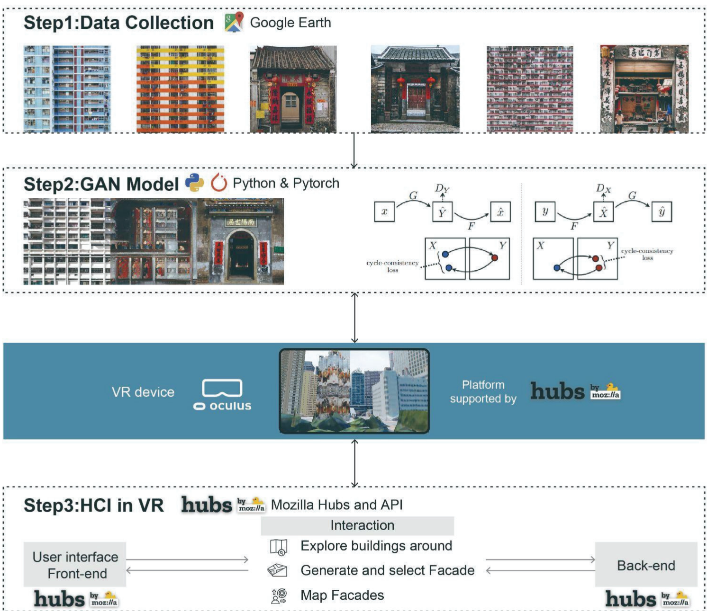


Figure 1. Participatory facades design architecture

This research project focuses on the indigenous Hakka settlements in Hong Kong. The methodology on participatory GANs-assisted automatic generation of façade design comprises three steps (see Figure 1) - dataset collection and processing, style customisation, and performance testing. In the dataset construction part, the façade photos of historical Hakka settlements are collected in high resolution and composed into a dataset that embeds cultural information as well as the normal Hong Kong high-density residential facades photo dataset. In the training and testing process, CycleGAN is used to train both Hakka settlements' façades and Hong Kong's exceptional high-density residential facades to build architectural style transfer models. In the next step of this project, we deployed the generated model as a social engagement and interactive public communication tool on urban culture, and the collaborative generation results will be adapted to virtual districts to test performance.

3.1. FACADE GENERATION

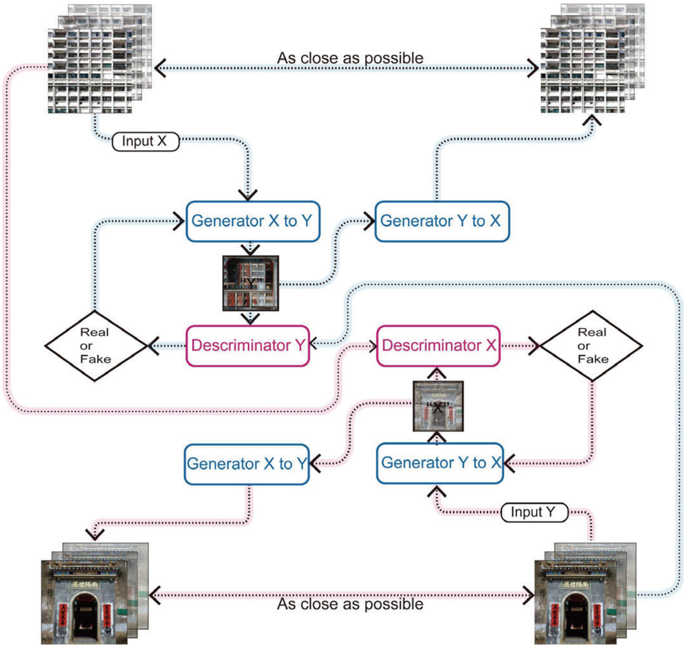


Figure 2. The architecture of generated facades by CycleGAN

The two tiles facades datasets are collected from Google Street View. We collected both residential areas of Hakka settlements and Hong Kong modernist facades settlements. In order to effectively synthesise images, we pre-processed the dataset using OpenCV camera calibration algorithm to autocorrect image distortion to ensure the quality of images. By cropping and resizing the images, we removed any unnecessary background to make sure the size of 256 by 256 pixels, is appropriate for the model. Also, the datasets were augmented by mirroring and rotation.

In the training process, two data tiles are fed into the network separately. We split the dataset into a training set (200 images) and a validation set (20 images). The training set was used to train the model, while the validation set was used to evaluate the performance of the model during training.

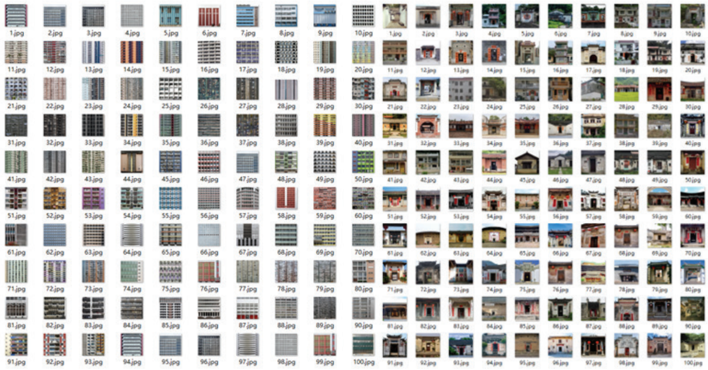


Figure 3. Image data preparation (a left) training set A from modern Hong Kong facades and (b right) training set B from Hakka facades

We defined the architecture of the CycleGAN model and involved two generators and two discriminators - Gxy and Gyx, and 2 discriminators - Dx and Dy. i.e., X to Y to capture the Hakka cultural element from training sets, while the discriminator Dy starts to evaluate the similarity (see Figure 2).

Deep learning framework, TensorFlow and Tensorflow-examples (1.15.0) were used to construct the generators and discriminators (Zhu et al., 2020; Sun, Zhou & Han, 2022). Then, a suitable loss function and optimisation algorithm should be used during training. To train the model effectively, we choose a loss function that captures the quality of the generated images, and an optimisation algorithm that can effectively adjust the model's parameters to minimise this loss.

While training, image X is passed through generator Gxy, which generates image Y^. The generated picture Y^ is passed through generator Gyx, which loops through to generate picture X^. The mean absolute error (MAE) is calculated between X and X^.

Forward cycle consistency loss: X –> generator Gxy (X) –> generator Gyx (Gxy (X)) ~ X^

Backward cycle consistency loss: Y –> generator Gyx (Y) –> generator Gxy (Gyx (Y)) ~ Y^

We used the formula (Zhu et al., 2020; Sun, Zhou & Han, 2022):

$$Loss_{GAN}(G_{xy}, D_y, X, Y) = E_{y\sim pdata(y)}[\log D_y(y)] + E_{x\sim pdata(x)}[\log (1- D_y(G_{xy}(x))] \quad (1)$$

Cycle consistency loss was introduced to push F(G(X)) ≈ X and G(F(Y)) ≈ Y

The formula for the cycle-consistency loss used in a CycleGAN model is typically based on the L1 distance between the input and generated images after they have been transformed back and forth between the two domains (2)(3) (Zhu et al., 2020; Sun, Zhou and Han, 2022).

$$Loss_{cyc}(G_{xy}, G_{yx}) = E_{x\sim pdata(x)}[||G_{yx}(G_{xy}(x)) - x||_1] + E_{x\sim pdata(x)}[||G_{xy}(G_{yx}(y)) - y||_1] \quad (2)$$

$$Loss_{identity}(G_{xy}, G_{yx}) = \lambda_x * E_{x\sim pdata(x)}[||G_{yx}(G_{xy}(x)) - x||_1] + \lambda_y * E_{x\sim pdata(x)}[||G_{xy}(G_{yx}(y)) - y||_1] \quad (3)$$

The overall loss function used in a CycleGAN model is therefore a combination of the GAN loss, the cycle-consistency loss and identity loss, weighted by the regularisation parameters lambda (4) (Zhu et al., 2020; Sun, Zhou and Han, 2022):

$$Loss(G_{xy}, G_{yx}, D_x, D_y) = Loss_{GAN}(G_{xy}, D_y, X, Y) + Loss_{GAN}(G_{yx}, D_x, Y, X) + Loss_{cyc}(G_{xy}, G_{yx}) + \lambda_i * Loss_{identity}(G_{xy}, G_{yx}) \quad (4)$$

The values of lambda were chosen between 10 and 50 during training to improve the performance of the model. We tested 100, 200, and 300 epochs for the training process and evaluated the generated power by using the 300 epochs model.

Lastly, once the model has been trained, we synthesised new facade images by providing it with input images of the appropriate type (i.e., the Hakka style). The model then generated a transformed version of the input image, formulating a mixed-style facade. We also manipulated the training dataset at the same time, replacing part of the facades of Hakka settlements with Hong Kong modernist facades to reduce the synthetic ratio and overfitting as an attempt.

3.2. HCI WORKFLOW IN VR

We built a platform for various participants to take a tour and participate in the design by selecting and projecting the facade results generated by CycleGAN onto the buildings they have willing to regenerate according to their own purpose and preference in the digital village.

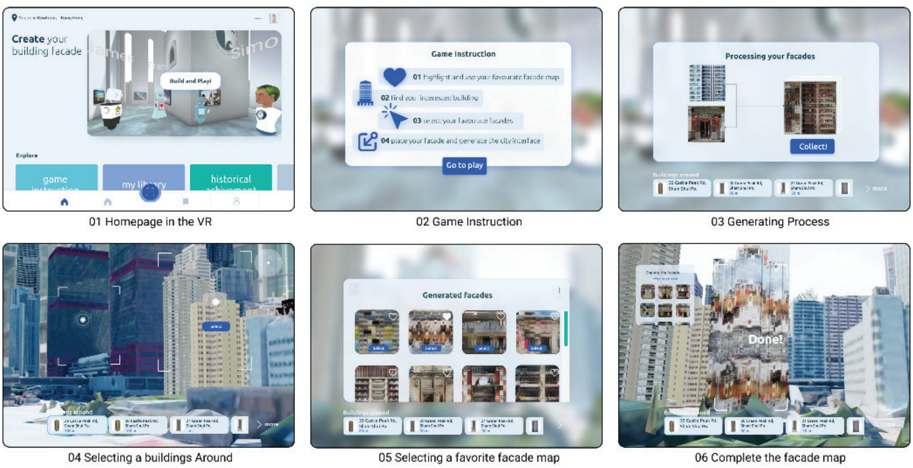


Figure 4. The user experience of the core functions in VR

The HCI development contains three parts. Firstly, we built a Hong Kong-based city scene for the tour and interaction in the digital twin environment. We used Spoke to export the Blender scene to the Mozilla hubs, which enables the tour with VR devices. Secondly, a series of user-friendly interfaces (UI) in the VR was designed for citizens to play and interact with. Thirdly, we designed multiple interactive functions concerning the user experience, which embedded the real-time CycleGAN style transfer algorithm to achieve real-time generation. The user experience includes exploring the buildings around, selecting the facades, and selecting and mapping the facades (See Figure 3). For visualising the generated facades in 3D view, we used the projecting UV map function in Blender 2.91.0, an open-source 3D modelling software (See Figure 4). The visualisation integrated both a bird's eye view and a panorama view, which provides citizens with an immersive and comprehensive view of the generated facades.

The digital version of this reproduction has the capability of simulating different design scenarios and evaluating their impact on the conservation of cultural heritages by allowing the local community and design team to collaborate and explore different options. This HCI in VR can involve conducting workshops and other engagement activities with members of the community to gather their input and ideas for the design of the village. By involving the community in the design process, it is possible to incorporate their cultural values and traditions which can help to preserve the unique character of the Hakka community to accommodate the preferences of the citizens.

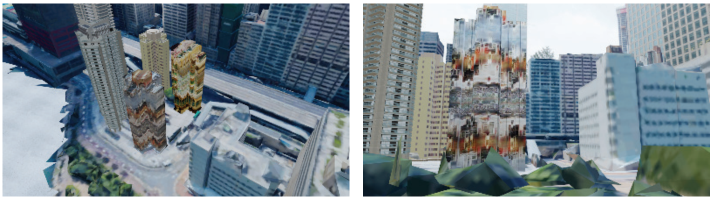


Figure 4. Demonstration of bird's eye view and 360° panorama view of generated design proposals

4. Discussions

This research demonstrated a participatory workflow by deploying the CycleGAN style transfer model of Hakka settlements into an HCI system. The CycleGAN style transfer model proved useful as a tool to stylise facades as design options using cultural images. The HCI workflow enables citizens' participation in the customisation of street facades, to ensure the engagement of the community making and education on the preservation of urban cultural heritage. Through the community engagement, there will be certain potential to see greater citizen awareness of the conservation of abandoned settlements and Hakka inhabitants gaining a sense of belonging in contemporary environments.

There are a few difficulties that need to be addressed. Researchers need to take citizens - both indigenous Hakka communities and majority groups as actors rather than data subjects in the co-learning process, which includes the identification of design problems and appropriate informative scopes (Ng, 2022). In addition, designers should take the level of guidance and level of interactivity into account to provide design options. Also, the design of beginner-friendly and user-friendly interfaces can add value to digital tools by bridging physical and virtual interfaces, which are the key dialogue windows for communication. Furthermore, the harmonic approach of the building mixed with Hakka settlements and Hong Kong modernist buildings is relatively mediated by the algorithm. The algorithm fine-tuning is also an important step to construct identities for specific inhabitants by keeping recognizable original fabrics.

In the next stage, the design practices can be divided into three phases. In the first phase, community engagement with VR will be conducted. Designers should take the role of facilitator to provide design options and guidance, and also help citizens to visualise their visions and preferences. In the second phase, the customised design toolkit with neural networks helps us to understand the synthesis of cultural heritage with co-defined data by architects and citizens to increase the sense of achievement, integrating building functions and facade design and incorporating them into VR as immersive digital twins. In the final step, data output should be taken as a feedback loop for the next stage of community engagement.

Our future direction will be optimising the CycleGAN style transfer model by enlarging the dataset and adding labelled images into training. The image segmentation network to label the collected facades can be added to the user interface, which provides the citizens more flexibility in the co-design process.

5. Conclusion

This research presented a GANs-assisted real-time participative street facade design tool in VR environments. The design of toolkits can facilitate the generation of cultural facades and design options for citizen participation and public communication, with the purpose of generating a sense of identity, shared expression, and ensuring urban culture inheritance, using user-friendly interfaces in VR. This research specifically focused on the Hakka settlements that are under threat of being demolished, using digital twins to preserve the urban cultural heritage. As we embrace the digital, our environment explodes in complexity with various opportunities. The Digital revolution is accelerating the evolution of our living environment, providing efficient accessibility of collaborative interaction tools to co-design our cities. The participatory community forms and evolves dynamically in the metabolism in cities, which extends the consideration of urban cultural inheritance, quality of life, sustainability and well-being.

References

Arnstein, S. R. (1969). A ladder of citizen participation. *Journal of the American Institute of planners*, 35(4), 216-224, https://doi.org/10.1080/01944366908977225.

Ashwath, B.,(2020), CycleGAN Facades-Segmentation Translation, https://www.kaggle.com/code/balraj98/cyclegan-facades-segmentation-translation-pytorch/notebook.

Audouze, F. (2002). Leroi-Gourhan, a philosopher of technique and evolution. *Journal of Archaeological Research*, 10(4), 277-306, DOI:10.1023/A:1020599009172.

Avinash, H.,(2022), The GAN Zoo. https://github.com/hindupuravinash/the-gan-zoo.

Campo, M. (2021). Architecture, language and AI-language, attentional generative adversarial networks (AttnGAN) and architecture design. *In Proceedings of the 26th CAADRIA Conference (Vol. 1, pp. 211-220).*

Chowdhury, S., & Hanegraaf, J. (2022). Co-presence in Remote VR Co-design: Using Remote Virtual Collaborative Tool Arkio in Campus Design. DOI:10.52842/conf.caadria.2022.2.465

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. Paper presented at: *Advances in Neural Information Processing Systems, https://doi.org/10.48550/arXiv.1406.266.*

Jacobs, J. (1961). *The Death and Life of Great American Cities.* Randoms House, New York. Book Unpublished resources.

Flobak, E., Wake, J. D., Vindenes, J., Kahlon, S., Nordgreen, T., & Guribye, F. (2019, May). Participatory design of VR scenarios for exposure therapy. *In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1-12), https://doi.org/10.1145/3290605.3300799.

Khairadeen Ali, A.; Lee, O.J. Facade Style Mixing using Artificial Intelligence for Urban Infill. Preprints 2021, 2021040690, doi: 10.20944/preprints202104.0690.v1.

Kim, D., Guida, G., Garcia, J. L., & Lopez, D. C. Y. Participatory Urban Using Generative Adversarial Networks, https://doi.org/10.52842/conf.caadria.2022.2.485.

Kong, Y. C. (2018). Transformation of a Hakka Village in Hong Kong: A Case Study of the Lo Wai Village, *quanqiukejiayanjiu*, (10), 209-234.

Lin, C. Y. (2022). Reconstruction of the Hakka Family Space: The Case of the Preservation of Meinung's Kiu Tiam Gui Fo Fong, *quanqiukejiayanjiu*, (19) 105-134

Loyola, M., Rossi, B., Montiel, C., & Daiber, M. (2019). Use of virtual reality in participatory design. *Blucher Design Proceedings*; Editora Blucher: Porto, Portugal, 449-454, DOI: 10.5151/proceedings-ecaadesigradi2019_156.

Lynch, K. (1964). *The image of the city.* MIT press.

Negroponte N. (1969). 'Toward a Theory of Architecture', *Journal of Architectural Education*, Blackwell Publishing on behalf of the Association of Collegiate Schools of Architecture, Inc., March 1969, https://doi.org/10.47982/spool.2022.1.03.

Ng, P. Fernandez, A. Kalema, N. (2022) 'Digital Commons: Decolonial Urbanism in Datafied Society'. *Conference: IFOU 2022.*

Ng P, Li Y, Zhu S, Xu B and van Ameijde J (2022), Digital common(s): the role of digital gamification in participatory design for the planning of high-density housing estates. Front. *Virtual Real*. 3:1062336. doi: 10.3389/frvir.2022.1062336

Peedicayil, J. (2001). The importance of cultural inheritance. *Medical hypotheses*, 56(2), 158-159.

Raman, S. (2014). Sense of belonging. *Encyclopedia of quality of life and well-being research.* Springer. https://doi. org/10.1007/978-94-007-0753-5_2646.

Sun, C., Zhou, Y., & Han, Y. (2022). Automatic generation of architecture facade for historical urban renovation using generative adversarial network. *Building and Environment*, 212, 108781, DOI:10.1016/j.buildenv.2022.108781.

Wiener, J. M., & Franz, G. (2004, October). Isovists as a means to predict spatial experience and behavior. *In International Conference on Spatial Cognition*. (pp. 42-57). Springer, Berlin, Heidelberg, https://doi.org/10.1007/978-3-540-32255-9_3.

Wong, Y. C. (2013). The Population Numbers Challenge Rekindling Hong Kong's Magic (Part II), Hong Kong Economic Journal, https://wangyujian.hku.hk/?p=3557&lang=en.

Zhu, J. Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. *In Proceedings of the IEEE international conference on computer vision* (pp. 2223-2232), DOI: 10.1109/ICCV.2017.244.

CHROMA

A Human-Centric AR Application Connecting Color Space And Users

ANASTASIA KARASPILIOU[1], ANNA KARAGIANNI[2], DIMITRIS ANDREADAKIS[3], HELENA G. THEODOROPOULOU [4], NIKOLAOS C. SPANOUDAKIS[5], CHRISTOS GEROTHODOROS[6], VASILIKI GEROPANTA[7], and PANAGIOTIS PARTHENIOS[8]
[1,2,3,4,5,6,7,8] *Technical University of Crete, Greece*
[1]*akaraspiliou@tuc.gr,*
[2]*annakaragianni@tuc.gr,*
[3]*dandreadakis@tuc.gr,*
[4]*etheodoropoulou1@tuc.gr,*
[5]*nispanoudakis@tuc.gr,*
[6]*cgerothodoros@tuc.gr*
[7]*vgeropanta@tuc.gr,*
[8]*pparthenios@tuc.gr*

Abstract. In the last few years, digital innovations such as AR, VR and sensing technologies have had a great impact in the sector of cultural heritage, offering new immersive standardized experiences to their visitors. Following this observation, this paper seeks to bring into light the theoretical background and research methodology of 'Chroma', a project that lies at the intersection of theories and empirical observations related to color, architecture, human - centric AR and human behaviour in a monument in Chania, Greece. Based on the hypothesis that color has the ability to alter spatial experience, and that different sound frequencies can intensify this experience, the paper aims at testing AR as a possible technology to study different sensual experiences in the monument, measure them and categorize them according to their emotional and cognitive impact. Thus, it builds on a methodology of work where a vast number of different colors and their combinations integrated in an AR app enables users to generate data at a conscious and subconscious level on a suggested site and becomes ground for further exploration.

Keywords. Mobile AR, Human-centric, Color Interaction, Spatial Perception

1. Introduction

Color has been a science for several hundred years. Newton's discovery that light produces color put an end to much speculation surrounding the nature around it and

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 483-492.

encouraged objective, scientific study. Albers developed an "experimental method for studying and teaching color", a method based on the idea that only by observing color in the field of tension of the environment can one begin to understand the nature of color. "In visual perception a color is almost never seen as it really is - as it physically is. This fact makes color the most relative medium in art. To use color effectively it is necessary to recognise that color deceives continually. To this end, the beginning is not a study of color systems." (Albers, 2013). Colors affect our bodily functions, our mind and our emotions through the energy generated by light. Studies have shown the positive effects of colors in terms of brain development, creativity, productivity, and learning. When color is transmitted from the eye to the brain, the brain releases a hormone that affects emotions, mental clarity, and energy levels. The negative and positive psychological effects of colors can be observed in people depending on the combination in which they are used.

Moreover, sounds affect us physiologically in a very powerful way. Since hearing is our primary sense of warning, a sudden sound will set a process in motion. The effects of sound on users are physiological, psychological, cognitive, and behavioural (Treasure, 2017). Sound transmits well in water. Human bodies are 70% water, thus becoming good conductors of sound. Sound is related to feelings and can cause stress and negativity. As such, sound can be an attractor or a repeller of an experience acting as an enabler of emotions and sounds.

The emergence of new immersive technologies has promoted the design and development of new experiences that enhance the spatial perception of cultural heritage monuments. Specifically, AR technologies enable the juxtaposition of new content on the physical space through a digital layer that contains computer-generated content. Furthermore, AR can become the means to better understand and evaluate an architectural idea for the purpose of intervening or redesigning-reusing an existing space, improving current practices of visualizing the design and conceptual process (Dania et al, 2021). For this reason, the specific project aims at uniting the above mentioned under a research experiment where each factor is studied differently and autonomously and altogether may open up the discussion for human behavior and interaction.

1. Methodology

The experiment of this case study tool place at the old port of Chania, inside the monument of Yali Tzamisi in Chania. This monument is chosen because of its location as a pivotal point in Chania, Crete so that it can be easily visited, making the experiment resemble the idea of a pop-up event. In addition, it is a well-known building, with its own history, which constitutes an experience evoking personalized emotions. A place experienced by diverse groups of visitors, causing them different reactions is suitable for investigating the effect of spatial changes on human reactions. Recognizing its visitability, its materiality, its architecture, it is an ideal location for public experimental data collection based on the design of the present AR project unfolded in the previous section.

For this paper the authors followed a methodology of two phases. The first phase included all the preliminary studies that were necessary for the preparation of the mission, and it consists of the following steps: a) Empirical observations about color, space and human behaviour at the selected monument without the use of technology, b) Literature review on the history, impact, and theories of color in the interior spaces. A number of scholars, books and publications are studied so as to better comprehend the combination of color, psychology and human experience in space. From the study, the concepts of Meerwein, Rodeck, Mahnke (2007) on people's different reactions to color and established universal color concepts that meet the expectations of different user groups as well as of Adams, 2017, in the history and cultural connections of each colors, and of Best (2012) on the impact of color on localities and on the issues surrounding the use of color, from the basic principles of what color is to its important applications in a variety of industries, were highlighted as guiding throughout the experiment process, and c) 3D scanning and modelling of the selected monument and d) Designing an AR app in Unity in two parts. From each step, a number of elements were defined and classified such as the palette of colors that stands as the background tools for the AR app, and the construction of the interactive parts of the AR named, the space - color - and sound.

The second phase refers to an experiment organized by the research team inside the selected monument, during June 2022 for educational reasons to support the diploma thesis of one of the authors on how to use color in an AR app as to offer an alternative experience of the monument. On this occasion, participants used the AR app, and among a sample of 32 people, 13 users wore the neurowave sensor as agents to measure brain activity during the experiment, while all of them replied to pre-defined questionnaires related to their new experience of the space. The group was formed by 19 female and 13 male participants, including: a) students from the Technical University of Crete, School of Architecture, external participants, and other visitors. Some participants were related to architectural science, others from different geographic areas, which reveals the interest and curiosity the program is raising. In the end, the results of the above were further elaborated statistically and assisted in drawing the conclusions of the specific exercise.

2. The Research Project

2.1. COLORS AND SOUNDS AS A SPACE MODIFIER

The theoretical background of this research is based on empirical observations on how we perceive colors, shapes and emotions and what connections might exist between the different human brain parts. It is known that the limbic system is the part of the brain responsible for behavior and emotional responses. For example, the limbic system consists of the hypothalamus, which not only controls emotional reactions but is also involved in regulating body temperature. This fact also explains why red and blue colors can also affect our body temperature (Ho et al., 2014).

Sound also promotes certain states of consciousness. According to Jefferies and Lepp (2012), ultra-gamma frequencies are suitable for extremely intense information processing and concentration. Based on the fact that frequencies from the Ultra-Gamma can be associated with certain emotions and state of consciousness (Jefferies, Lepp, 2012), the team assigned frequencies to colors to render the whole experience more tangible for the user (Table 1). Building on previous research (Karaspiliou, 2022), each color was matched with an ultra-gamma frequency that acts as the intensifier of the experience. Specifically, the frequency of 100 Hz is used as a neutral agent, thus being assigned to the white color, higher frequencies are assigned to yellow and blue rooms, while lower frequencies correspond to red and black rooms. The association of colors, frequencies and sentimental status shapes the hypothesis and becomes the ground for further exploration.

Colors					
Rooms	Red Room	Yellow Room	Blue Room	Black Room	White room
Description	Anger is a reaction to threats or stressors in the environment. Anger emanates from the amygdala, which stimulates the hypothalamus, just like the response to fear.	Happiness refers to a general state of well-being or contentment. Imaging studies suggest that the happiness response originates in part from the limbic cortex.	Relaxation means that the body and mind are free of tension and anxiety. Relaxation is a form of mild ecstasy that emanates from the frontal lobe of the brain, where the posterior cortex sends signals to the frontal cortex via a mild sedative.	Like anger, it helps us respond appropriately to threatening situations that could harm us. This reaction is triggered by stimulation of the amygdala, followed by the hypothalamus. This process produces hormones such as adrenaline and cortisol. When these hormones enter the bloodstream, you may notice some physical changes, such as an increase in heart rate, breathing rate, blood sugar and sweating.	Neutral is a state in which all our emotions are on the same wavelength.
Emotions/Status	Anger, energy, increase pulse	Happiness, energy, increase pulse	Relaxation	Fear	Neutrality
Sound Frequency/ Hz	150 Hz	540 Hz	432 Hz	100 Hz	250 Hz

Table 1: Theory around colors, emotions, and frequencies (Karaspiliou, 2022)

2.2. COLORS AND SOUNDS AS A SPACE MODIFIER

Yali Tzamisi is the only surviving mosque in the city of Chania, that dates to the second half of the 17th century and is a listed archaeological site of Crete. Located at the front of the Venetian Port of Chania, Yali Tzamisi is a cubic building, covered by a large hemispherical dome supported by four elaborate stone arches. On its western and northern sides, it is surrounded by a portico covered by six small domes. The northern portico is used as an information point for the city visitors. For a few years, the monument hosted the Archaeological Museum of Chania, while nowadays, it serves as a venue for exhibitions and venues.

Initially, the experience was designed considering the architectural spatiality of the Mosque. Yali Tzamisi has two main entrances, a north and a western one, both having a sea view. To ensure that all users would follow a predesigned path, the western part was considered as the main entrance, while the northern one was made inaccessible for users. During the second step, the Mosque was 3D scanned and set up into the Unity platform as a point cloud model (Figure 1). The third step included the design and implementation of the immersive part of the application. The western part of the monument was virtually divided into five different rooms, each one 'painted' in one of the basic colors: red, yellow, blue, black, and white (Figure 2). Based on the hypothesis, the experience was enhanced by the frequencies enabled when users entered each virtual color room.

Figure 1. Vuforia scan.

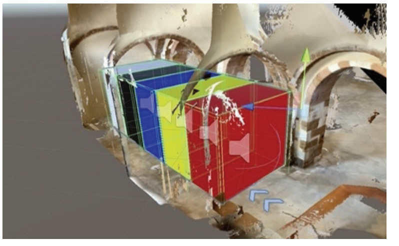


Figure 2. Colored boxes in Unity

At the fourth step, the building was used as a "canvas" with specific predefined color palettes available to the user-visitor. Six predefined sets of colors were empirically created and grouped into color palettes: 1. Van Gogh, 2. The Mediterranean, Islam, Baroque, Bauhaus, Impresionism. The color palettes were created through the collection of a vast number of artifacts related to the specific Art movement or environment combined with the empirical observation on which colors repeatedly appear in these artifacts. Namely, at this stage, the user selects the desired colors and the application projects them on the real walls through a digital layer that implements the AR functionality. At this point, users could move to the center of the main area, either seated or standing, and using an iPad, they could follow instructions on the three consecutive experiments. Through this exploration, the application intended to educate users on three fields: 1) how colors affect space, 2) how colors affect each other and 3) how geometry, space and colors are interrelated. As such, the palettes acted as a connecting tissue of the three experiments. The first experiment reflected how the spatial experience of a certain color is transformed when this color is juxtaposed on different colors (Figure 3, Figure 4).

Figure 3. Second experiment.

Figure 4. Second experiment.

Following the UI instructions, users could first choose a palette, then pick three colors and explore how the first color was perceived when juxtaposed to different colored canvases. Along the same lines, at the second experiment, users would pick four different colors and see the different ways of color combination within the projected AR layer on the interior of the Mosque (Figure 5).

Figure 5 (left). Color Projection on the interior of the Mosque.
Figure 6 (right). Reflective stained-glass dome inspired by "Nuit de Noël", Matisse, 1958.

Each color combination would be projected for fifteen seconds until the end of the experiment. Regarding the last part of the experiment, the research team tried to alter the spatial experience by virtually reconstructing the solid structure of the dome of the building, digitally replacing it with its transparent clone with reflective stained glass inspired by "Nuit de Noël", Matisse, 1958 (Figure 6). The goal was to simulate the sun, its light and its movement during the day by reconstructing the interior of the building showing how the colors and light could change the spatial experience.

2.3. AR TECHNOLOGY & USER INTERFACE

The digital tool in use, was created with the real-time development platform Unity and Vuforia SDK and utilizes Augmented Reality (AR) technology. The application is aimed at mobile devices and the system includes two subsystems / functions, the stereoscopic mode, and the full screen mode. The digital content being used to augment the selected place is color enhanced with sound as described in the previous section.

The combination of the place, technology and content sets the following experience: Initially, in the first stages/ virtual rooms, the user is required to experience full immersion, while tactile interaction with the application is not necessary. Thus, at this stage of the experiment, the subsystem/sublevel of the application's stereoscopic function is implemented. Specifically, the user places the mobile phone in special AR glasses and begins the predetermined passing through the fully colored rooms augmented with the predetermined audio background, following the sequence: red room, blue room, yellow room, black room, and finally white room (Figure 7). The second stage of the experiment requires user interaction with the application's interface and personalized dynamic management of the experiment's content (color selection, viewing angle, target surface). At this stage, the user has the mobile device in her/his hands and the application runs in full screen mode. Therefore, with the pre-designed color palettes at her/his disposal, the user selects colors and digitally transforms the space in real time into an experimentally personalized canvas.

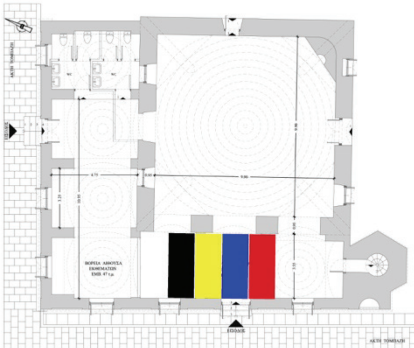


Figure 7. Floorplan of the existing building with the location of the colored boxes.

The overall process aims to collect information based on the user experience, leading to a pool of data and, through its systematic analysis, providing answers regarding the original research question. Data is acquired in two ways: questionnaires for the end users and through a special headset with biosensors (MindWave Mobile 2 Headset) that record the electrical activity of the user's brain during the experiment.

3. Data Collection & Analysis

The first set of data collection was implemented through questionnaires that were answered by 32 users in total, 19 men and 13 women. The first set of questions asked users to describe the feeling that can be most related to the experience of each room. At this point, it seemed that all rooms except for the white one demonstrated a clear connection with a specific feeling or similar feelings. 78% of the users related the red room with stress, while the yellow room was primarily associated with joy (41%) and secondarily with calmness (25%). The blue room was associated with calmness by 63% of the users while the black room was linked to fear by 44% of the users and sadness by 19%. The white room had a diverse range of answers. Specifically, 34% associated the experience with stress, 25% with calmness, and 22% thought it was neutral (Figure 8).

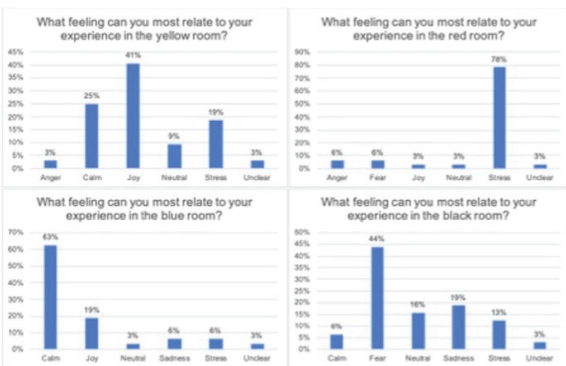


Figure 8. Data resulting from the questionnaires

The second set of questions intended to validate the findings of the first part through a different set of questions. For example, the question "which room scared you" validated that fear is connected to black, as 66% of the users chose the black virtual room. The question "which room gave you positive feelings?" confirmed that blue (38%) and yellow (38%) are associated with positive feelings, while red (3%) and black (3%) are linked to such feelings. As a result, blue and yellow are the colors of the rooms in which guests had the most frequent positive emotions, black and red for the least frequent positive emotions (white was found somewhere in the middle). Again, the darker the rooms, the less often for positive emotions. Black, red, yellow, white is, in (descending) order, the colors of the rooms that cause negative effects from most to least often (the darker the color, the more often it causes negative emotions).

Statistically, by checking the null hypothesis that black rooms do not generate more fear than other-colored rooms and do a one-sided two proportions z-test, this null hypothesis was rejected with a significance level of 0.05 as we have a p-value of 0.045. As a result, the fact that black rooms generate more fear that the other colored rooms combined was proven. Parallel to the data collection through the questionnaires, when visiting the virtual color rooms (third part of the research process), users were asked to wear a headset that recorded their brain activity and depicted values from 0 to 100 for two parameters: attention and meditation (Figure 9). The users' brain activity was recorded and juxtaposed on the screen recording of the mobile AR application (Figure 10&11).

Figure 9. Data extracted through the Headset - Attention / Meditation (axis X = time, axis Y = users), in addition to the amount of time the viewer stayed in each colored room (Red colored boxes is red colored room, blue-colored boxes is blue room etc.)

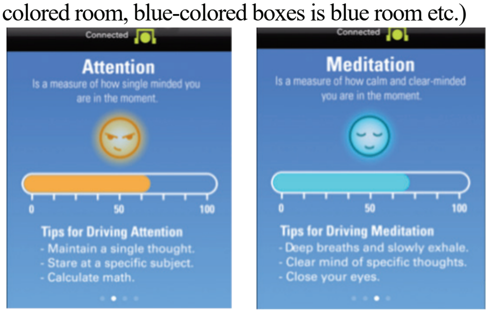


Figure 10&11. Headset Interface – Meditation/Attention

The research team gathered two types of data, based on the AR application: the conscious data extracted through the questionnaires, and the subconscious data, extracted through the headset that recorded the meditation and attention factors. The conscious data clearly showed the correlation of colors with certain feelings, while, according to the subconscious data (based on ANOVA) we could not find a statistically significant difference of the mean values of the two variables (meditation – attention) based on the visited room.

5. Discussion and Further Research

Overall, the current research detected differential evidence on the ways that colors can be associated with brain activity. On the one hand, the data extracted by the questionnaires validated the hypothesis and revealed a clear correlation of color with a specific feeling. On the other hand, the ANOVA of on the measured brain activity demonstrated that the designed immersive experience did not trigger a substantial difference at the average brain activity of each room. The negligible difference of compared meditation and attention averages in each room produces vague results on the correlation of color, sound and meditation or attention levels through biometric data. This finding probably flows from the following facts: a) the experience was designed as a continuous transition from one virtual room to the other, thus encouraging users to stay in every virtual room only for a few seconds, b) the biosensor used has a limited sensitivity and c) the human brain is not affected by different kinds of color in the case that the AR layer contains color embedded in space and not as a form of a specific stimuli. As such, for further exploration and research, the above limitations need to be tested. Moreover, combining the same colors with different frequencies is another field that could be further tested and compared to the acquired data.

In this case, the combination of color and sound are used as a flexible means of shaping space that can give meaning, context, and identity, support architectural form or act subversively against it. In this sense, 'Chroma' is intended to be a pop-up installation, so it can be used in any type of building, indoors or outdoors, in any location, etc. At this strand of the research, the team articulated the initial hypothesis based on empirical findings on the influence of color and sound to the spatial experience. This hypothesis served as the ground for the design and development of the AR application. Through the process of extracting user-generated data, the team was able to collect both conscious and subconscious cognitive data associated with the sensual perception of space. This methodology can be further elaborated to demonstrate the correlation of user-generated data with color, sound, and space, opening up new horizons for the typological classification of user experience assessment based on color and sound.

References

Adams, S., (2017). The Designer's Dictionary of Color. Harry N. Abrams.
Albers, J. (2013). Interaction of color. New Haven: Yale University Press, 978-0300179354.
Best, J., (2012). Colour Design. Elsevier Science.
Dania P., Theodoropoulou, H. G., Karagianni, A., Geropanta, V. and Parthenios, P. (2021). Enhancing User Experience through Interaction Design - Rethinking the municipal agora of Chania through AR narratives. Proceedings of the 39th eCAADe Conference - Volume 2, University of Novi Sad, Novi Sad, Serbia, (8-10 September 2021), pp. 263-272, from, https://doi.org/10.52842/conf.ecaade.2021.2.263.
Ho H., Doorn, G. Takahiro Kawabe, Watanabe, J., Spence, C. (2014). Colour-Temperature Correspondences: When Reactions to Thermal Stimuli Are Influenced by Colour, from https://doi.org/10.1371/journal.pone.0091854
Karaspiliou A. (2022). Psychology of color in the visual arts and architecture, https://doi.org/10.26233/heallink.tuc.91371.
Jefferies K., Lepp, A (2012). An investigation of extraordinary experiences, from https://www.researchgate.net/publication/251534596_An_investigation_of_extraord inary_experiences
Jastreboff PJ (1990). Phantom auditory perception (tinnitus),https://doi.org/10.1016/0168-0102(90)90031-9.
Meerwein, Rodeck, Mahnke, (2007). Color. Springer Verlag Ny.
Sutton, T. and M, B., (2004). The Complete Color Harmony. Rockport Publishers. Treasure J. (2017). How to be Heard: Secrets for Powerful Speaking and Listening. Mango Media.
Wolfgang, J. and Eastlake, D., (1840). Theory of Colours. MIT Press.
Vischer, J. (2008). Towards a user-centered theory of the built environment, Building Research & Information, 36:3, 231 – 240, DOI: 10.1080/09613210801935472.

PARAMETRIC DESIGN AND ANALYSIS OF DYNAMIC LOUVER FOR OPTIMIZED DAYLIGHTING IN HIGH-RISE OFFICE BUILDINGS

ALIREZA JAHANARA[1] and JAE YONG SUK[2]
[1] The University of Texas at San Antonio
[2] University of California, Davis
[1] a.alirezajahanara@gmail.com, 0000-0001-6228-7914
[2] jysuk@ucdavis.edu, 0000-0002-0765-714X

Abstract. This paper presents an innovative approach to optimize daylighting in high-rise office buildings, through parametric analysis of dynamic shading system designs. This study concentrates on a kinetic shading angle optimization plot to determine the optimum angles of the dynamic horizontal louver shading systems at specific times with integrated operationalization equations and parametric performance simulations. Solar irradiance and daylighting were considered as performance metrics in this research, which investigated the integration process using the operationalization method in order to find an optimal rotational angle of dynamic horizontal louver shading systems at the specified time. In this study, dynamic horizontal louver shading systems were positioned in different orientations (Southeast and Southwest) to evaluate the effect of the shading's tilt angles on daylighting. To quantify the daylight quality, maximum and average illuminance were obtained from the raw illuminance on the work plane. At the end, the outcomes of the analyses as well as the optimized angle of the dynamic louver shading were compared to a base case with no shading, and the results prove that dynamic louver with the support of an operationalization method to find optimum angle and testing with parametric performance software leads to optimizing the daylighting performance, enhancing it by approximately 14%.

Keywords. Parametric Design, Dynamic Louver, Solar Irradiance, Daylighting

1. Introduction

Construction façades perform a crucial task in managing various elements of outdoor environments such as temperatures and natural daylight(Chi, Moreno, and Navarro 2018). Temperature and natural daylighting change during the day according to the sun's rotation(Chi, Moreno, and Navarro 2018). Similarly, one of the main factors for an individual's underperformance in a building is visual discomfort in the form of direct sunlight in one's direct perimeter of view(Foroughi, Asadi, and Khazaeli 2021). It is consequently important to use a thorough methodology on building envelopes which revolves around the comfort level of the inhabitants (Eltaweel et al. 2020). It is however very difficult to meet the human comfort needs (daylight plus thermal) while dynamically altering the environmental stimuli to confront kinetic shading systems(Le-Thanh et al. 2021a).

The integration of kinetic louvers with building envelopes is one of the most pioneering factors in climate adaptive construction envelopes(Jayaweera, Rajapaksha, and Manthilake 2021a). A kinetic louver could optimize the solar radiation and internal daylight over the construction envelope concerning the sun's position. They can respond to a variety of disturbances in the sun's radiation(Jayaweera, Rajapaksha, and Manthilake 2021a). Likewise, dynamic louver operation could be based on sensory feedback or based on a predetermined schedule, and it could consist of a manual override type for allowing the residents to take control of the process(Al-Masrani and Al-Obaidi 2019).

This paper seeks to offer a methodology used for horizontal dynamic louver shading to find an optimum rotation angle while accounting for solar irradiance and daylight. With this methodology, it is possible to calculate and find the optimal rotation angle of a horizontal dynamic louver shading that elevates the indoor luminance while getting to an appropriate level of solar irradiance and counting for the effect of daylight optimization on saving energy. For the case study, horizontal dynamic shading was applied to a high-rise building in downtown Houston, TX. The main part of this research utilizes the Grasshopper modeling tool and its plugins for parametric performance evolution, and a horizontal dynamic shading is also simulated parametrically and applied to the office windows. The simulated office rooms with the discussed climate adaptive building envelopes could be rotated to face the two main cardinal directions (Southwest and Southeast) to create various design scenarios.

2. Literature Review

2.1. 2.1 SCENARIOS, ISSUES, AND OBSTACLES TO CONTROL DYNAMIC SHADING

Dynamic shading equipment reacts to the environmental variations in four steps: sensing, data-processing, actuating (Alkhatib et al. 2021). In sensing as well as data processing, a variety of weather stimuli are linked to dynamic shading operations as control variables. They are also used in the actuating procedure, occasionally serving conflicting objectives(Waseef 2017). As a result, there are opposite dynamic shading operation patterns based on the construction location, orientation, and the season(Alkhatib et al. 2021). That is why various occupant-oriented elements and weather stimuli must be studied concurrently to create optimum dynamic shading process developments; however, this method becomes extremely complicated(Le-Thanh et al. 2021b). The efficacy of dynamic shading technology can vary considerably depending on the types and thresholds used in dynamic shading processes and weather stimuli (Rubeis et al. 2019).

3. Methodology

This study incorporates parametric design, parametric performance simulation, and mathematical equations to find an optimized shading device angle. The methodology of this study is split into four major stages and is presented in Figure 1, which shows precise details for daylight optimization of the kinetic shading system model.

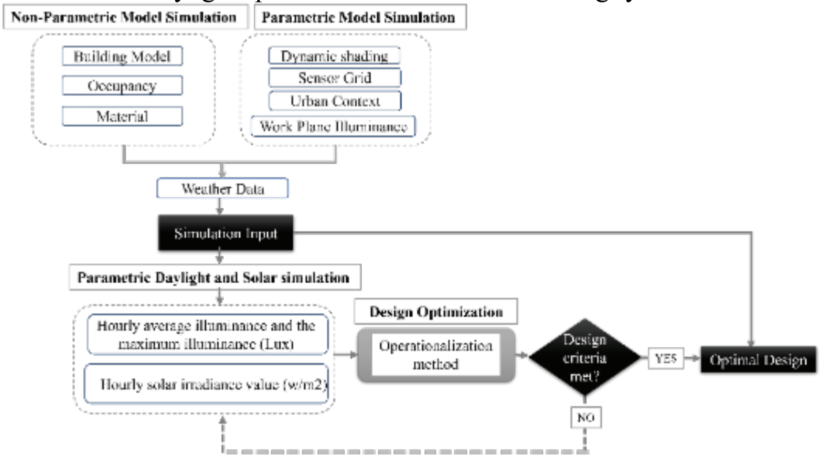


Figure 1: Research framework

Figure 1 illustrates how the research method provides input variables for parametric designs and how to adapt and modify these parameters throughout the proposed simulations.

3.1. CONSTRUCTION, URBAN CONTEXT, AND DYNAMIC SHADING SCENARIO

In the first phase, the building model, dynamic shading, and urban context were divided into two sections: parametric and non-parametric modeling. Grasshopper, a graphical algorithm editor, was selected for parametric modeling and was integrated with Rhino 3D for the non-parametric modeling part of the process. The geometric building layouts of high-rise buildings (e.g. preferred window area and façade dimension), interior configurations (e.g. room size, ceiling thickness), and façade system properties (e.g. thickness, type) were set up as non-parametric variables in Rhino. The real urban context in Houston, Texas was created using the Elk plugin in Grasshopper, which creates street maps and topographies, applying information from Shuttle Radar Topography Mission (SRTM) information from NASA/Jet Propulsion Laboratory and OpenStreetMap.org.

The primary concept of the parametric louver is to rotate around the x-axis in response to the movements of the sun. The dynamic louver is attached to the window without any distance between the window and the shading. It starts at a 40cm distance from the floor level and ends with the last louver attached to the ceiling. It can revolve in a clockwise or anticlockwise path to manage the daylight. The horizontal dynamic louver panel was then set up for this research with circumrotation angles of 0, 15, 30, 45, 60, 75, and 90 degrees (Figure 2).

The dynamic shading system is divided into 9 dynamic louvers, and these louvers cover the whole surface of the window. When the louver is at 0 degrees it means that it is in a fully open position, and when at 90 degrees, it means it is fully closed. Since there is a 13 cm distance between the panels, even when the louver is totally closed, the sun can still penetrate inside the space. The target of 13 cm distance between the panels was set so as not to interrupt the view and close the visual comfort when dynamic shading is totally closed. The depth of the louvres was set at 30 cm, with a thickness of 8 cm and a width of 11.30 m. The horizontal dynamic louver was then applied to the Southwest and Southeast sides of the windows of the simulation model located on the last floor of the high-rise building.

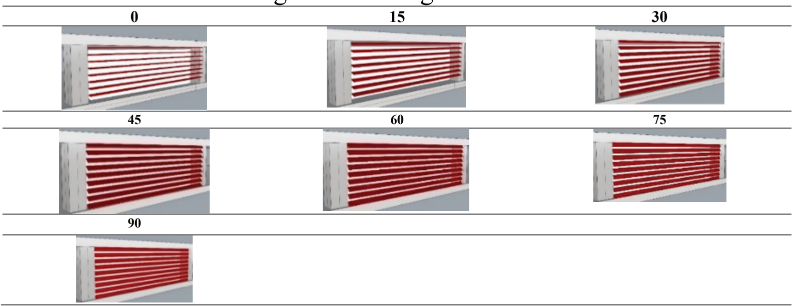


Figure 2: Diagram of dynamic horizontal louver movement at 0, 15, 30, 60, 75, and 90 degrees

3.2. PARAMETRIC PERFORMANCE SIMULATION

In the second stage, two separate parametric performance simulations were used for calculating the performance of dynamic louvers. Daylighting and solar irradiance were simulated using different parametric software packages. Daylighting evaluation as well as hourly solar analyses both took place at 9am, at 12pm and at 3pm on June 21st for two different orientations.

3.2.1. Daylighting Model

The objective of this stage was to evaluate the daylighting performance of the dynamic louver by altering the rotation angle of the dynamic louver on June 21st at different times. To quantify the daylight quality, two principals were obtained from the raw illuminance on the work plane: the maximum illuminance and the average illuminance. The maximum amount of illuminance was set at 2000 lx and an average illuminance of 500 lx was suggested for the work plane. DIVA for Grasshopper with Annual daylight simulation was used to find the hourly maximum illuminance and the hourly average illuminance for all the sensors distributed throughout the indoor space. 138 sensors were positioned through a horizontal measurement grid with 0.5m space between them and were placed 0.76cm above the floor to calculate the horizontal illuminance. To evaluate the number of satisfied sensors based on the rotation angle, a sum of eight simulations involving one with no dynamic shading (base case with glass) were carried out. The material properties for dynamic louvers showed a 4% transmittance. The reflectivity of the interior wall was 50% and the reflectivity of the ceiling was 80%. These parameters are adapted to case-specific studies and are generic. The glazing was double glazing with a visible transmittance of 0.800 and a solar heat gain coefficient (SHGC) of 0.764. Louvers were characterized utilizing the “sheet metal” material created by the DIVA plugin for Grasshopper.

3.2.2. Solar Irradiance

Evaluating solar irradiance has a crucial role in determining the angle of dynamic louvers for daylight optimization. The dynamic louver angle offering reduced solar irradiance is studied as the optimal situation during summer, and the dynamic louver angle receiving increased solar irradiance is the optimal condition during winter. DIVA for Grasshopper is used in this research to create the hourly solar irradiance. 1,800 sensors were positioned on the simulation plane to gather the solar irradiance hourly. The window, with and without dynamic shading, revolving between 0, 15, 30, 45-, 60-, 75-, and 90-degrees angles, was then analyzed.

4. Operationalization Method

The third phase is applying the operationalization process. The aim of this phase is to combine the outcomes of the different performance simulations for solar irradiance and daylighting, which have different units, and to convert all features to be on a comparable scale. To implement the operationalization process (max-min), the initial simulation information should be transformed into a normalized rate scale from 0 to 1 so that the separate simulation outcomes could be compared as an identical measure. To do so, two equations were defined for operationalizing solar irradiance and daylighting with the same target. In the daylight simulation phase, the meaning of the operationalized value of ‘0’ is when the daylit area is the smallest among all the cases, and the operationalized value of ‘1’ shows the maximum performance which has the maximum daylit area.

In the following operationalization equations Figure 3, solar is represented with S and daylight is represented with D. Furthermore, OS represents the original value of solar irradiance and OD represents the original value of daylighting simulation. Max shows the highest value of the original information, and Min shows the lowest value of the original information.

$$D = \frac{OD - Min\ D}{Max\ D - Min\ D} \quad (1) \qquad S = \frac{OS - Min\ S}{Max\ S - Min\ S} \quad (2)$$

Figure 3: Operationalization equations for daylight and solar irradiance

Based on the intensity of the solar irradiance, the definition of an operationalized value can change. For example, the intensity of solar irradiance in July is different than December. In July we need to minimize the solar irradiance in comparison to December when we need to maximize it. Accordingly, the operationalized value for July was defined as -1 to minimize the solar value, and the value of 1 is set for December to maximize the solar value.

4.1. EVALUATION OF THE FRAMEWORK AND THE CASE STUDY

To evaluate the above-mentioned framework for integrated parametric modelling and building performance analyses, a high-rise building located in downtown Houston was chosen as the case study. The selected building, CenterPoint Energy Plaza (formerly

Houston Industries Plaza), is surrounded by many high-rise buildings. This building has a height of 226.00 m (741 ft.), a length of 55.07 m (180.68 ft) and a width of 49.27 m (161.65 ft). The building was then modelled in Rhino and Grasshopper with 47 floors.

The surrounding buildings were modelled with Elk2 Grasshopper and were included in the site model to realize the impact of the shadow of the urban context in the daylight optimization process of the dynamic façade. The Ladybug Radiation Analysis Tool uses the sun position for every hour of the year to determine how much solar radiation the exterior surfaces receive.

4.2. CASE STUDY ORIENTATION

For this case study, each side of the building was divided into four office rooms. We chose two offices, one located in the Southeast, and the other one in the Southwest. The simulation was conducted on the 47th floor of the high-rise building, in the south façade. The reason for choosing Floor 47 was based on the results of Ladybug, which indicated that as this floor has four open sides, the sun penetrates deeply inside it while the neighboring buildings have no effect on this floor by their shadows. Each of the simulated offices is 12 m deep, 11.20 m wide, and 3.7 m high.The glass is 3.30 m high and 10.80 m wide.

4.3. RESULT OF DAYLIGHT SIMULATION

The parametric daylight performance analysis performed in "DIVA" for Grasshopper generated results for the indoor daylighting requirements, and the visual comfort of the occupants was evaluated based on the amount and quality of the daylight. In Tables 1 and 2, the result of the hourly daylight simulation is shown for the base case of a fully glazed window, with and without a dynamic louver shading. The optimization target is an average illuminance of 500 lx and a maximum of 2000 lx.

In the southeast orientation, with a dynamic louver at an angle of 0° to 15° degrees, most of the area has an acceptable illuminance starting at 9 a.m. except for around 3 p.m. when the illuminance level goes over 3000 lx. If the angle is set to 30° degrees, the illuminance gradually reaches its optimal level at 9 a.m. Compared to all other angles, an angle of 45° denotes the illuminance is at its optimal level at 9 in the morning. Angles of 60° and 75° decrease the amount of illuminance in the indoor space from 12 to 3 p.m. and do not improve the daylit area. At 90 degrees where the louver totally closes, the 13 cm distance between louver panels, helps the indoor space to have a better luminance level.

At a 0° angle in the southwest orientation, from 9 a.m. to 12 p.m., most area has over 3000 lx illuminance level. At 9 a.m., with angles of 15° and 30°, most of the area was still over 2500 lx. In both orientations, angles of 0° and 75° were proven to be ineffective angles for the daylight requirement of workspaces.

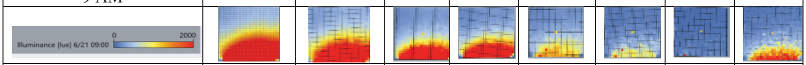
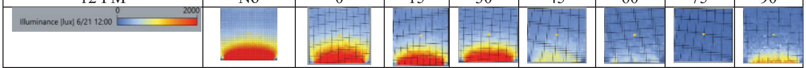
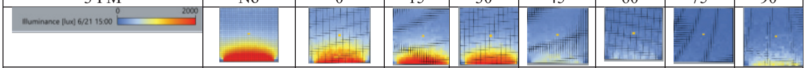
9 AM	No louver	0	15	30	45	60	75	90
								
Maximum illuminance	3620	3016	2592	2542	2011	2308	1630	2540
Average illuminance	4828	1465	1198	1137	503	337	65	463
12 PM	No	0	15	30	45	60	75	90
								
Maximum illuminance	6132	3348	3360	2024	928	832	176	1472
Average illuminance	1150	868	721	502	334	194	43	264
3 PM	No	0	15	30	45	60	75	90
								
Maximum illuminance	3696	2120	2236	2048	748	620	152	1024
Average illuminance	845	657	556	505	281	162	36	190

Table 1: Daylight performance simulation result on June 21st for southeast

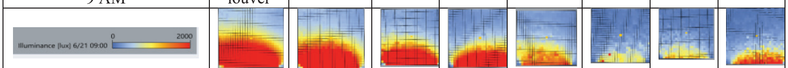
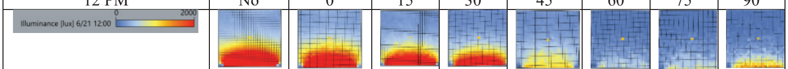
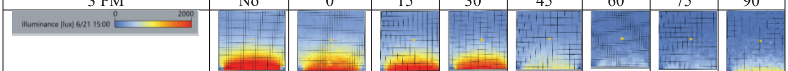
9 AM	No louver	0	15	30	45	60	75	90
								
Maximum illuminance	3748	2849	2801	2724	2018	1236	1432	2080
Average illuminance	5435	2288	1423	1264	522	313	333	1253
12 PM	No	0	15	30	45	60	75	90
								
Maximum illuminance	8208	4292	5080	3336	1688	924	864	2014
Average illuminance	1435	1010	936	779	440	207	202	501
3 PM	No	0	15	30	45	60	75	90
								
Maximum illuminance	3496	2048	2424	2012	1021	664	548	1100
Average illuminance	819	636	587	518	329	153	141	182

Table 2: Daylight performance simulation result on June 21st for southwest

4.4. RESULT OF SOLAR IRRADIANCE

The Solar radiation simulation in this study is presented through hourly analyses. In Tables 3 and 4, solar irradiance is demonstrated at both orientations with different intensity levels in the surface of the glass. The variance is related to the dynamic horizontal shading rotation angle of the louver. Tables 3 and 4indicate that solar irradiance is reduced by the movement of a dynamic louver.

For the purposes of the current study, evaluations were performed in the month of June. As a result, the angle that produces the minimum solar irradiance value should be considered for a city like Houston, Texas. For example, 30° is the optimum rotation position for kinetic louver in the Southeast orientation, at 12pm on June 21st , to produce the lowest solar irradiance (Table 3). On the other hand, in the Southwest orientation at 12pm, the angle of 90° is considered the optimum angle (Table 4).

Table 3: Solar irradiance simulation result on June 21st for southeast

Solar Irradiance	**Point in time**	**No louver**	**0**	**15**	**30**	**45**	**60**	**75**	**90**
W/m2	9 AM	435.12	175.98	164	149.42	68.80	123.90	117.88	132.82
	12 PM	189.016	174.92	173.01	162.56	173.46	174.37	177.38	162.59
	3 PM	150	150.68	147.912	146.44	144.54	146.55	161.92	146.47

Table4: Solar irradiance simulation result on June 21st for southwest

Solar Irradiance	**Point in time**	**No louver**	**0**	**15**	**30**	**45**	**60**	**75**	**90**
W/m2	9 AM	518	192.3	182.2	163.1	141.2	142.8	142.4	143.9
	12 PM	301	187.9	183.1	178.4	176.0	174.3	173.0	144.6
	3 PM	135	153.3	150.2	145.8	146.3	148.6	159.2	160.1

4.5. CASE STUDY OPERATIONALIZATION METHOD

The environmental factors (solar irradiance and daylighting) were also employed in the operationalization process. According to the equations established Figure 3 and the explanation in section 3.3, since the simulation was performed in June, the operationalized value of the solar irradiance was defined between -1 and 0 to minimize the solar irradiance during summer.

	0		**15**		**30**		**45**		**60**		**75**		**90**	
	OD	OS	OD	OS	OD	OS	OD	OS	OD	OS	OD	OS	OD	OS
9 AM	3016	175.98	2592	164	2542	149.42	2011	68.80	2308	123.90	1630	117.88	2540	132.82
Operationalization	1	-1	0.69	-0.89	0.66	-0.75	0.27	0	0.49	-0.51	0	-0.46	0.66	-0.60
Result	0		-0.2		-0.09		0.27		-0.02		-0.46		0.6	

	0		**15**		**30**		**45**		**60**		**75**		**90**	
	OD	OS	OD	OS	OD	OS	OD	OS	OD	OS	OD	OS	OD	OS
12 PM	3348	174.92	3360	173.01	2024	162.56	928	173.46	832	174.37	176	177.38	1472	161.20
Operationalization	1	-0.83	1	-0.71	0.58	0	0.24	-0.74	0.21	-0.80	0	-1	0.41	-0.09
Result	0.17		0.29		0.58		0.5		-0.59		-1		-0.32	

	0		**15**		**30**		**45**		**60**		**75**		**90**	
	OD	OS	OD	OS	OD	OS	OD	OS	OD	OS	OD	OS	OD	OS
3 PM	2120	150.68	2236	147.91	2048	146.44	748	144.54	620	146.55	152	161.92	1024	146.47
Operationalization	0.94	-0.27	1	-0.10	0.91	0	0.29	-0.12	0.22	-0.01	0	-1	0.42	0
Result	0.67		0.9		0.91		0.17		0.21		-1		0.42	

Table 5: Operationalization for solar irradiance and daylighting result on June 21st for southeast

	0		15		30		45		60		75		90	
	OD	OS	OD	OS	OD	OS	OD	OS	OD	OS	OD	OS	OD	OS
9 AM	2849	192.3	2801	182.2	2724	163.1	2018	141.2	1236	141.8	1432	142.4	2180	143.9
Operationalization	1	-1	0.97	-0.80	0.92	-0.43	0.48	0	0	-0.01	0.12	-0.02	0.52	-0.05
Result	0		0.17		0.49		0.48		-0.01		0.1		0.47	
	0		**15**		**30**		**45**		**60**		**75**		**90**	
	OD	OS	OD	OS	OD	OS	OD	OS	OD	OS	OD	OS	OD	OS
12 PM	4292	187.9	5080	183.1	3336	178.4	1688	176.0	924	174.3	864	173.0	2014	144.6
Operationalization	0.81	-1	1	-0.89	0.59	-0.78	0.20	-0.73	0.01	-0.69	0	-0.66	0.27	0
Result	-0.19		0.11		-0.19		-0.53		-0.68		-0.66		0.27	
	0		**15**		**30**		**45**		**60**		**75**		**90**	
	OD	OS	OD	OS	OD	OS	OD	OS	OD	OS	OD	OS	OD	OS
3 PM	2048	153.3	2424	150.2	2012	145.8	1021	146.3	664	148.6	548	159.2	1100	160.1
Operationalization	0.80	-0.52	1	-0.31	0.78	0	0.25	-0.03	0.06	-0.20	0	-0.94	0.29	-1
Result	0.28		0.69		0.78		0.22		-0.14		-0.94		-0.71	

Table 6: Daylight performance simulation result on June 21st for southwest

Simulation performance values are calculated at every given time and according to the rotation of the dynamic louver angle. The calculation intended for operationalization is obtained by initially subtracting the lowest value from the variable which is going to be normalized. Then the lowest value is subtracted from the highest value, and the previous outcome is split by the latter.

Tables 5 and 6 indicate the optimal angles for each case in the southeast and southwest orientations. Each optimal result was determined through the operationalization process. For instance, after applying the operationalization method for the southeast orientation at 9 a.m., the most effective rotational angle the solar irradiance performance, while controlling for the daylight condition, was at 45 degrees.

5. Discussion and Conclusion

To enhance the daylight optimization in a building, a dynamic horizontal shading system is an attractive option. Since this option is very hard to be optimized, creating a new process to assess the implementation of a dynamic horizontal shading is essential to gather the relevant data . This study then aims to integrate the parametric design, the parametric environmental software, and the mathematical equation that assists in the development of a dynamic louver design and its dynamic movement controls. The analyzed metrics in this methodology were the solar irradiance and daylighting The results of evaluating dynamic louver can then guide designers towards informed solutions.

This research has investigated the optimization process of two different metrics for solar irradiance and daylighting, utilizing the operationalization procedure to achieve an optimum rotational dynamic horizontal shading angle at any given time. It also aimed for decreasing the solar gains on the glazing of high-rise buildings while providing a specific illuminance level on work plane. In addition, the results of hourly solar irradiance and hourly maximum luminance were utilized as operationalized values to get to the final simulation outcomes. At the end, based on the outcomes of this study, it was determined that the rotational motion of a dynamic horizontal shading system has a significant effect on the daylight and solar irradiance conditions.

References

Ahmad, E.,Su, Y., Lv, Q., and Lv, H., 2020. “Advanced Parametric Louver Systems with Bi-Axis and Two-Layer Designs for an Extensive Daylighting Coverage in a Deep-Plan Office Room.” Solar Energy 206(June): 596–613. https://doi.org/10.1016/j.solener.2020.06.035

Alkhatib, H., P. Lemarchand, B. Norton, and D. T.J. O’Sullivan. 2021. “Deployment and Control of Adaptive Building Facades for Energy Generation, Thermal Insulation, Ventilation and Daylighting: A Review.” Applied Thermal Engineering 185(November): 116331. https://doi.org/10.1016/j.applthermaleng.2020.116331.

Al-Masrani, S. M., and Al-Obaidi, K. M., 2019. “Dynamic Shading Systems: A Review of Design Parameters, Platforms and Evaluation Strategies.” Automation in Construction 102(March): 195–216. https://doi.org/10.1016/j.autcon.2019.01.014.

Chi, Doris A, Moreno, D., Navarro, J., 2018. “Correlating Daylight Availability Metric with Lighting, Heating and Cooling Energy Consumptions.” Building and Environment 132: 170–80, DOI:10.1061/(ASCE)AE.1943-5568.0000339

Foroughi, R., S Asadi, and Khazaeli, S., 2021. “On the Optimization of Energy Efficient Fenestration for Small Commercial Buildings in the United States.” Journal of Cleaner Production 283: 124604. https://doi.org/10.1016/j.jclepro.2020.124604

Jayaweera, N., Rajapaksha, U., and Manthilake, I., 2021. “A Parametric Approach to Optimize Solar Access for Energy Efficiency in High-Rise Residential Buildings in Dense Urban Tropics.” Solar Energy 220: 187–203,https://doi.org/10.1016/j.solener.2021.02.054

Luan, L., et al. 2021. “Optimal Design of an Origami-Inspired Kinetic Façade by Balancing Composite Motion Optimization for Improving Daylight Performance and Energy Efficiency.” Energy 219: 119557. https://doi.org/10.1016/j.energy.2020.119557

Luan, L., et al. 2021. “Optimal Design of an Origami-Inspired Kinetic Façade by Balancing Composite Motion Optimization for Improving Daylight Performance and Energy Efficiency.” Energy 219: 119557. https://doi.org/10.1016/j.energy.2020.119557.

Rubeis, T. et al. 2019. “A Novel Method for Daylight Harvesting Optimization Based on Lighting Simulation and Data- Driven Optimal Control.” 16 th IBPSA International Conference and Exhibition: 1036–43. DOI:10.26868/25222708.2019.210494

Waseef, A., 2017. “Towards a New Classification for Responsive Kinetic Facades.” Proceedings of the Memaryat International Conference “MIC 2017” (April).

Urban Analytics, Urban Modelling and Simulation

CONVOLUTIONAL NEURAL NETWORK (CNN) SUPPORTED URBAN DESIGN TO REDUCE PARTICLE AIR POLLUTANT CONCENTRATIONS

ZISHEN BAI[1] and CHENGZI PENG[2]
[1,2]*School of Architecture, The University of Sheffield.*
[1]*zbai5@sheffield.ac.uk, 0000-0002-8098-4729*
[2]*c.peng@sheffield.ac.uk, 0000-0001-8199-0955*

Abstract. PM2.5 has become a significant factor contributing to the haze outbreak in mainland China, which has negative impacts for public health. The current agility of CFD-based modelling to reveal in real-time the changes in PM2.5 concentrations in response to (proposed) changes in urban form limits its practical applications in the design processes. To support urban design for better air quality (AQ), this study presents a machine learning approach to test: (1) that the spatial distribution of PM2.5 concentrations measured in an urban area reflects the area's capacity to disperse particle air pollution; (2) that the PM2.5 concentration measurements can be linked to certain urban form attributes of that area. A Convolutional Neural Network algorithm called Residual Neural Network (ResNet) was trained and tested using the ChinaHighPM2.5 and urban form datasets. The result is a ResNet-AQ predictor for the city centre area in Beijing which had one of the highest air pollution levels within the Beijing-Tianjin-Hebei region. The urban area covered by the ResNet-AQ predictor contains 4,000 grid cells (approx. 25.3 km x 25.3 km), of which 1,200 (30%) cells were selected randomly for testing. The ResNet-AQ prediction accuracy achieved 87.3% after 100 iterations. An end-use scenario is presented to show how a social housing project can be supported by the AQ predictor to achieve better urban air quality performance.

Keywords. PM2.5, Urban Form Indicators, Image Classification, Convolutional Neural Network, Open Urban Data

1. Introduction

In recent years, the detrimental effects of air pollution on human health have attracted attention (Abdalla and Peng, 2021). China is ranked the most polluted country for air pollution among 175 countries globally (China MEE, 2020). PM2.5 consists of a high concentration of toxic and hazardous substances, which can be transported over long distances and have a long residence time in the atmosphere. PM2.5 significantly impacts atmospheric environmental quality and public health (Wang, 2017a).

Recent research on Urban Air Quality has focused on two areas. (1) Studies of

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 505-514.

traditional end-of-pipe treatments to reduce air pollutant emissions (Li and Zhou, 2019). A typical strategy was to reduce vehicle emissions by controlling urban road layouts (Rodríguez et al., 2016; Yuan et al., 2018). (2) Studies on reducing the pollutant stock in the air. A common approach was to improve air pollutant dispersion and deposition (Wang et al., 2017b). The effect of urban form on wind speed has been investigated via computational fluid dynamics (CFD), which further estimated the influence of wind speed on pollutant dispersion (Jia et al., 2021). Zhai et al. explored the effects of vegetation coverage on particle sedimentation based on computational fluid dynamics simulations (Zhai et al., 2022). Conventional CFD modelling requires detailed descriptions of the particle and fluid fields. CFD models may perform poorly due to inaccurate settings of urban form parameters and meteorological conditions (Jurado et al., 2021). Based on CFD outputs, a deep-learning neural network to predict air pollution was developed by Jurado et al. (2022) recognising that the CFD models were based on assumed parameters and could not fully reflect the real-world complexity. Due to the high computational cost, CFD remains problematic for architects and urban designers to evaluate urban air quality performance of design proposals in an interactive manner.

In this paper, we present a study of a modelling framework for urban air quality prediction to aid urban design based on satellite sensing data. In particular, near-surface PM2.5 concentrations derived from the Aerosol Optical Depth datasets have been made widely available by various agencies. Here, the main idea was to model the relationship between satellite imaging of PM2.5 concentrations and urban form characteristics at a city scale through Convolutional Neural Network (CNN) learning. A set of nine urban form indicators for characterising urban form was identified in our literature review, including Road Surface Density, Road Network Density, Natural Surface Ratio, Terrain Elevation, Building Density, Building Length, Building Volume, Average Building Height, and Building Height Difference. Both PM2.5 concentrations and urban form characteristics were pre-processed for CNN training and testing. Thus, urban design proposals drawn as master plans can be evaluated rapidly by the resultant CNN model to reveal the likely changes in urban air quality performance (i.e., PM2.5 concentrations) due to the proposed designs (changes).

2. Methods and materials

Figure 1 shows an overview of the workflow. The proposed modelling framework consists of three main parts: (1) Urban form data; (2) Urban PM2.5 concentration data; and (3) Training and testing Machine Learning algorithms. A major challenge was the acquisition and pre-processing of the urban form and urban air quality data required for the machine learning modelling. The quality (accuracy and resolution) of the datasets will determine the efficacy of the resultant model trained and tested as a location based urban air quality predictor, given the urban characteristics as inputs.

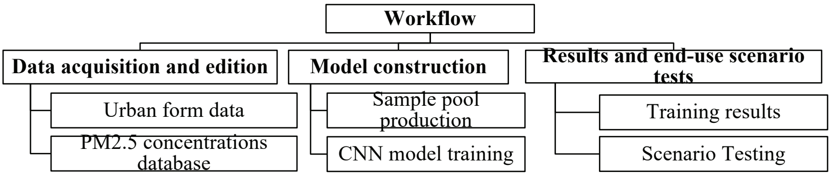


Figure 1. Workflow overview

2.1. URBAN FORM DATASET

A city map containing vector information is required to compute the nine urban form indicators. We used Open Street Map to obtain the urban form data in the ESRI Shapefile format (Marsudi, 2019). However, open street map does not provide valid building height information (Bernard et al., 2022). An alternative geo-data source, Amap (amap.com), one of the most popular web mapping service providers in China (Sun et al., 2020), provided building height data. A transformer application, Bigemap, was used in the study to acquire and transform the initial data from OSM. Furthermore, the NCDDC (National Coastal Data Development Center, https://www.ncddc.noaa.gov/) provides territorial elevation data, which can be input into ArcMap (version 10.8.1) directly as vector data.

Urban vector maps require the definition of the Geographic Coordinate System and Projected Coordinate System in ArcMap to make all processing work occur in the same coordinate system (Jekeli, 2006). In this study, the Chinese Geodetic Coordinate System 2000 was used, and the Gauss-Krüger projection method, currently expected in China, was used as the projected coordinate system (Yang, 2009).

The purpose of the urban form data editing is to generate a large set of training samples for image classification. Nine options of the grid resolution were examined for the study based on the fishnet function in ArcMap (Figure 2). While the low sample resolution (2 km x 2 km) can accommodate a wider urban area, the resultant sample size cannot be sufficient for machine learning. In contrast, the high-resolution sampling (50 m x 50 m) can produce the sample pool of a satisfactory size, but the small spatial coverage is inappropriate to the urban form indicators. Based on the assessment of the nine grid resolutions (Figure 2), the 400 m x 400 m resolution was chosen in this study to produce the final sample pool (N=4,000) for machine learning.

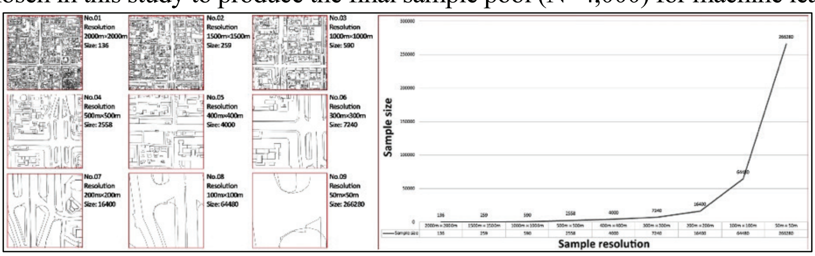


Figure 2. Urban grid size and resolution test

The next stage of the urban form data preparation involved colour labelling in

ArcMap, including the three vector layers labelled with different colours for defining urban form. Figure 3 shows the labelled sample images based on the Hexcone model in the RGB colour system. The red colour field shows the building height based on the three-metre gradient labels in the study area. Roads are in white colour, and natural surfaces are in green, reflecting the road surface density, road network density and natural surface density of the study area. Furthermore, the territorial height data provided by the NCDDC was labelled with a three-metre gradient of grey-scale shifts, which shows urban terrain changes in relative height.

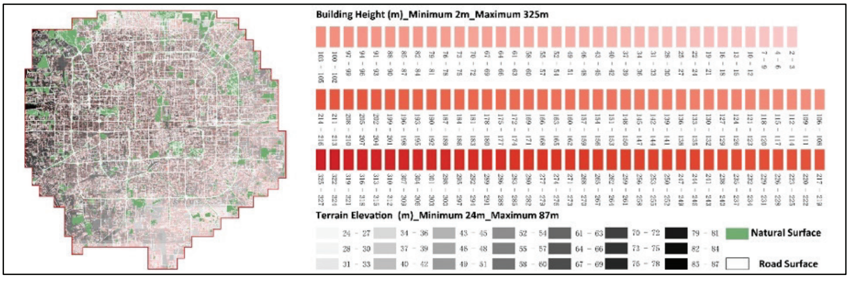


Figure 3. The urban form map of the study area tagged in ArcMap (v.10.8.1)

2.2. PM2.5 CONCENTRATION DATASET

Wei et al. (2021) employed the Moderate Resolution Imaging Spectroradiometer Multi-Angle implementation of Atmospheric Correction algorithm to estimate near-surface PM2.5 concentrations at a 1 km resolution. The output is the high resolution PM2.5 dataset for China (ChinaHighPM2.5), which was accessed for the Beijing central urban area in this study.

The image classification model (see Section 2.3) was trained on the urban form data samples (N=4,000), and each urban form sample (cell) was assigned a PM2.5 concentration value according to ChinaHighPM2.5. The assignment requires resampling where urban form grid cells crossing multiple PM2.5 concentration boundaries. Figure 4 shows the resampling method based on weighted average sum to ensure consistency in assigning PM2.5 concentrations to the urban from grid cells.

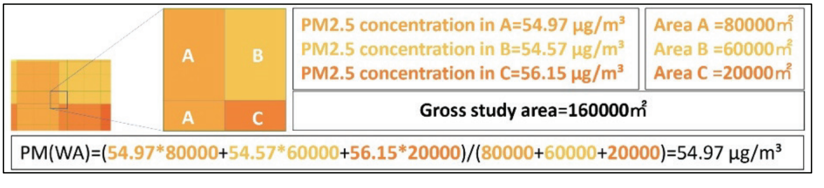


Figure 4. A weighted average algorithm for samples which cross concentration boundaries

The air quality monitoring network of the China Environmental Monitoring Center, CnOpenData, and Qingyue Open Data Center of Environment jointly provide the download service for accessing the ground based air quality monitoring data in China (He et al., 2021). There are 15 monitoring stations within the study area, providing monthly PM2.5 data for six years (2015-2020). The edited PM2.5 dataset prepared for the study area was then evaluated with three statistical measurements:

Determination (R^2), Root Mean Square Error (RMSE), and Mean Absolute Error (MAE). The sample size (N) is 90, and the statistical check gives R^2=0.96, RMSE=4.01 μg/m³ and MAE=2.98 μg/m³. The performance of the reassigned PM2.5 data is close to the source data provided by ChinaHighPM2.5 with R^2= 0.94, RMSE=5.07 μg/m³, and MAE=3.72 μg/m³ (Wei et al., 2021).

The training samples were further assessed the correlation of the PM2.5 with the urban form to identify the strongest correlation dataset. It is likely that the urban from samples and PM2.5 concentrations bear no correlation because of the instability of geospatial datasets and data latency (Zufle et al., 2020). A regression analysis was applied to determine the regression coefficient R^2 between urban form indicators and the annual mean PM2.5 concentration values from 2015 to 2020. Figure 5 shows the regression coefficients of PM2.5 concentrations and nine urban form indicators used in the study during the six-year period (2015-2020). The regression coefficient for the PM2.5 data in 2019 was the highest (R^2=0.30). As the study aims to achieve more sensitive classifiers in the image classification, the PM2.5 range is important, and the degree of fluctuations in the winter PM2.5 concentrations in 2019 shows a highest regression coefficient (R^2=0.37). Winter heating in northern China consume large quantities of fossil fuels and significantly increase PM2.5 concentrations (Liang et al., 2015). Thus, the winter months of January, February, November and December in 2019 were chosen for image classification training.

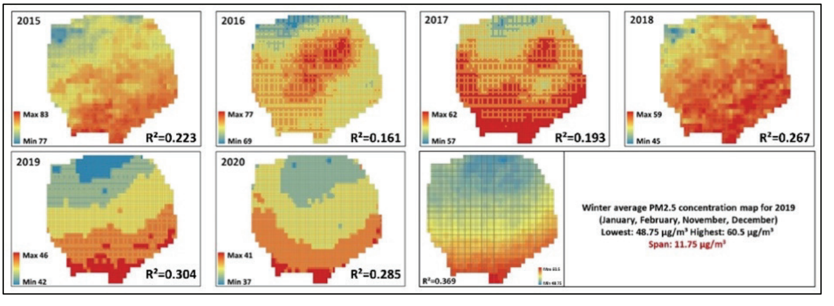


Figure 5. Regression coefficients of PM2.5 concentrations and urban form indicators in the study area for six years from 2015 to 2020

2.3. IMAGE CLASSIFICATION MODEL CONSTRUCTION

The study adopts a supervised classification learning approach to classify the input images (i.e., the urban form characterisation) using spectral features obtained from the training samples (Das, 2017). The urban form samples were applied for supervised classification training, and the category was based on the PM2.5 values of winter season 2019.

For efficient image printing of the samples, a simplified Python program was written (https://github.com/ZishenBai/GIS.git/). The program allows the printout to be saved based on a range of PM2.5 concentration levels. The advantage of the Python coding is efficient production and sample outputs and the program code is easily adjustable and reusable. The core import module for batch printing and classification saving is the Search-Cursor in ArcMap 10.8. 1. Search-Cursor is typically used to create read-only cursors on element classes or tables and allows the use of where clause or field to restrict the query and sort results.

The image classification models were constructed and trained with the support of an online open resource no-code AI platform, EasyDL (https://ai.baidu.com/easydl/). EasyDL is an AI service development platform provided by Baidu Brain with deployable machine learning and deep learning algorithms (Singh, 2016). The main algorithm adopted for this study is the Deep Residual Learning algorithm, coded as ResNets on the EasyDL platform. ResNets has a strong track record, having won the best performance award at the ILSVRC 2015 computer vision competition (He et al., 2015). Due to the limitation of the sample pool size, an incremental training strategy was used in this study to improve the CNN model generalisation capability. In the urban form samples, a system of colour gradients was applied to different urban form indicators, meaning that warping or colour changes could potentially mislead the training models. An incremental training strategy with randomly occurring XY translation (TranslateX; TranslateY) with a 10% probability was selected for this study (Figure 6).

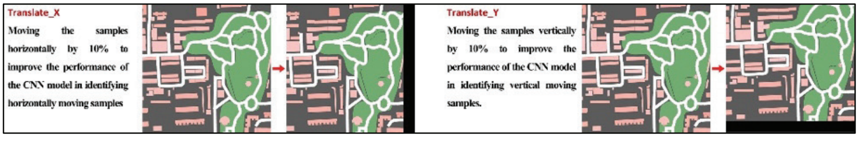


Figure 6. An incremental training strategy, based on a 10% XY translation

The final step was configuring the training environment based on the Baidu Public Cloud deployment service which provides the inference and prediction capabilities following the REST API style. The basic parameters of the model training server were Tesla P40, Video Mem: 24GB, CPU: 12 Cores, RAM: 40G.

3. Results and discussion

3.1. TRAINING RESULTS

In machine learning, image classification training generally involves optimising multiple training strategies, which can significantly affect the training results. In this study, all tests were based on the same training algorithm and test sample pool. The image classification model (function) was trained based on the supervised learning approach of the CNN framework using the ResNets algorithm. Table 1 shows the test results based on PM2.5 concentrations equally divided into five categories (levels), achieving an overall accuracy of 59.20%. The test report showed 93 samples confounded in Level_02 and Level_03, while only 7 were confounded in Level_02 and Level_05. This suggests that PM2.5 concentration levels cannot establish a significant relationship with the samples. Thus, the levels were reduced to four.

Table 1. Misclassification list for the test training

Label	Top 5 misidentified label size				Accuracy	Size
level_04	[79] level_05	[85] level_03	[46] level_02	[31] level_01	53.80%	260

level_02	[93] level_03	[78] level_01	[46] level_04	[07] level_05	52.50%	237
level_03	[93] level_02	[85] level_04	[35] level_01	[15] level_05	57.10%	264
level_01	[78] level_02	[35] level_03	[24] level_05	[31] level_04	66.70%	249
level_05	[79] level_04	[24] level_01	[15] level_03	[07] level_02	65.80%	185

Based on the test training results, the final model classification was defined as four categories according to four ranges of PM2.5 concentrations (Table 2). For each classification category (PM2.5 concentration level), 30% of the samples were randomly selected as the test set. After removing invalid samples, 3,800 samples (out of 4,000) were used, including 2,660 for training and 1,140 for testing. With 100 iterations of training (runtime: 5 hours and 16 minutes), the model achieved an accuracy of 87.30%.

Table 2. Classification results checklist

Category	Level_01	Level_02	Level_03	Level_04
PM2.5 concentrations	48.75μg/m³-52.00μg/m³	52.01μg/m³-54.00μg/m³	54.01μg/m³-57.00μg/m³	57.01μg/m³-60.50μg/m³

3.2. AN END-USE SCENARIO

To illustrate how the Beijing ResNets-AQ Predictor as trained may be applied in urban design practice, the Baiziwan social housing project, located on the fourth ring road in the Beijing city centre (39.89 ° N, 116.51 ° E), was evaluated as an end-use test scenario. The affordable housing project was completed by MAD Studio in 2021 (Iype, 2021). The ResNets-AQ Predictor was applied to compare the PM2.5 concentration levels of the urban area before and after the Baiziwan social housing project. Figure 7 shows the outcome of the testing. It shows that the PM2.5 concentration of the site before the social housing project was estimated at Level_03 (probability of 56.06%) and Level_04 (41.07%), indicating the winter PM2.5 concentrations would be in the range of 54.01μg/m³ - 60.50μg/m³ (probability of 97.13%). The PM2.5 concentration levels after the Baiziwan Social Housing Project was estimated at Level_02 (64.38%) and Level_03 (30.90%), indicating the winter PM2.5 concentrations would be in the range of 52.01μg/m³ - 57.00μg/m³ (probability of 95.28%). The PM2.5 concentration values and predicted probabilities for the top two levels of the test results were used to assess the effect of the social housing project in terms of PM2.5 concentrations through a weighted average estimation. The results show that the PM2.5 concentration at the site was 55.24 μg/m³ (before) and 51.27 μg/m³ (after), indicating that the social housing project leads to about 7.19% reduction in PM2.5 concentrations. As such, this end-use test scenario shows an example of how the ResNets-AQ Predictor can be applied to generate rapid air quality assessments of urban design proposals without going through site-specific CFD modelling and simulation. The runtime invested in the ResNets-AQ predictor training and testing delivers the agility required of interactive design processes.

Figure 7. Assessment results (a) Site before regeneration, (b) Baiziwan social housing project

The numerical predictions can be complemented by adding links to related air quality literature to further improve the usability of the ResNets-AQ Predictor for rapid evaluation of existing urban areas as well as proposed interventions. For this test use scenario, firstly, the social housing project has significantly reduced the building density on the site, decreasing wind obstruction (Yang et al., 2020). The new buildings were designed with reduced windward sides and increased distances between buildings, which potentially increased the wind speed within the site and accelerated PM2.5 dispersion rates (Yang et al., 2020). Furthermore, the project has significantly increased natural surface areas, raising the sedimentation impact of vegetation on PM2.5 (Wang et al., 2014; Cheng et al., 2015). Vertical landscaping can reduce dust pollution due to increased wind speeds by high-rise buildings. However, the project has increased road density, potentially increasing PM2.5 concentrations. There is a synergistic effect between the road surface density and road network density indicators, and generally, a wider road network with lower density has the potential to decrease PM 2.5 concentrations (Wang et al., 2017b). In conclusion, the renewed urban form with reduced PM2.5 concentrations is primarily attributable to the change in wind speed and the increase in vegetation area. Thus, the ResNets-AQ predictive evaluation was considered to be in an agreement with an observation of the urban form features introduced by the Baiziwan social housing project.

4. Conclusion and further work

The study presents an urban air quality modelling tool built on the CNN-based image classification technique. The main purpose of the tool is to enable rapid assessment of urban form performance in terms of PM2.5 concentrations during high ambient air pollution months. A contribution of the study is the development of a novel modelling framework that creates links between satellite aerosol sensing data and urban form characterisation at an urban scale. Unlike the conventional CFD approach, the data-driven deep learning method can provide the agility required in rapid interaction with design decision-making. It can be anticipated that designers will be able to interact with a ResNets-AQ evaluation/prediction platform via mobile devices and receive likely PM2.5 concentration levels of their design proposals.

Following this study, there are a number of areas to be further investigated. Firstly, a ResNets-AQ predictor is location based, that is, the Beijing ResNets-AQ predictor can only be applied to urban areas/projects in Beijing. However, the modelling methods and techniques are expected to be applicable to other cities.

Further work on other cities' datasets will provide such a verification. Secondly, the non-linear relationship between PM2.5 concentrations and urban form, and the autocorrelation among some indicators (Haining, 2003), should be considered in further development. Thirdly, further integration with Multi-Objective Optimisation could extend the utility of an urban air quality platform by providing not only assessment but also recommended steps or moves to achieve improvements.

Researchers have recently proposed training pix2pix models with a Generative Adversarial Network to build generative machine learning frameworks (Yao et al., 2021). A future multi-objective optimisation framework realisation could make a breakthrough by moving into the emerging field of interaction design with AI, or machine learning assisted design.

References

Abdalla, T. E., & Peng, C. (2021). Evaluation of housing stock indoor air quality models: A review of data requirements and model performance. *Journal of Building Engineering*, 43, 102846. https://doi.org/10.1016/j.jobe.2021.102846

Bernard, J., Bocher, E., Le Saux Wiederhold, E., Leconte, F., & Masson, V. (2022). Estimation of missing building height in OpenStreetMap data: a French case study using GeoClimate 0.0. 1. *Geoscientific Model Development Discussions*, 1-47. https://doi.org/10.5194/gmd-15-7505-2022

Ministry of Ecology. (2022). *2020 Press Conference Records of Ministry of Ecology and Environment, the People's Republic of China. Springer Nature*. Taken from: http://www.mee.gov.cn/hjzl/sthjzk/zghjzkgb/

Rodríguez, M. C., Dupont-Courtade, L., & Oueslati, W. (2016). Air pollution and urban structure linkages: Evidence from European cities. *Renewable and Sustainable Energy Reviews*, 53, 1-9. https://doi.org/10.1016/j.rser.2015.07.190

Cheng, Y., He, K. B., Du, Z. Y., Zheng, M., Duan, F. K., & Ma, Y. L. (2015). Humidity plays an important role in the PM2. 5 pollutions in Beijing. *Environmental pollution*, 197, 68-75. https://doi.org/10.1016/j.envpol.2014.11.028

Das, T. (2017). *Machine Learning algorithms for Image Classification of hand digits and face recognition dataset. Machine Learning*, 4(12), 640-649.

Haining, R. P., & Haining, R. (2003). *Spatial data analysis: theory and practice*. Cambridge university press. http://dx.doi.org/10.1017/CBO9780511754944

He, K., Zhang, X., Ren, S., & Sun, J. (2015). *Deep residual learning for image recognition*. cite. arXiv preprint arxiv:1512.03385. https://doi.org/10.1109/CVPR.2016.90

He, Q., Gao, K., Zhang, L., Song, Y., & Zhang, M. (2021). Satellite-derived 1-km estimates and long-term trends of PM2. 5 concentrations in China from 2000 to 2018. *Environment international*, 156, 106726. https://doi.org/10.1016/j.envint.2021.106726

Iype, J (2021) *MAD's first social housing project Baiziwan integrates community into urban fabric*. Available at: https://www.stirworld.com/see-features-mad-s-first-social-housing-project-baiziwan-integrates-community-into-urban-fabric/ (Accessed: 17 August 2022).

Jekeli, C. (2006). Geometric reference systems in geodesy. Division of Geodesy and Geospatial Science, *School of Earth Sciences*, Ohio State University, 25.

Jia, B., Liu, S., & Ng, M. (2021). *Air quality and key variables in high-density housing. Sustainability*, 13(8), 4281. https://doi.org/10.3390/su13084281

Jurado, X., Reiminger, N., Vazquez, J., & Wemmert, C. (2021). On the minimal wind directions required to assess mean annual air pollution concentration based on CFD results. *Sustainable Cities and Society*, 71, 102920. doi.org/10.1016/j.scs.2021.102920

Jurado, X., Reiminger, N., Benmoussa, M., Vazquez, J., & Wemmert, C. (2022). Deep learning methods evaluation to predict air quality based on Computational Fluid

Dynamics. *Expert Systems with Applications*, 203, 117294. https://doi.org/10.1016/j.eswa.2022.117294

Liang, X., Zou, T., Guo, B., Li, S., Zhang, H., Zhang, S., Huang, H., & Chen, S. X. (2015). Assessing Beijing's PM 2.5 pollution: severity, weather impact, APEC and winter heating. Proceedings of the Royal Society. A, *Mathematical, Physical, and Engineering Sciences*, 471(2182), 20150257. https://doi.org/10.1098/rspa.2015.0257

Li, F., & Zhou, T. (2019). Effects of urban form on air quality in China: An analysis based on the spatial autoregressive model. *Cities*, 89, 130-140. https://doi.org/10.1016/j.cities.2019.01.025

Marsudi, I. (2019). *Extraction of Open Street Map to Produce Digital Maps. In IOP Conference Series: Earth and Environmental Science.* Vol. 366, No. 1, p. 012025. IOP Publishing. https://doi.org/10.1088/1755-1315/366/1/012025

Singh, J. (2016). *Baidu unveils open-source deep learning platform PaddlePaddle*.=

Sun, J., Wang, H., Song, Z., Lu, J., Meng, P., & Qin, S. (2020). Mapping essential urban land use categories in nanjing by integrating multi-source big data. *Remote Sensing*, 12(15), 2386. https://doi.org/10.3390/RS12152386

Wang, Q., Sandberg, M., Lin, Y., Yin, S., & Hang, J. (2017). Impacts of urban layouts and open space on urban ventilation evaluated by concentration decay method. *Atmosphere*, 8(9), 169. https://doi.org/10.3390/atmos8090169

Wang, S., Zhou, C., Wang, Z., Feng, K., & Hubacek, K. (2017). The characteristics and drivers of fine particulate matter (PM2. 5) distribution in China. *Journal of Cleaner Production*, 142, 1800-1809. https://doi.org/10.1016/j.jclepro.2016.11.104

Wang, X., Liu, C., Kostyniuk, L., Shen, Q., & Bao, S. (2014). The influence of street environments on fuel efficiency: insights from naturalistic driving. *International Journal of Environmental Science and Technology*, 11(8), 2291-2306. https://doi.org/10.1007/s13762-014-0584-1

Wei, J., Li, Z., Lyapustin, A., Sun, L., Peng, Y., Xue, W., Su, T., and Cribb, M. (2021). Reconstructing 1-km-resolution high-quality PM2.5 data records from 2000 to 2018 in China: spatiotemporal variations and policy implications, Remote Sensing of Environment (252), 112136, doi.org/10.1016/j.rse.2020.112136.

Yao, J., Huang, C., Peng, X. I., & Yuan, P. F. (2021). *Generative design method of building group-Based on generative adversarial network and genetic algorithm*. Proceedings of the 26th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2021, Volume 1, 61-70.

Yang, Y. (2009). *Chinese geodetic coordinate system 2000*. Chinese Science Bulletin, 54(15), 2714-2721. https://doi.org/10.1007/s11434-009-0342-9

Yuan, M., Song, Y., Huang, Y., Hong, S., & Huang, L. (2018). Exploring the association between urban form and air quality in China. *Journal of Planning Education and Research*, 38(4), 413-426. https://doi.org/10.1177/0739456X17711516

Yang, J., Shi, B., Zheng, Y., Shi, Y., & Xia, G. (2020). Urban form and air pollution disperse: Key indexes and mitigation strategies. *Sustainable Cities and Society*, 57, 101955. https://doi.org/10.1016/j.scs.2019.101955

Zhai, H., Yao, J., Wang, G., & Tang, X. (2022). Impact of Land Use on Atmospheric Particulate Matter Concentrations: A Case Study of the Beijing–Tianjin–Hebei Region, China. *Atmosphere*, 13(3), 391. https://doi.org/10.3390/atmos13030391

Züfle, A., Trajcevski, G., Pfoser, D., & Kim, J. S. (2020, June). *Managing uncertainty in evolving geo-spatial data. In 2020 21st IEEE International Conference on Mobile Data Management (MDM) (pp. 5-8)*. IEEE. https://doi.org/10.1109/MDM48529.2020.00021

SEPARATION OF TALL GREENERY COMPONENT IN 3D CITY MODELS BASED ON LIDAR DATA

PAWEŁ RUBINOWICZ[1]
[1] *West Pomeranian University of Technology in Szczecin, Poland*
[1] *pawel@rubinowicz.com.pl, ORCID: 0000-0003-0377-272X*

Abstract. The research analyses the possibility of separating tall greenery as a spatial component of the city in a 3D model based on LiDAR data. This applies in particular to Digital Surface Models (DSMs). The paper presents methods to generate theoretical DSMs of a city without tall greenery and tall greenery only as a component extracted from the 3D model. The first method is based on the use of GIS data, including 2D building outlines. The second method requires additional manual contouring of tall greenery. Both methods have been applied by the author in the planning practice in several cities in Europe. Results of the research are discussed in the article based on the example of Szczecin, Poland. It includes the preparation process, visualisations of theoretical DSM models (buildings without tall greenery and tall greenery only) and their application in urban analyses concerning e.g. protection and development of the cityscape. All simulations have been performed using C++ software developed by the author.

Keywords. 3D City Models, Digital Cityscape Analysis, Urban Planning, Visual Impact, Dsm, Lidar.

1. Introduction

The paper presents methods used to generate theoretical DSM models of a city with and without tall greenery as a component extracted from the 3D model. The term 'tall greenery', also referred to in literature as 'urban forest' (Münzinger, Prechtel and Behnisch 2022) – refers to trees in parks or squares in the city, trees planted in rows along streets, trees next to residential or commercial buildings and others. In this sense, tall greenery is a component of a spatial structure, which in many cities is comparable to buildings in terms of its volume (Figure 1). This is important for shaping the urban composition, building ecological links, and for the organisation of human living space (Li, Wenbo and Wei 2015). The ability to separate tall greenery in a 3D model determines the feasibility of multifaceted digital urban analyses.

Section 2 describes the 3D city model (DSM) and other GIS data used and outlines assumptions of the two methods that are used to separate tall greenery. Section 3 analyses the application of these methods using the example of the city of Szczecin, Poland. Section 4 compares these methods. The results are cross-referenced with other scientific studies and prospects are outlined for the development of the methods.

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 515-524. © 2023 and published by the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA), Hong Kong.

Section 5 presents a summary of the research aimed at the application of the methods in urban planning.

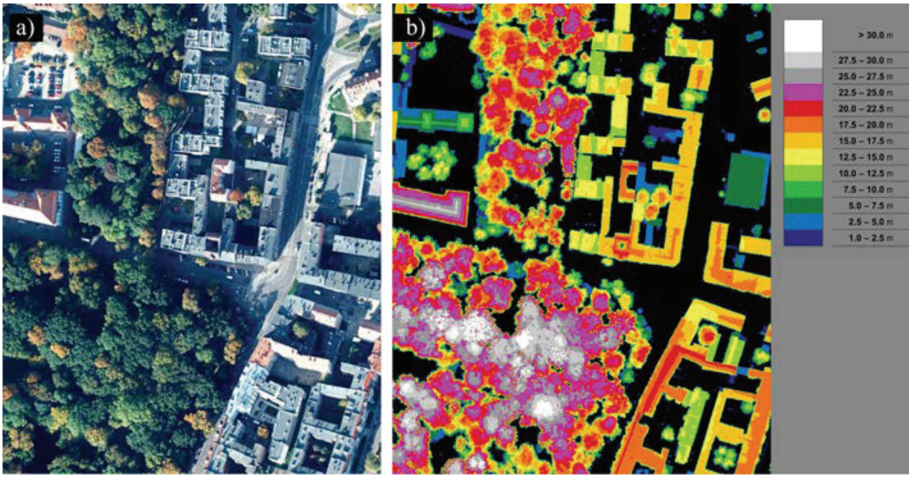


Figure 1. Tall greenery, an important component of the city space – analysis based on the example of Szczecin, Poland: a) orthophoto map; b) heights of buildings and tall greenery.

2. Materials and Methods

2.1. DIGITAL SURFACE MODELS (DSM)

The Digital Surface Model (DSM) is the primary resource used in the research. They map the spatial structure of the city with high accuracy. They are based on LiDAR data from ALS flights. The DSM point cloud is mapped on a regular grid. The 50 cm DSM grid for the city of Szczecin was used in this study (Figure 2a). The DSM outlines buildings with equal accuracy, and the same applies to tall greenery and other elements with a maximum height measurement error of less than 15 cm. DSM data are devoid of their semantics, which is, among other things, the foundation of CityGML models (Kolbe and Gröger 2003; Kutzner, Chaturvedi and Kolbe 2020). The DSM does not even include a basic classification of LiDAR areas (Morsy, Shaker and El-Rabbany 2017). However, together with the Digital Terrain Model (DMT), it can be an extremely valuable tool for the visualisation of the city, as well as for multi-faceted urban analyses.

2.2. COMPLEMENTARY 2D GIS DATA

For the purpose of isolating tall greenery, 2D GIS resources were used as supplementary data in one of the methods presented. Such resources are available worldwide in a wide variety of formats. They accurately map property issues, i.e. plot boundaries and building outlines. Other data, for example, land use, are usually considerably less precise. This includes areas of tall greenery. In the study for Szczecin, the publicly available BDOD10k resource (Rubinowicz 2017) was used, including a 2D building outlines layer (Figure 2b).

2.3. SOFTWARE

The research used C++ software that has been developed by the author since 2014. In the first phase of development, the software developers focused on the use of CityGML models. In the following years, in order to make use of the available digital resources, it was extended and adapted to process DSMs. The software allows DSM data to be visualised and used in complex urban analyses. These are used at their full accuracy resulting from ALS/LiDAR measurements and DSM grid density in the form of a continuous geometric surface (Rubinowicz 2019) displaying a 3D model of the city (Figure 2c).

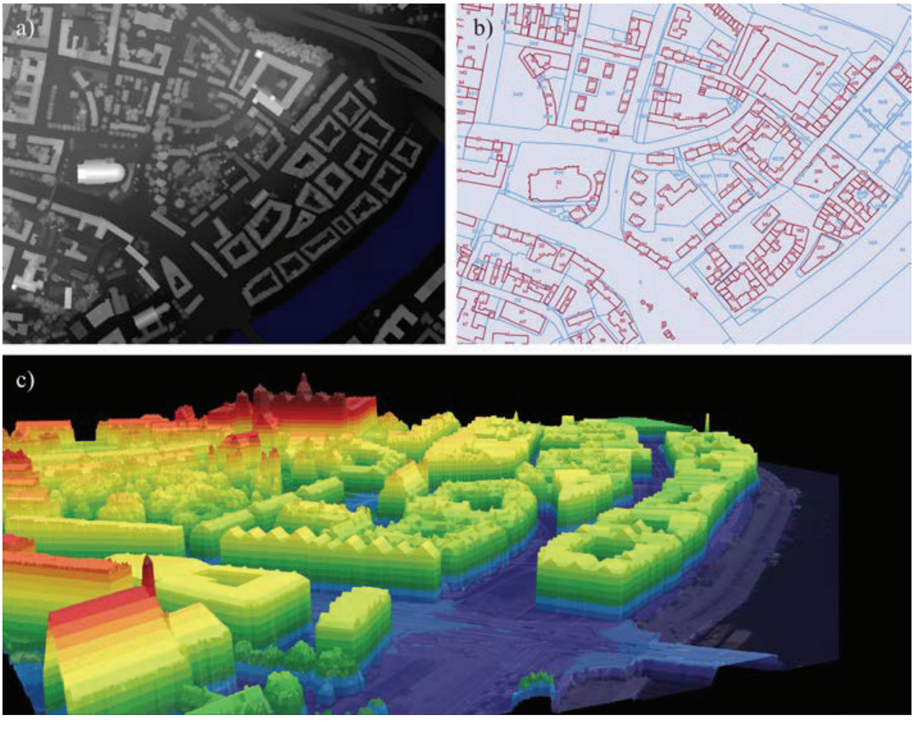


Figure 2. Digital data and tools used in the study: a) DSM model in the frame showing Szczecin, Poland, b) 2D GIS data view - BDOT10k, c) DSM visualisation in the C++ software.

2.4. METHODS DEVELOPED

This paper presents two methods that allow DSM models to be transformed to separate buildings and tall greenery. The first method, discussed in section 3.1, is based on the use of 2D GIS data including outlines of building. It can be applied automatically for very large urban areas. It allows the generation of a DSM model that includes only the built-up area of the city. In urban studies, such a model already provides a basis for analysing tall greenery through comparative analyses with the original DSM model. The second method, discussed in section 3.2, requires additional manual contouring of tall greenery areas on a dedicated digital layer and based on DSM and DTM data. It allows both the generation of a DSM model that includes the built-up area and a model that includes tall green only. The assumptions of the method and its implementation process are shown in the diagram below (Figure 3).

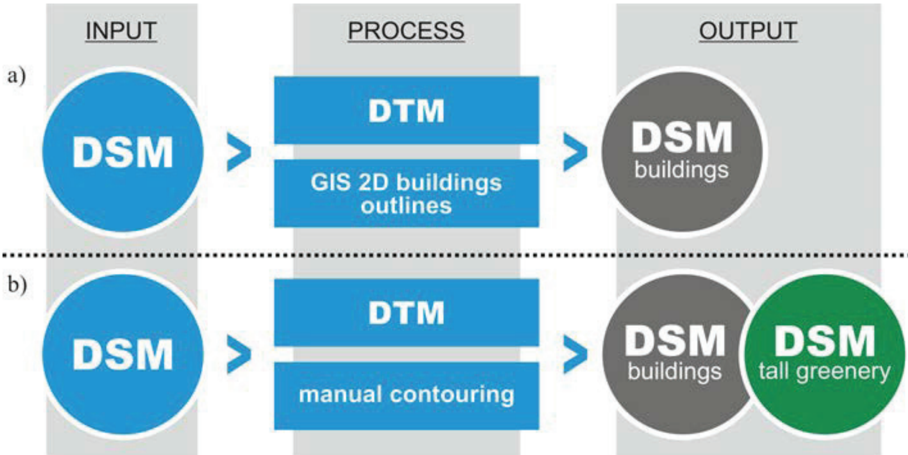


Figure 3. Assumptions of the methods for the separation of tall greenery in the DSM: a) the method explained in section 3.1, b) the method explained in section 3.2.

3. Results

3.1. SEPARATION OF TALL GREENERY USING 2D BUILDING OUT-LINES

An analysis of the method applicability was carried out on the example of a 36 km2 (6.0 km x 6.0 km) section of Szczecin, Poland. A visualisation of the entire area is shown in Figure 5. It encompasses the city centre with a diverse urban structure: quarter of historical buildings, multi-family and single-family houses, cemetery and parks, and a waterfront. A DSM 0.5 m grid model was used, which contains 144 million points for the surveyed part of the city, and a DTM model with 1.0 m grid accuracy. Outlines of buildings are derived from BDOT10k GIS data (Figure 4a, marked in white) and the analysis area contains a total of approximately 12,500 objects. DSM points located outside the building outlines are downscaled to the DTM level. The procedure is carried out automatically. The necessary calculations are fast, as these are done using the software (described in section 2.3) on a standard PC (6-core I7-8700, 3.70GHz, 32GB RAM) in a few seconds. As a result, a DSM model can be generated that represents only the built-up area of the city without tall greenery (Figure 4c).

3.2. SEPARATION OF TALL GREENERY BY MANUAL CONTOURING

An analysis of the method applicability was carried out using the example of a corresponding part of the city (described in section 3.1). Moreover, it used the same DSM and DTM resources. An essential element included additional contours of tall greenery (Figure 4a – marked in green). The data were entered manually, on a prepared

2D digital substrate depicting the structure of the city based on the DSM and DTM (such as Figures 4a or 1b). The contour marking can be carried out on various GIS or CAD platforms. It only requires a general indication of areas, without detailed contouring. Based on the research discussed, it takes between 0.5 and 2.5 hours (depending on the type of urban fabric) to prepare contours for 1 km2 of a city in a 0.5m DSM. Using the data, a 3D model of the city including only the built-up area can be developed (Figures 4d, g). This is slightly different and more accurate than the model obtained with the previous method (section 3.1, Figure 4c, f). Importantly, it is also possible to generate a 3D model that includes tall greenery only (Figures 4e, 5a).

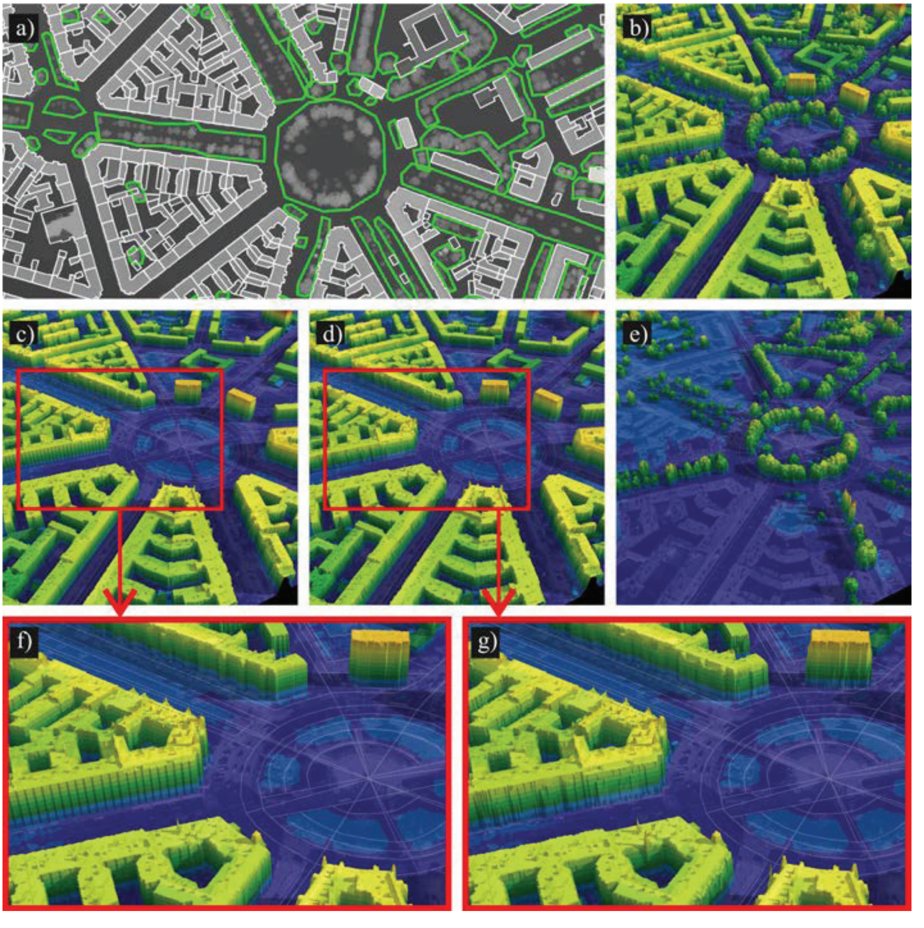


Figure 4. Input data and results of the discussed analyses on a small part of the city: a) projection with indication of GIS (BDOT10k) building outlines marked in white and manual high greenery outlines marked in green, b) full DSM model, c) building model according to GIS outlines, d) building model according to the method discussed in section 3.2, e) model including tall greenery only; f-g) c and d zoomed in

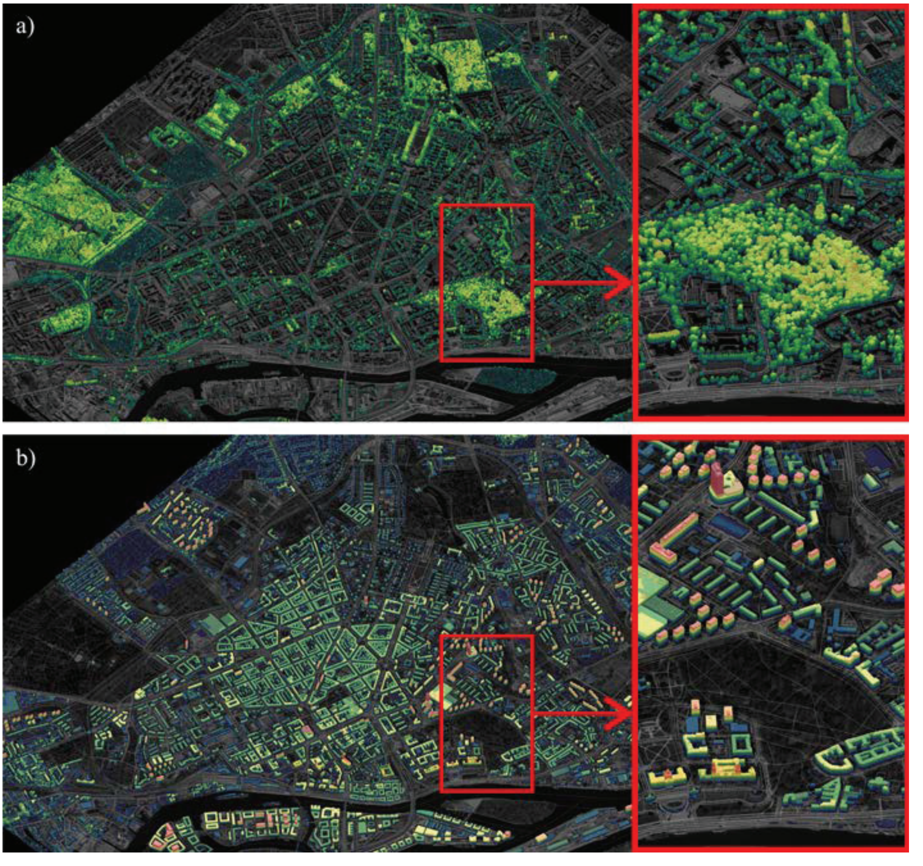


Figure 5. Results of analyses for the full area of Szczecin (Poland): a) model including only tall greenery, b) model including only buildings

3.3. APPLICATION OF URBAN ANALYSIS MODELS

Results of the methods considered for the full analysis Szczecin (36 km2) are shown in Figure 5. The juxtaposition of buildings and tall greenery separately provides valuable observations for planning (Figure 5). The immediate effect is that it helps to determine the degree of land use, intensity of development, and the proportion of tall trees. The resulting models also provide a basis for deeper analyses that allow for a meaningful assessment of the proportion of tall greenery in the urban structure. Another field of application, which is important for urban development, is cityscape analysis. This can be used from the protection of the existing cityscape, e.g. using the Visual Protection Surface (Czyńska and Rubinowicz 2015; Rubinowicz 2020), to the analysis of the impact of planned investment, including tall buildings, using the Visual Impact Size (Czyńska 2015, Czyńska 2018). In each of these simulations, it is important to include tall greenery. Its impact varies seasonally and depends on the translucency of trees. Figure 6 shows a comparative analysis of VIS simulations.

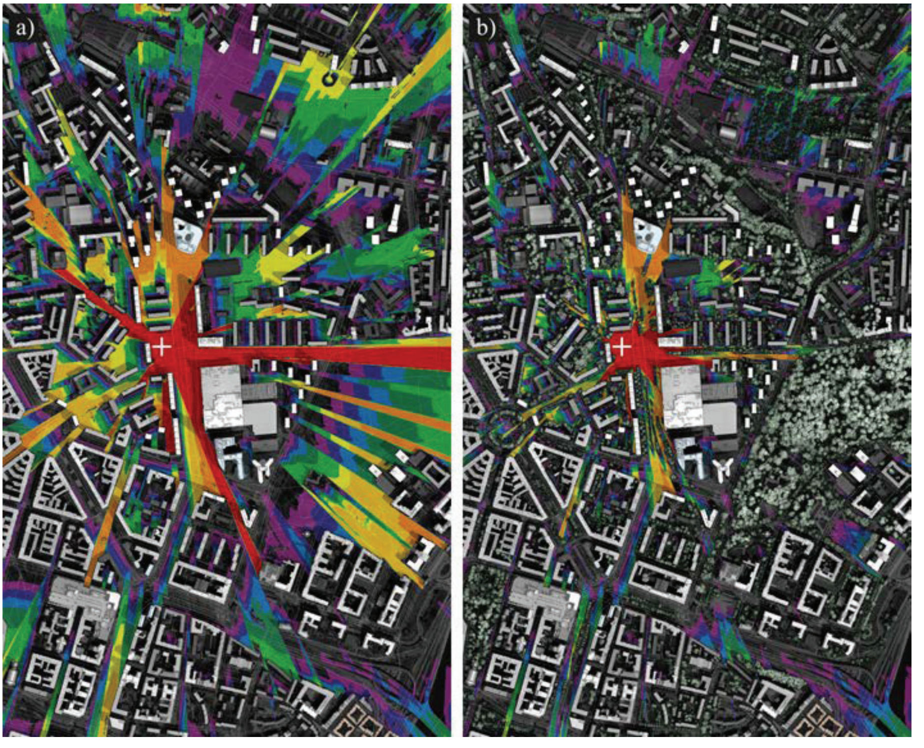


Figure 6. VIS analysis for a location of a new tall building in the centre of Szczecin:
a) VIS in the model including buildings only (without tall greenery), b) VIS in the full DSM model

4. Discussion

4.1. COMPARISON OF METHODS

The paper discusses two methods that can be used for separating tall greenery in DSM models. The first one (section 3.1) can be developed automatically using 2D building outlines. The method helps to develop a DSM model including buildings only (Figure 7a). It does not, however, allow for the creation of a DSM model that shows separately tall greenery. The negative of the model includes not only tall greenery, but also narrow parts of façades (Figure 7b). This is the result of the discrepancy between a geodetically precise outline of a building and the DSM point cloud plotted on an orthogonal grid. Another shortcoming of the method is that the available 2D outlines usually represent the ground level of buildings and not their contours 'seen from above'. This may lead to errors (Figure 7c, d). However, these errors affect only a small number of objects and are relatively minor. The second method (Section 3.2) offsets these errors and allows for a more complete use of DSM data. It helps to produce a more accurate model of buildings and a model of tall greenery. It does, however, require additional work, i.e. resource assessment and manual input of contours to separate buildings from tall greenery.

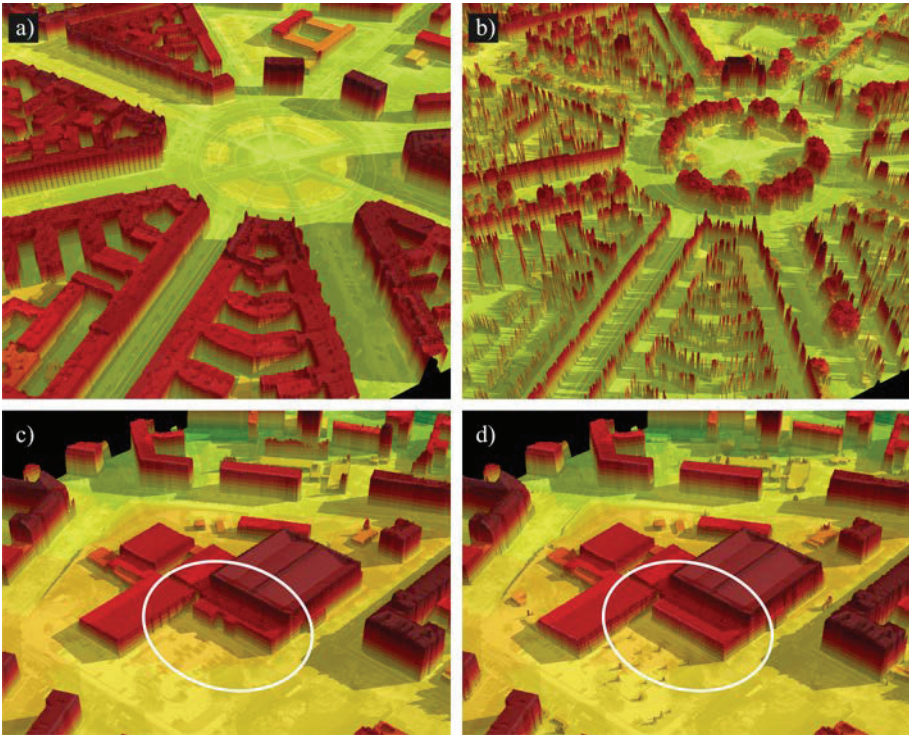


Figure 7. Application of the methods discussed: a) model that includes buildings only on the basis of their 2D GIS outlines, b) negative of the model, c) model of buildings obtained with the first method (section 3.1), d) model of buildings with the second method (section 3.2)

4.2. OTHER METHODS

DSM and DTM models are derived from LiDAR data. Irregular and denser point cloud is simplified to orthogonal grids. However, LiDAR (LAS) data contain more information, including colour and classification into terrain, buildings, vegetation, and unclassified components. In addition, the vegetation can be further distinguished as low, medium and high, which corresponds to the concept of separating tall greenery discussed in the article (Morsy, Shaker and El-Rabbany 2017; Yan, Shaker and El-Ashmawy 2015). However, data are scattered and heterogeneous. Thus, it is not to automatically indicate areas of tall greenery.

Application of LiDAR, DSM, DTM data for the developing 3D tall greenery models has been the subject of various scientific research e.g. studies for Oslo identifying ecosystem condition indicators (Hanssen et al. 2022, Gobeawan et al. 2018), just to mention the achievements of the Dresden team, that enables tree crown reconstruction with geometric primitives overlaid on the point cloud (Münzinger, Prechtel and Behnisch 2022). Although the complexity of the analyses undertaken by Dresden-based researchers is significantly larger and more advanced than the solutions discussed in the article, the resulting model does not increase the accuracy and specificity of 3D city model necessary e.g. for the application of the VIS analysis

(Figure 6). Solutions presented in the article includes application of DSMs as full complexity 3D models (consist of hundreds of millions of points and triangles mapping DSM). No geometric generalisation can increase the original accuracy of DSM, LiDAR based input data.

4.3. RESEARCH PROSPECTS

One way forward for the research in question is to use other 3D models, such as the reality-mesh-models used in Google Earth (Rubinowicz 2019). Their accuracy is potentially superior to that of the DSM. A significant difficulty is the lack of 'open source' access. However, LIDAR data are available. The classification of each cloud point included in them may help to partially automate the identification of areas with tall greenery and thus simplify the manual contouring process in the DSM to make necessary corrections. On the other hand, the use of 3D models of tall greenery and buildings separately (Figure 5) opens up the possibility to compare volumes of tall greenery and buildings for different parts of the city. It also facilitates the study of green area accessibility or ecological connections. DSM models are being updated as a result of the technology development at a reduced cost. This creates prospects for these methods to be applied to study spatial changes in a city, not only with regard to new buildings, but also to monitor the status of tall greenery (tree pattern, cuttings and new plantings).

5. Conclusion

Tall greenery (also referred to in literature as 'urban forestry') is a component of the city that is important for the urban composition. It forms ecological links and organises a human living space. This paper demonstrates the possibility of using the DSM to analyse tall greenery. Two methods are presented that enable to transform DSM to separate buildings and tall greenery. The first method is based on the use of GIS data, including 2D building outlines and allows for the automatic creation of theoretical DSM models that include buildings only. The second method requires additional manual contouring of tall greenery. Although more time-consuming, it provides for a better stratification of the DSM and the creation of a 3D model of the city that includes tall greenery only. The applicability of both methods is analysed in detail using the example of the city of Szczecin, Poland. The article shows examples of model application in cityscape analyses. Prospects for further research are presented as well. The methods discussed were previously used in the planning practice in several cities in Europe. The simulations were created using C++ software developed by the author. However, they can be reproduced and developed using other GIS/CAD software.

References

Czyńska, K. (2015). Application of Lidar Data and 3D-City Models in Visual Impact Simulations of Tall Buildings. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 2015. https://doi.org/10.5194/isprsarchives-XL-7-W3-1359-2015

Czyńska, K. (2018). High precision visibility and dominance analysis of tall building in cityscape - on a basis of Digital Surface Model. *Computing for a better tomorrow: Proceedings of the 36th eCAADe Conference*, ISBN: 978-94-91207-15-0, pages: 481-488.

Czyńska, K., Rubinowicz, P. (2015). Visual protection surface method: Cityscape values in context of tall buildings. In: K. Karimi, L. Vaughan, K. Sailer, G. Palaiologou, T. Bolton (eds), *Proceedings of the 10th International Space Syntax Symposium*, (pp. 142:1-142:10), London.

Gobeawan, L., Lin, E. S., Tandon, A., Yee, A. T. K., Khoo, V. H. S., Teo, S. N., … Poto, M. T. (2018). MODELING TREES FOR VIRTUAL SINGAPORE: FROM DATA ACQUISITION TO CITYGML MODELS. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLII-4/W10, 55–62. https://10.5194/isprs-archives-XLII-4-W10-55-2018

Gruber, M., Schachinger, B., and Mostafa, M. (2022). The next generation vexcel imaging 3d city modelling using directly georeferenced data, *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLIII-B1-2022, 309–316. https://doi.org/10.5194/isprs-archives-XLIII-B1-2022-309-2022.

Hanssen, F., Barton, D., Venter Z., Nowell M., Cimburova, Z. (2022). Utilizing LiDAR data to map tree canopy for urban ecosystem extent and condition accounts in Oslo. *Ecological Indicators*, vol. 130, https://doi.org/10.1016/j.ecolind.2021.108007.

Kolbe, T. H. & Gröger, G (2003). Towards Unified 3D-City-Models. In *Proc. of. ISPRS Commission IV Joint Workshop on Challenges in Geospatial Analysis, Integration and Visualization II*, Stuttgart.

Kutzner, T., Chaturvedi, K. & Kolbe, T.H. (2020). CityGML 3.0: New Functions Open Up New Applications. *Journal of Photogrammetry, Remote Sensing and Geoinformation Science* 88, 43–61. https://doi.org/10.1007/s41064-020-00095-z.

Li, H., Wenbo, Ch., Wei, H. (2015). Planning of Green Space Ecological Network in Urban Areas: An Example of Nanchang, China. *International Journal of Environmental Research and Public Health* 12, no. 10: 12889-12904. https://doi.org/10.3390/ijerph121012889.

Morsy, S., Shaker, A., El-Rabbany, A. (2017). Multispectral LiDAR Data for Land Cover Classification of Urban Areas. *Sensors* 17(5):958. https://doi.org/10.3390/s17050958.

Münzinger, M., Prechtel, N., Behnisch, B. (2022). Mapping the urban forest in detail: From LiDAR point clouds to 3D tree models. *Urban Forestry & Urban Greening*, vol. 74, https://doi.org/10.1016/j.ufug.2022.127637.

Rubinowicz, P. (2017). Generation of CityGML Lod1 City Models Using BDOT10k and Lidar Data. *Space & Form* 31, https://doi.org/10.21005/pif.2017.31.A-03.

Rubinowicz, P. (2019). Visualisation of DSM as 3D-Mesh for Urban Analyses. In Cocchiarella L. (eds), *Proceedings of the 18th International Conference on Geometry and Graphics (ICGG 2018, Milano). Advances in Intelligent Systems and Computing*, vol 809. Springer, (pp: 1013-1022). https://doi.org/10.1007/978-3-319-95588-9_87.

Rubinowicz, P. (2020). Sustainable development of a cityscape using the Visual Protection Surface method – optimization of parameters for urban planning. *Proceedings of the 25th International Conference of the Association for Computer-Aided Architectural Design Research in Asia, CAADRIA 2020*. https://doi.org/10.52842/conf.caadria.2020.1.863.

Yan, W., Y, Shaker, A., El-Ashmawy, N. (2015). Urban Land Cover Classification Using Airborne Lidar Data: A Review, *Remote Sensing of Environment*, 158, 295-310. https://doi.org/10.1016/j.rse.2014.11.001.

ANALYST PATTERNS OF INFLUENCE BETWEEN A COMMERCIAL DISTRIBUTION AND NEIGHBOURHOOD DYNAMIC IN A RESIDENTIAL NEIGHBOURHOOD

BOWEN CHEN[1] , PUI KUAN LAO[2] , ZHIYI DOU[3], WAI-SHAN QIU[4] and DAN LUO*[5]
[1,2,3,5] *The University of Queensland.*
[4] *Department of City and Regional Planning, Cornell University.*
[1]*b.chen1@uqconnect.edu.au, 0000-0003-4387-5379*
[2]*p.lao@uqconnect.edu.au, 0000-0001-9523-3525*
[3]*zhiyi.dou95@gmail.com, 0000-0002-0356-144X*
[4]*wq43@cornell.edu, 0000-0001-6461-7243*
[5]*d.luo@uq.edu.au, 0000-0003-1760-0451*

Abstract. The spatial distribution of urban commercial spaces significantly impacts the overall efficiency and vibrancy of adjacent neighbourhoods. As such, it is an important factor to consider during urban development. This study aims to examine the patterns of impact between commercial distributions and neighbourhood dynamics in a residential neighbourhood, based on the case study of a highly populated, thriving commercial, and culturally rich area situated in Mong Kok, Hong Kong. In this research, a series of numeric evaluations and statistical analyses of liveability and vibrancy metrics are presented, uncovering the tension created by existing commercial forms and local living patterns. This research started with multi-dimensional data mining, such as accessing planning data using Geographic Information Systems (GIS), perception data using Street View Images (SVI), and business performance data from Google; secondly, analysing the data via machine learning (ML) algorithm and statistical correlation to identify correlations overlaid with a mapping of spaces of measurable characteristics. The goal is to establish a measurable evaluation of the relationship between commercial vibrance and urban features that can further inform the impact of urban design strategies on fostering the vitality of community commercial centres.

Keywords. Mong Kok, Model Learning Machine (ML), SVM, PSPnet, MaskRCNN, POI, Commercial Vibrance, Heatmap Correlation, Visualization, QGIS, Google Maps Information

1. Introduction

A vibrant residential neighbourhood in urban analysis is commonly defined as a place that shows diverse and frequent consumer activities (Wang et al., 2021). For instance,

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 525-534.

many researchers have found that the growth of commercial activities is significantly faster within transportation corridors than outside of them (Ma et al., 2023; Abe et al., 2014; Foth, 2010). Commercial vibrancy is traditionally seen as a broad measure of urban success; however, cities have specific needs that must be met to ensure successful spatial development and community vitality. Every city has its own unique set of requirements that must be taken into consideration when evaluating its economic health and overall well-being. It is essential to recognise the individual needs of each city and make specific requirements to ensure their long-term prosperity. Some papers in recent years, such as the one by Liu, Gou, and Xiong (2022), have focused on utilising cultural politics or demographics to evaluate urban vitality. However, as urban planners, our qualities are those related to land use, amenities, and infrastructure. This article will concentrate on a more granular approach to urban vibrancy, rather than considering it as a general concept, our research aims to identify the specific characteristics that make up each type of urban form. By doing so, we hope to gain a better understanding of the nuances of urban vibrancy and the ways in which it can be enhanced.

Through an examination of Mong Kok's urban context, this article seeks to understand how the different commercial forms of Mong Kok contribute to and shape the city's infrastructure and design quality. The findings of this research have the potential to provide valuable insight into the dynamics of urban planning and development. The study identified six commercial types, of which one has three different evaluation metrics. A quality measurement was then created for each type of business form, providing a comprehensive understanding of the strengths and weaknesses of each option. This will enable business owners to make an informed decision when selecting a business structure that is best suited to their individual needs. A comprehensive research project was conducted in which forty different urban features were collected and analysed. A regression model was then used to investigate the relationship between various business forms and the aforementioned urban features. The results of the study provide valuable insight into how different business forms can be impacted by a variety of urban features.

2. Study Area

A case study was conducted in Mong Kok, Hong Kong, to validate this approach. It is located on the Kowloon Peninsula's southern tip (Xue et al., 2001). The area is densely populated, with a population density of 130,000 people per square kilometre (Aranda-Mena et al., 2018). Mong Kok is a popular tourist destination for both local and foreign visitors. It is a place that combines features of old and new, east and west. The district is home to several markets, including the Flower Market, the Goldfish Market, and the Jade Market (Aranda-Mena et al., 2018). Mong Kok is also a food-lover's paradise, with a wide variety of restaurants and street food stalls, making it a busy and vibrant area. As a transport hub, Mong Kok has a comprehensive and diverse transport network. As demonstrated in Figure 1 and Figure 2, Mong Kok is the most representative of Hong Kong's traditional urban civilization, with complex land use and a diverse urban structure when compared to other precincts.

Industry	POI subclass	Quantity	Proportion
Retail	supermarkets, grocery stores, variety stores and pharmacy	115	8%
Catering	Chinese and foreign restaurants	1156	80%
Finance	banks, and insurance agents	49	3%
Recreation	theatres, amusement parks, gym, KTV	35	2%
Auto service	car repair/sales and motorcycle service	26	2%
Accommodation	hotels and Airbnbs	56	4%

Figure 1: Classification of commercial POI in Mong Kok

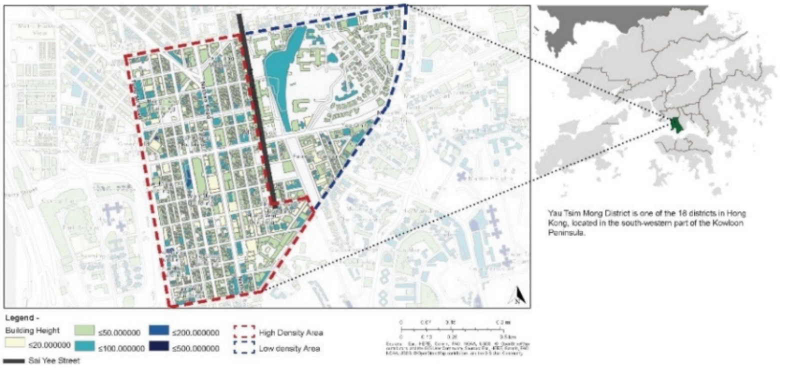


Figure 2: The map shows an area of highly diverse buildings on the west side of Sai Yee Street

3. Methodology

3.1. WORKFLOW FRAMEWORK DIAGRAM

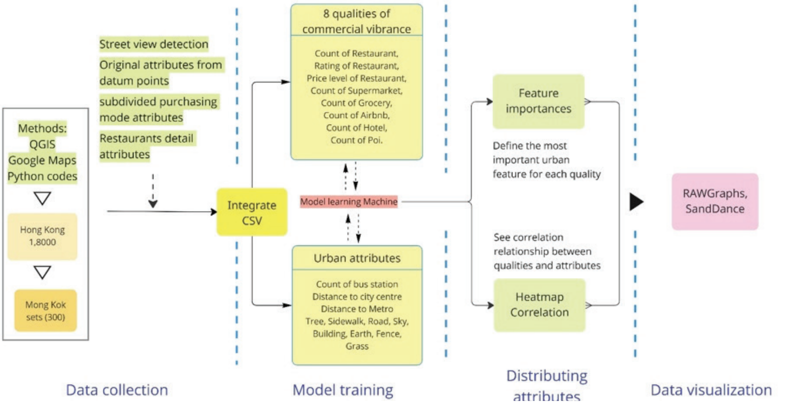


Figure 3： Workflow of research and analysing methodology

Figure 3 clarifies the overall workflow. The research starts with data collection, which narrows down the primitive 18,000 datum points that are computed from QGIS and Google Maps. Then for augmenting data, a variety of features including street view detection, restaurant detail and purchasing mode are extracted via different urban analysis models, which are calculated and assigned as attributes according to their proximity and are integrated into the CSV file. Eight attributes are selected as the qualities reflecting commercial vibrance, and the correlation between qualities and urban attributes is tested via Model Learning Machine. Following the above procedure, in order to distribute attributes, the most important urban features for each quality are defined, and an integral heatmap is computed for revealing the overall correlation. In order to better convey the findings, the final visualization step is to describe graphical views of the analysed data.

3.2. DATA COLLECTION

As shown in Figure 4, for collecting additional data, four categories are defined based on their different research methodologies.

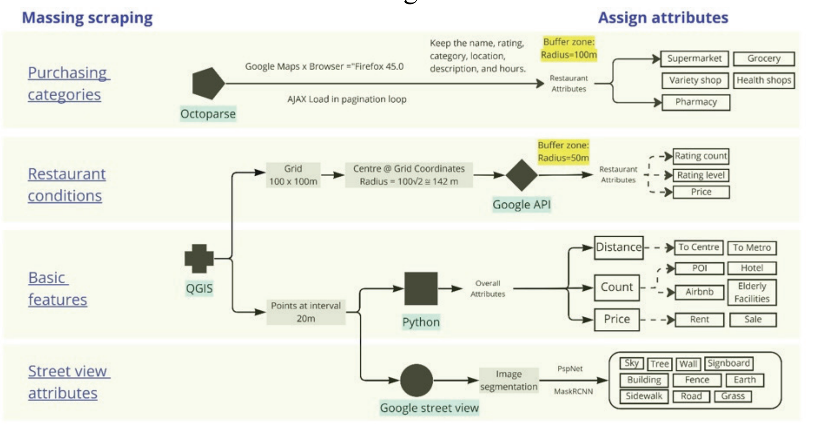


Figure 4: Four approaches of data extraction

3.3.1 Purchasing mode attributes (First row in Figure 4)

"Supermarket", "Grocery", "Variety Store", "Health Facility", and "Pharmacy" have been detected as the main five representative categories of Mong Kok's purchasing action, which also encapsulate the daily routines of purchasing behaviours. These shopping modes are analysed via Octoparse, creating a pagination loop with AJAX Load and selecting the expected sections (Supermarket, Grocery, etc,.) enabling Octoparse to scrape the targeted information, then to split the whole CSV file into five individual ones and input them into QGIS for further analysis.

3.3.2 Restaurants detail attributes (Second row in Figure 4)

The great prosperity and high diversity of restaurants are reckoned as one of the biggest features of Mong Kok district, which directly reflects the status of commercial distribution in this district. A 100 x 100 m grid with 54 nodes was created in QGIS, and the radius was set to be 142 m (100√2m) to cover the whole research area. By applying Google Maps with coordinates and radius in the URL, information about restaurant rating count, rating levels, and price levels were collected and saved in CSV. By manually culling duplicated items, 1156 restaurants' information could be processed to the next step.

The theory of assigning the targeted attributes to datum points is based on the vicinity between target points and datum points. For restaurants, a buffer zone with a radius of 50 m is created at each datum point, and for purchasing modes, a radius of 100 m is applied; values of the feature points that fall into this circle are calculated, and the mean value is appended at the end of the table as a new field of attribute. The same steps are processed for each datum point, and its urban values are extended by analysing the performance of nearby restaurants and purchasing modes.

3.3.3 Datum data and Basic attributes (Third row in Figure 4)

18,000 geographic points in Hong Kong are collected via the Quantum Geographic Information System (QGIS) based on the rule of “every 20 metres, one point." Then, by using Google Maps and Python coding, a variety of urban features of these points were computed. Besides the geographic coordinates (longitude and latitude) and the distance, several other features are also detected, digitised into numbers, and added to the table as quantifiable attributes for further analysis. These features contain the counting type, the price type, and the distance type. Then 300 points are selected from Yau Tsim Mong (217) and Kowloon City (83) to cover the area-subjective and saved in a CSV file.

3.3.4 Street View detection and processing (Forth row in Figure 4)

Based on the geographic coordinates from QGIS and the local road data, the CSV file was read in Python codes, and the Street View Imagery (SVI) of each datum point was processed. These screenshots were identified as "paths” in MaskRCNN and PSPNet codes; targeted street features were defined in the code as “keywords” (Zhu et al., 2021; Yang and Guo, 2022). These elements include buildings, sidewalks, signboards, sky, roads, trees, plants, grass, wall, earth, and fence, which are the main recognisable features; the values are reparameterized in decimals from the proportion of the street view image and recorded in CSV.

3.3. MODEL TRAINING AND DISTRIBUTIONS ATTRIBUTES

Eight detected attributes are selected as the qualities to reflect the commercial vibrance and reparameterized in the range of 0–1; these include the Rating, Count, and Price of restaurants and the counts of POI, Grocery, Hotels, Airbnb and Supermarkets. Detected elements from Image segmentation and other physical features are tested as inputs into Machine Learning Model to reveal the correlation between existing physical conditions and subjective commercial vibrance (Nugrahaeni and Mutijarsa, 2016, Qiu et al., 2021).

Figure 5 shows the types and the results of model learning machines, which test R^2, RMSE, and MAE for each of the eight qualities, and Figure 6 contains eight sub-figures sorted by the eight qualities, which show the ranking of urban features in terms of their importance. Figure 7 shows the correlation in the heatmap (Gu et al., 2016), which demonstrates that the distribution of Airbnb, Hotel, and POI are deemed to be highly correlated in terms of the commercial vibrance in the context of the Mong Kok district, but the bus station leaves a negative correlation to those three qualities. The distance to the metro and the distance to the city centre negatively correlate with all the commercial vibrance-defined eight qualities; clearly the more remote and with poor transport connections a place becomes, the less vibrant it will be in terms of a commercial environment. While the density of bus stations is the most important factor for the three tourist features (Airbnb, Hotel, POI), it presents a negative correlation with them.

	Q1_Count of Restaurant			Q2_Rating of Restaurant			Q3_Price level of Resaurant			Q4_Count of Supermarket			Q5_Count of Grocery			Q6_Count of Airbnb			Q7_Count of Hotel			Q8_Count of POI		
Model	R^2	RMSE	MAE	R^2	RMSE	MAE	R^2	RMSE	MAE	R^2	RMSE	MAE	R^2	RMSE	MAE	R^2	RMSE	MAE	R^2	RMSE	MAE	R^2	RMSE	MAE
KNN	0.5	0.09	0.06	0.41	0.29	0.22	0.26	0.28	0.2	0.31	0.16	0.11	0.25	0.17	0.11	0.58	0.09	0.06	0.67	0.09	0.06	0.57	0.05	0.04
SVM	0.39	0.1	0.08	0.35	0.3	0.23	0.39	0.25	0.19	0.14	0.18	0.14	0.02	0.2	0.14	0.59	0.09	0.08	0.67	0.09	0.07	0.31	0.07	0.06
Random Forest	0.47	0.1	0.07	0.51	0.27	0.19	0.39	0.25	0.17	0.25	0.17	0.12	0.37	0.16	0.1	0.91	0.04	0.03	0.96	0.03	0.02	0.85	0.03	0.02
Decision Tree	0.46	0.1	0.06	0.43	0.28	0.15	0	0.32	0.2	0.1	0.19	0.1	0.17	0.21	0.1	0.79	0.06	0.04	0.75	0.08	0.06	0.7	0.05	0.03
OLS	0.43	0.1	0.08	0.46	0.28	0.22	0.38	0.26	0.2	0.15	0.18	0.14	0.13	0.19	0.13	0.76	0.07	0.06	0.79	0.07	0.06	0.63	0.05	0.04
Gaussian	0.44	0.1	0.08	0.44	0.28	0.23	0.36	0.26	0.2	0.16	0.18	0.14	0.13	0.18	0.13	0.76	0.07	0.06	0.79	0.07	0.06	0.63	0.05	0.04
Voting Selection	0.48	0.09	0.07	0.52	0.26	0.2	0.37	0.26	0.19	0.25	0.17	0.13	0.26	0.17	0.11	0.89	0.05	0.04	0.92	0.05	0.04	0.81	0.04	0.03
ADA Boost	0.3	0.11	0.09	0.4	0.29	0.26	0.34	0.26	0.21	0.17	0.18	0.15	0.14	0.18	0.15	0.88	0.05	0.03	0.93	0.04	0.03	0.83	0.03	0.03
Bagging Regression	0.37	0.1	0.08	0.35	0.31	0.23	0.39	0.25	0.19	0.14	0.18	0.14	0.05	0.19	0.14	0.58	0.09	0.08	0.66	0.09	0.07	0.28	0.07	0.06

Figure 5: Nine Model Learning Machines about the correlation between features and qualities

Figure 6: Result of feature importance from Model Learning Machine

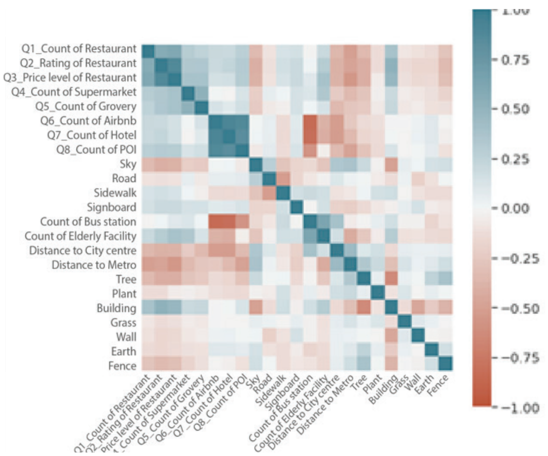


Figure 7: Heatmap correlation between 8 qualities of commercial vibrance and urban features

4. Results

4.1. THE RELATIONSHIP BETWEEN COMMERCIAL VIBRANCY AND URBAN PUBLIC FEATURES

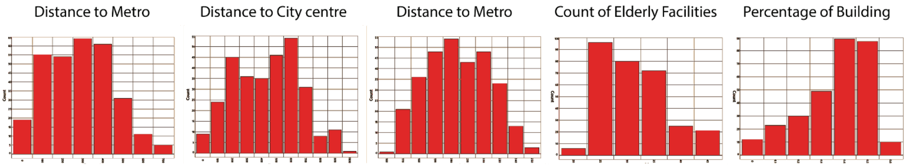


Figure 8: The distribution of the five most important urban public features

Through the above analysis and testing, five out of eight urban qualities show the significance in influencing urban vibrance, Figure 8 presents the condition of Mong Kok in terms of the five urban qualities.

Different types of businesses have different needs and therefore require different urban layouts to accommodate them. The findings of Figure 9 indicate that the area of Mong Kok in Hong Kong has a positive relationship between the number of restaurants and the distance to the nearest metro station. This is indicative of the importance of the metro station for commercial activity in Mong Kok and suggests that the metro station is an integral part of the local economy. The proximity of a restaurant to public transportation also has a major influence on the prices that restaurants charge. In general, restaurants located closer to metro stations tend to have higher prices than those located further away. This is because customers who are closer to public transportation are more likely to frequent the restaurant, resulting in higher demand and, therefore, higher prices. Additionally, restaurants located further away from public transportation may be able to offer lower prices since they are less likely to receive as much traffic. The correlation figures also indicate that there is a strong relationship between the proximity of a restaurant to transportation links and the city centre and the reviews the restaurant receives. This suggests that restaurants located close to transportation links and the city centre can benefit from increased visibility and accessibility, resulting in more positive reviews from customers. Furthermore, this could lead to an increase in customer satisfaction, resulting in improved business performance.

Moreover, as seen in Figure 10, the bustling district of Mong Kok has been found to have a positive correlation between the number of supermarkets and their proximity to the metro, the elderly facilities, and the city centre. As the proximity to these areas increases, so does the number of supermarkets. This is beneficial for both businesses and consumers, as it ensures easy access to products and services for the elderly who may not be able to travel far, as well as those living near the centre who can conveniently access a variety of stores. Furthermore, this association between supermarkets and metro stations helps to create a more vibrant and diverse atmosphere in the area. On the contrary, the most crucial aspect of the number of grocery stores in Mong Kok is their proximity to the city centre. The number of bus stations is the second significant element, while the distance to the Metro is the third, which is the inverse of the key feature for restaurants. The easy access to fresh food and other essential items that grocery stores provide is essential for the well-being of the city's citizens. Having grocery stores that are conveniently located close to the city centre ensures that individuals can access these necessities quickly and easily, saving time and money. The number of grocery stores and their proximity to the city centre are thus key factors in providing a healthy and prosperous environment for the citizens of the city. The money that it generates goes back into local businesses, which support local communities and help them grow in size over time. This means that people living in these communities are employed locally, which helps to keep crime levels low and helps to strengthen community bonds between residents.

Distance / Place	Grocery	Supermarket	Health facility	Pharmacy	Variety store	Restaurant
50m	0	0	1	0	0	18
100m	0	2	2	1	1	110
150m	1	4	2	2	2	229
200m	6	8	6	7	4	418

Figure 9: Distribution of commercial industry based on the distance to the metro

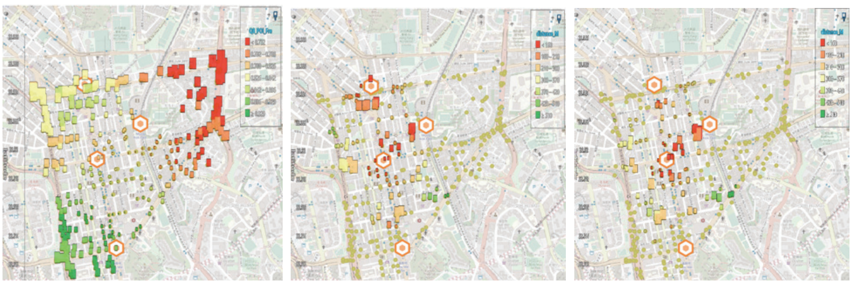


Figure 10: Metro distribution and Set "Distance to city centre, Supermarket and Grocery"

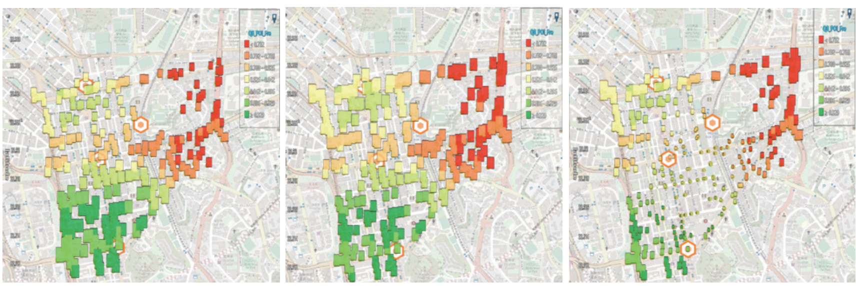


Figure 11: POI distribution and Set "Airbnb, Bus Station and Distance to centre"

The findings of Figure 11 also demonstrate that as the number of bus stops in Mong Kok increases, the number of Airbnbs, hotels and POI located nearby decreases. This could be due to a variety of factors, such as increased competition for tourists' attention or an overall decrease in the demand for hotel and Airbnb accommodations in Mong Kok. The decreased presence of Airbnb and hotels in certain areas may be attributed to the higher population of permanent residents in those areas. With more people living in these areas, there is a higher concentration of bus stations and other public transportation options, resulting in fewer visitors and, subsequently, fewer Airbnb and hotel accommodations. On this basis, it can be revealed that tourism accommodation usually depends upon transportation, but traffic planning always facilitates the POI. Such a result has significant implications for urban planning and tourism, as it suggests that public transportation could be a viable alternative for travellers who are looking to stay in the region. It also indicates that local policymakers should consider the potential effects of increased bus stops on the tourism industry when making decisions about the allocation of resources.

The paper developed the correlation and statistical regression between different commercial forms and urban contexts based on the above research methodology and data analysis. In conclusion, different commercial qualities can be influenced by a multiplicity of different urban contextual factors in terms of their vibrance. According to Figure 6, these eight commercial qualities can be roughly classified into three groups: the three dimensions of the restaurant and supermarket are sensitive to the feature DM; grocery is more sensitive to the feature DC; and the other three tourism-related qualities (Airbnb, hotel, and POI) are more reliant on the feature CS. It is obvious that “distance to the metro” is the biggest factor that affects the catering industry and supermarket distribution. Moreover, Airbnb, hotels, and points of interest (POIs) are important components of the tourism industry. Their distribution is more widespread in Mong Kok, and the most sensitive feature of these components is the number of bus stops in the area, as this is key to ensuring that visitors can easily reach their destinations in a dense area like Mong Kok. This is because buses can access more areas than the metro, which typically follows a more limited route. Bus stops are often placed in areas that are not easily accessible by the metro, allowing individuals to travel to and from those locations.

5. Reflections

5.1. INCOMPLETENESS/ DISINTEGRATE OF SCRAPPABLE INFORMATION

This research is still in its early stages, but it has the potential to transform the way we understand and improve cities. The detailed data of the restaurant is sourced from Google Maps rather than the professional UE App (User Experience Application); therefore, some restaurants are unsourced due to their small scale or ordinariness, which presents insufficiencies in food scoring and feedback per capita consumption.

5.2. HONG KONG’S FLUIDITY

Due to the high variability of government policies and commercial formats in Mong Kok, the research objects and supporting data in this paper have certain time limitations. There may still be undetected dismantlement and reconstruction projects. Therefore, the results of this research merely reflect the current commercial distribution and data collection in Mong Kok, Hong Kong.

5.3. LIMITATION AND FUTURE DEVELOPMENT

The analysing tools in this paper include Python coding, SVI with Image segmentation PSPnet, MaskRCN, Octoparse, and Machine Learning Model etc., which themselves have certain limitations for data monitoring and capture. For example, PSPnet and MarkRCN can only analyse the street images extracted from Google Street View at a specific time, and the street images can only represent the street conditions within the picture frame at a certain viewing height. The recognition of objects in the analysis is also limited to some extent by the capacity of these tools, which might lead to some inaccuracy in reflecting all target identification.

Although Hong Kong is a very distinctive setting, we may apply this commercial reflection of urban activities technique in the future to assess the realistic natural gathering of commercial vitality. This method could be used elsewhere and compared to other cities to determine if there are differences in urban organisation or structure; therefore, applying this technique to other cities is our next step.

Data Accesses: The data that support the findings of this study are available from the corresponding author Dr Dan Luo, upon reasonable request.

References

Abe, S., Nakagawa, D., Matsunaka, R., & Oba, T. (2014). Study on the factors to transform underused land focusing on the influence of railway stations in central areas of Japanese Local cities. Land Use Policy, 41, 344–356. https://doi.org/10.1016/j.landusepol.2014.06.021

Foth, N. M. (2010). Long-Term Change Around SkyTrain Stations in Vancouver, Canada: A Demographic Shift-Share Analysis. The Geographical Bulletin, 51(1), 37.

Gu, Z., Eils, R., & Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics, 32(18), 2847–2849. https://doi.org/10.1093/bioinformatics/btw313

Liu, H., Gou, P., & Xiong, J. (2022). Vital triangle: A new concept to evaluate urban vitality. Computers, Environment and Urban Systems, 98, 101886. https://doi.org/10.1016/j.compenvurbsys.2022.101886

Ma, Z., Li, C., Zhang, P., Zhang, J., Liu, D., & Xie, M. (2023). The impact of transportation on commercial activities: The stories of various transport routes in Changchun, China. Cities, 132, 103979. https://doi.org/10.1016/j.cities.2022.103979

Nugrahaeni, R. A., & Mutijarsa, K. (2016). Comparative analysis of machine learning KNN, SVM, and random forests algorithm for facial expression classification. 2016 International Seminar on Application for Technology of Information and Communication (ISemantic), 163–168. https://doi.org/10.1109/ISEMANTIC.2016.7873831

Qiu, W., Li, W., Liu, X., & Huang, X. (2021). Subjectively Measured Streetscape Perceptions to Inform Urban Design Strategies for Shanghai. *ISPRS International Journal of Geo-Information, 10*(8). doi:10.3390/ijgi10080493

Wang, P., Liu, K., Wang, D., & Fu, Y. (2021). Measuring Urban Vibrancy of Residential Communities Using Big Crowdsourced Geotagged Data. Frontiers in Big Data, 4, 690970–690970. https://doi.org/10.3389/fdata.2021.690970

Yang, C., & Guo, H. (2022). A Method of Image Semantic Segmentation Based on PSPNet. Mathematical Problems in Engineering, 2022. https://doi.org/10.1155/2022/8958154

Zhu, X., Cheng, Z., Wang, S., Chen, X., & Lu, G. (2021). Coronary angiography image segmentation based on PSPNet. Computer Methods and Programs in Biomedicine, 200, 105897–105897. https://doi.org/10.1016/j.cmpb.2020.105897

Xue, C. Q. L., Manuel, K. K., & Chung, R. H. Y. (2001). Public space in the old derelict city area - a case study of Mong Kok, Hong Kong. Urban Design International (London, England), 6(1), 15–31. https://doi.org/10.1057/palgrave.udi.9000032

ASSESSING SPATIAL ACCESSIBILITY TO PUBLIC FACILITIES FOR VULNERABLE PEOPLE TOWARDS 15-MINUTE CITY IN HONG KONG

A Comparative Study of Sham Shui Po and Tin Shui Wai

FANGNING WU[1]
[1]*School of Architecture, The Chinese University of Hong Kong*
[1]*fangningwu@link.cuhk.edu.hk, 0000-0003-2825-6609*

Abstract. Due to the COVID-19 pandemic, people started to rely more on their communities and attach great importance to the accessibility of public facilities at a hyperlocal level. The 15-Minute City concept, first put forth by Carlos Moreno in 2016, gradually gained popularity worldwide during the age of pandemics. This human-centric concept aims to build complete neighbourhoods that meet the daily needs of residents within 15 minutes by using non-motorized transport. However, few studies focus on vulnerable groups such as children and the elderly for the assessment of 15-Minute City. Therefore, this paper provides an assessment framework for spatial accessibility to public facilities for vulnerable people from the perspective of 15-Minute City. It is measured in three aspects: spatial distribution characteristics, service population ratio and number of facilities through the comparison between Sham Shui Po and Tin Shui Wai in Hong Kong. The result shows that the accessibility to public facilities needed in daily life for children in Sham Shui Po and Tin Shui Wai is relatively reasonable, while the accessibility for the elderly needs to be improved. The research can provide references for evaluating and optimizing spatial planning to promote health and well-being in Hong Kong and other cities.

Keywords. 15-Minute City, Vulnerable People, Spatial Accessibility, Network Analysis, Human-centric, New Town

1. Introduction

For decades, people have been affected by climate change caused by excessive greenhouse gas emission and health issues due to the lack of exercise. The reliance on automobiles is one of the main reasons. Therefore, walking, cycling, and public transportation are widely recommended for their environmental, economic, and health benefits. In particular, measures such as quarantine, lockdown, and curfew greatly influenced residents' travel behaviour during the COVID-19 pandemic. Residents have significantly reduced long-distance travel and begun relying on their communities to meet their daily needs. Because of this, accessibility to public facility and public space has become increasingly critical. In this context, the emergence of the 15-Minute City concept attracted extensive attention from governments and scholars worldwide (Pozoukidou and Chatziyinnaki, 2021). It was proposed in 2016 by Carlos Moreno (Moreno et al., 2021) and was included in Paris mayor Anne Hidalgo's re-election declaration in 2020 (Peters,2020).

This people-oriented concept hopes that residents can access their daily needs from their housing within 15 minutes by non-motorized transport. It pays attention to the spatial equity of vulnerable group due to economic or physiological reasons. The vulnerable group for economic reason is mainly the low-income group. Travel cost has a more significant impact on their travel choices. Accordingly, the promotion of non-motorized transport through the 15-Minute City is conducive to improving the accessibility to public facilities for them. On the other hand, the vulnerable group for physiological reason mainly include the elderly, children and the disabled (Zhong et al., 2022). They need short-distance, safe and barrier-free travel in their daily life which could also be a critical aspect of the 15-Minute City.

As an international metropolis, Hong Kong has a vast population but scarce land resources. In order to disperse people in the overcrowded urban area and improve their quality of life, Hong Kong started to build new towns led by public housing development in the 1970s. Self-containment, one of the principles for new town construction (He et al., 2020), hopes that a new town can meet the basic needs of its residents in terms of housing, employment, education, recreation and other community facilities. It is consistent with the idea of 15-Minute City. However, the liveability in some of Hong Kong's new towns still needs to be enhanced, and vulnerable groups' needs must be prioritised. For the research of 15-Minute City assessment, few studies focus on residents in different types of housing and vulnerable groups such as children and the elderly.

Therefore, this paper provides an assessment framework for spatial accessibility to public facilities for vulnerable people towards 15-Minute City using buildings as the analytical units. Two low-income districts in Hong Kong were selected. One is Sham Shui Po (SSP) in the urban area, and the other is Tin Shui Wai (TSW) new town. A comparative study is conducted on the relationship between demand and supply for public facilities and vulnerable people. The research results can provide references for optimizing the distribution of public facilities to improve health and well-being in Hong Kong and other cities.

2. Previous Research on the Assessment of 15-Minute City

At present, network analysis is commonly used for 15-Minute City assessment. There are generally two methods of evaluation (Table 1). One is to calculate the residential area or population that can be covered by a 15-minute walk distance from the public facility (Capasso Da Silva et al., 2019; Ferrer-Ortiz et al.,2021; Zhong et al.,2022). And the other is to calculate the number of public facilities within a 15-minute walk from the housing (Ma et al., 2020; Zhuo et al.,2021). In theory, the latter is more in line with the 15-Minute City concept regarding housing as the centre. Furthermore, some existing studies use only one temporal range while others use multiple ranges that can be more precise. Most of the analytical units are parcels and few studies involve vulnerable groups.

Table 1. Summary of previous research

Centre of the walking range	Existing studies	Study population	Type of facility	Temporal range	Unit of Analysis
facility	Capasso Da Silva et al. (2019)	all	provisioning, education, entertainment, healthcare, open space, civic institution	20 mins	parcel
	Ferrer-Ortiz et al. (2021)	all	provisioning, education, entertainment, healthcare, open space, transportation	5,10,15 mins	parcel
	Zhong et al. (2022)	vulnerable group	provisioning, education, entertainment, healthcare, pension	5,10,15 mins	point
housing	Ma et al. (2020)	all	education, entertainment, healthcare, transportation	15 mins	subdistrict
	Zhuo et al. (2020)	all	provisioning, education entertainment, healthcare, open space	15 mins	parcel

3. Methodology

3.1. STUDY AREA AND OVERALL FRAMEWORK

This study selected two low-income districts: Sham Shui Po (SSP) and Tin Shui Wai (TSW) in Hong Kong as study areas (Figure 1). SSP is located in the northwest of the Kowloon while TSW, a third-generation new town, is situated in the northwestern part of the New Territory. And the accessibility to public facilities for vulnerable people is evaluated from three aspects: spatial distribution characteristics, service population ratio and total number of facilities (Figure 2).

Figure 1. Study area and their location

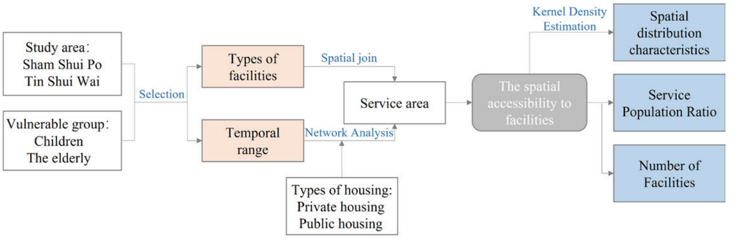


Figure 2. Methods and workflow

3.2. DATA PREPARATION

This paper uses the following data:

- Pedestrian network: The pedestrian network of SSP and TSW is obtained from the 3D pedestrian network in the Hong Kong Geodata Store.
- Building Footprint: The building footprint is obtained from the digital topographic map (iB1000) in the Hong Kong Geodata Store.
- Population for building groups: The number of children (younger than 15 years old) and the elderly (older than 65 years old) in building groups is collected from 2016 census data provided by Centamap Company Limited.
- Point of interest (POI): This paper selects the POIs of provisioning, education, healthcare, pension and entertainment (including POIs outside the study area but within 15-minute walking range from building groups). Through the Gaode map and Hong Kong Geodata Store, 6972 effective POIs were collected.
- Land Utilization in Hong Kong: This paper collects 2020 land utilization in Hong Kong (LUHK) data from Planning department and extract land use of open space and recreation in SSP and TSW.

3.3. METHODS AND ESTIMATION INDICES

3.3.1. Kernel Density Estimation

Kernel Density Estimation is a non-parametric method that "weights events within its sphere of influence according to their distance from the point" (Gatrell et al. 1996, p.259) so that the intensity is evaluated. This paper uses the kernel density estimation to analyse the clustering characteristics of public facilities in SSP and TSW.

3.3.2. Network Analysis

Network analysis is to make modelling of geographic network and urban infrastructure network which is commonly used in analysing traffic flows (Oh, 2007). First, the pedestrian road network was obtained from the Hong Kong geodata store. After breaking it, the length of each section was calculated and the walking speed was given to 53.3m/min(3.2km/hr) (The Department of Health in Hong Kong, 2016) to get the walking time of each section. Then, the points of building groups were imported and service areas were constructed (Figure 3). In the attribute of service area, the default interruption was set to 10 and 15 minutes. Finally, the service area buffers based on the existing pedestrian network were obtained.

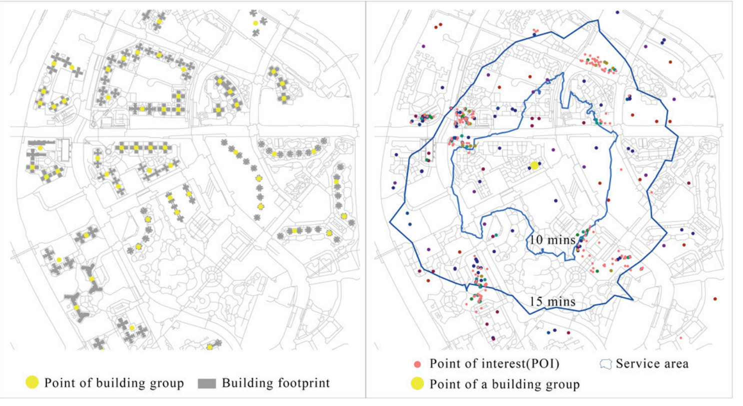


Figure 3. Service area for a building group

3.3.3. Construction of Index System

According to children's daily needs, this paper selects three major categories: provisioning, education, entertainment, and divides them into eight subcategories (Table 2). And based on the daily needs of the elderly, four major categories: provisioning, healthcare, pension and entertainment are chosen and divided into eleven subcategories (Table 3). The temporal ranges were set to be 10 or 15 minutes because of children's and the elderly's needs for short-distance travel. The number of facilities or areas of open space covered by service area was calculated. Then, Analytic Hierarchy Process (AHP) was used to determine the weight of each index and 12 experts from relevant fields were invited for scoring. The weights were calculated by yaahp software and consistency test was conducted. Finally, the final weights are shown in Table 2 and Table 3.

Table 2. Children's evaluation index system

Category	Subcategory	Temporal Range	Weights
Provisioning	Convenience store	10	0.1191
	Catering	10	0.0794
Education	Kindergarten	10	0.1716
	Primary school	10	0.1614
	Middle school	10	0.1215
Entertainment	Open space	10	0.1448
	Cultural facility	15	0.0888

	Sport facility	10	0.1134

Table 3. The elderly's evaluation index system

Category	Subcategory	Temporal Range	Weights
Provisioning	Supermarket	10	0.0955
	Convenience store	10	0.0911
	Market	10	0.0911
	Catering	10	0.0697
Healthcare	Hospital	15	0.1051
	Clinic	10	0.1051
	Pharmacy	10	0.0867
Pension	Day care centre and nursing home	10	0.1062
Entertainment	Open space	10	0.1041
	Cultural facility	15	0.0638
	Sport facility	10	0.0815

Service population ratio for children or the elderly is taken as an index to measure the proportion of children or the elderly that is served by a public facility in the specific temporal range according to the evaluation index system (Table 2, Table 3). The calculation formula is as follows:

$$\text{service population ratio of housing}(\%) = \frac{\text{population covered by service area}}{\text{total population of housing}} \times 100$$

$$\text{service population ratio of public housing}(\%) = \frac{\text{population of public housing covered by service area}}{\text{total population of public housing}} \times 100$$

$$\text{service population ratio of private housing}(\%) = \frac{\text{population of private housing covered by service area}}{\text{total population of private housing}} \times 100$$

Moreover, the total number of different types of facilities varies significantly. In order to make different types of facilities comparable, it is necessary to normalize the original data so that the results are in the range between 0 and 1. The calculation formula is as follows:

$$r'_{ik} = \frac{r_{ik} - r_{kmin}}{r_{kmax} - r_{kmin}}$$

where r_{ik} is the original value of the number of public facilities covered by service area (or area covered by service area for open space) of the k-th subcategory; r'_{ik} is the normalized value; r_{kmin} is the minimum value of data of k-th subcategory; r_{kmax} is the maximum value of data of k-th subcategory.

The evaluation value of spatial accessibility was obtained by multiplying the normalized value and weight of each public facility, and the total evaluation value was acquired by adding them. The calculation formula is as follows:

$$F = \sum_{i=1}^{n} r'_{ik} w_k$$

where F is the total evaluation value; n is the number of subcategories; r'_{ik} is the normalized value of each subcategory; W_k is the weight of each subcategory.

4. Results

4.1. SPATIAL DISTRIBUTION CHARACTERISTICS

It can be seen that provisioning and healthcare facilities in SSP tend to cluster along the Cheung Sha Wan Road with the highest degree of aggregation near the Dragon Centre and the metro station of Mong Kok respectively. And education, pension and entertainment facilities are evenly distributed. Compared to SSP, the density of facilities in TSW is relatively lower and distributed dispersedly. The provisioning and healthcare facilities in TSW are mainly clustered along the light rail line (Figure 4).

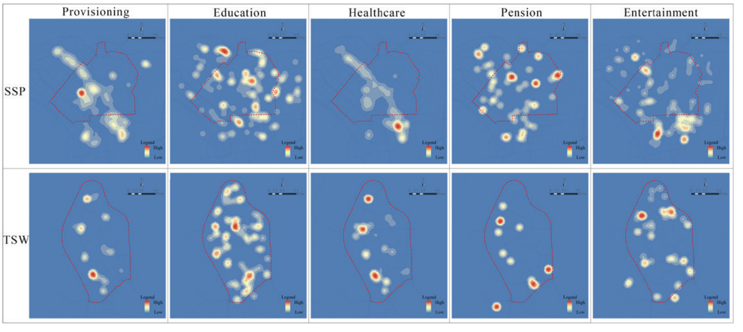


Figure 4. The spatial distribution characteristics of public facilities in SSP and TSW

4.2. ACCESSIBILITY ANALYSIS BASED ON SERVICE POPULATION RATIO

The service population ratio shows the relationship between demand and supply. By calculating the service population ratio, the inadequacy of facilities supply for vulnerable people can be accurately evaluated. As shown in Figure 5, all the service population ratios for children in SSP and TSW are higher than 90% and 80% respectively, indicating good accessibility in general. However, the service population ratios of private housing in TSW are significantly lower, especially for cultural facility, convenience store and sport facility.

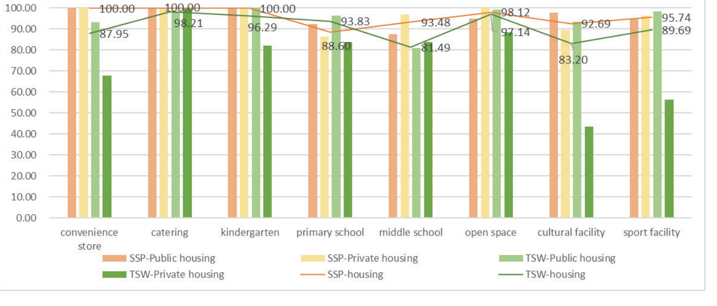


Figure 5. Service population ratio for children in SSP and TSW

Based on the service population ratio for the elderly (Figure 6), only 23.85% and 20.29% of the elderly in SSP and TSW can access to a hospital within 15 minutes by walking. Basically, the service population ratios of different facilities for the elderly in SSP are higher than 85% except hospital. Furthermore, the service population ratios for pension and cultural facility in private housing of TSW are much lower suggesting that their accessibility can be improved.

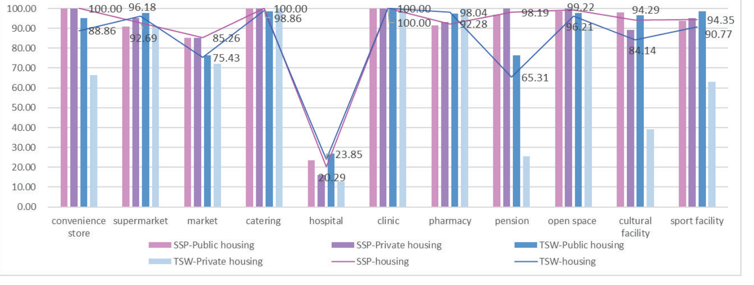


Figure 6. Service population ratio for the elderly in SSP and TSW

4.3. ACCESSIBILITY ANALYSIS BASED ON NUMBER OF FACILITIES

The accessibility analysis based on the total number of facilities shows the diversity of facilities for children or the elderly in their daily life. According to the total evaluation value for children, the private housing near the intersection of Cheung Sha Wan Road and Yen Chow Street has the highest score while the public housing of Fu Cheong Estate, Chak On Estate and Pak Tin Estate has lower scores in SSP. As for TSW, the public housing of Tin Heng Estate Block and Tin Yat Estate has a relatively higher score (Figure 7).

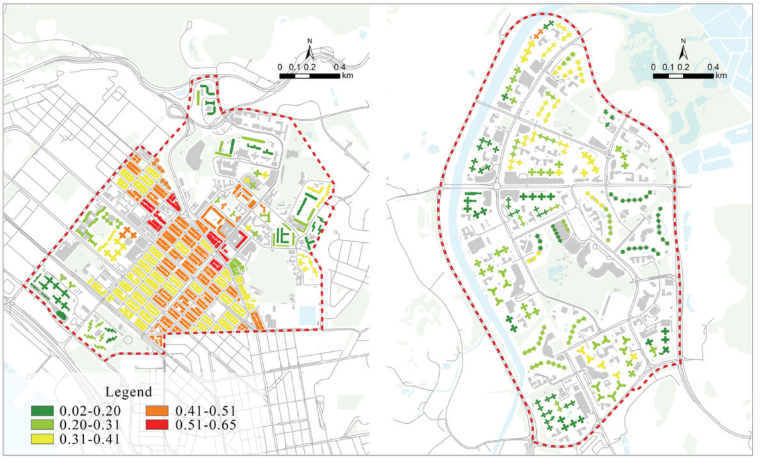


Figure 7. Total evaluation value for children in SSP and TSW

In terms of the total evaluation value for the elderly, the private housing between Yen Chow Street and Nam Cheong Street in SSP has the highest score. The majority of the total evaluation value for the elderly in TSW is at a lower level (Figure 8).

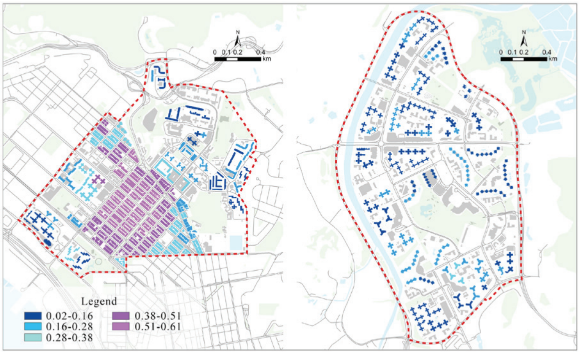


Figure 8. Total evaluation value for the elderly in SSP and TSW

The average evaluation value for both children and the elderly in the private housing in SSP is much higher than that in public housing in SSP and housing in TSW, indicating its good accessibility to public facilities (Figure 9).

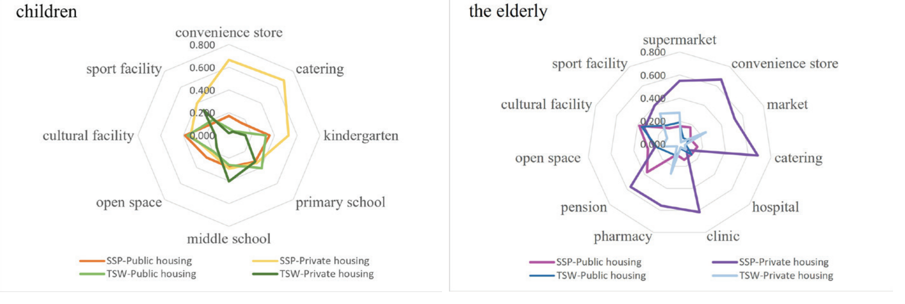


Figure 9. Average evaluation value for children and elderly in SSP and TSW

5. Discussion and Conclusion

The 15-Minute City is a human-centric concept aiming to improve the liveability of residents and tries to rethink the relationship between people, space and time. The evaluation of spatial accessibility to public facilities is the key to apply it to practice. And vulnerable people are supposed to be given prior attention. Consequently, this paper provides an evaluation framework for children and the elderly from three aspects: spatial distribution characteristics, service population ratio and total number of facilities. By comparing SSP and TSW, it can be seen that the accessibility to public facilities for children and the elderly in SSP is mostly higher than that in TSW. Among them, the accessibility to facilities of private housing in SSP is the highest, while that of private housing in TSW is the lowest.

According to the evaluation result, the spatial accessibility to public facilities for the underserved areas can be optimized by the following strategies:(1) Transforming the existing buildings and public spaces from monofunctional to multifunctional uses. (2) Establishing more decentralised local hubs that provide daily services for children and the elderly. (3) Improving the connectivity and safety of the pedestrian network.

This paper attempts to combine spatial accessibility with population and land use, but there is a limitation that the POI data cannot reflect the scale and capacity of the facilities. Additionally, we should pay more attention to the heterogeneity of people. It is expected that the needs of vulnerable group and their travel behaviour can be further explored, so as to optimize the distribution of public facilities and improve the pedestrian network.

References

Capasso Da Silva, D., King, D. A., & Lemar, S. (2019). Accessibility in practice: 20-minute city as a sustainability planning goal. *Sustainability*, 12(1), 129, https://doi.org/10.3390/su12010129.

Ferrer-Ortiz, C., Marquet, O., Mojica, L., & Vich, G. (2022). Barcelona under the 15-Minute City Lens: Mapping the Accessibility and Proximity Potential Based on Pedestrian Travel Times. *Smart Cities*, 5(1), 146-161, https://doi.org/10.3390/smartcities5010010.

Gatrell, A. C., Bailey, T. C., Diggle, P. J., & Rowlingson, B. S. (1996). Spatial point pattern analysis and its application in geographical epidemiology. *Transactions of the Institute of British geographers*, 256-274, https://doi.org/10.2307/622936.

He, S. Y., Tao, S., Ng, M. K., & Tieben, H. (2020). Evaluating Hong Kong's spatial planning in new towns from the perspectives of job accessibility, travel mobility, and work–life balance. *Journal of the American Planning Association*, 86(3), 324-338, https://doi.org/10.1080/01944363.2020.1725602.

Ma,W. J., Li, L, Gu, J. & Zhu, Y. G. (2020). Research on Spatial Layout and Accessibility of Infrastructural Public Service Facilities in the 15 Minute Life Circle, *Planners*, 36 (20), 11-19.

Moreno, C., Allam, Z., Chabaud, D., Gall, C., & Pratlong, F. (2021). Introducing the "15-Minute City": Sustainability, resilience and place identity in future post-pandemic cities. *Smart Cities*, 4(1), 93-111, https://doi.org/10.3390/smartcities4010006.

Oh, K., & Jeong, S. (2007). Assessing the spatial distribution of urban parks using GIS. *Landscape and urban planning*, 82(1-2), 25-32, https://doi.org/10.1016/j.landurbplan.2007.01.014.

Peters, A. (2020). Paris's mayor has a dream of 'the 15-minute city'. *World Changing Ideas*, Retrieved from https://www.fastcompany.com/90456312/pariss-mayor-has-a-dream-for-a-15-minute-city

Pozoukidou, G., & Chatziyiannaki, Z. (2021). 15-Minute City: Decomposing the new urban planning eutopia. *Sustainability*, 13(2), 928, https://doi.org/10.3390/su13020928.

The Department of Health in Hong Kong. (2015, January 1). The Exercise "Formula". *Health@work.hk Project*, Retrieved from: http://www.chp.gov.hk/archive/healthatwork/tc/contentf87f.html?MenuID=108

Zhong, J.H., Yi, F.R., He, Z.G., Cheng, G.R., & H, D. (2022). Research on Allocation of Community Service Facilities based on Evaluation of Individual Accessibility of Vulnerable People. *Journal of Geo-information Science*, 24(5),875-888, https://doi.org/10.12082/dqxxkx.2022.210496

Zhuo, B.L., Zhang, Z.M., Gao, H.Z., Wang, Z. &Liu, D.J. (2021). Research on the Convenience of 15-Minute Community Life Circles Based on the Road Network. *Urban Transport of China*, 19 (1),65-73+111

DIGITAL APPLICATION OF TYPO-MORPHOLOGY IN THE CONSERVATION AND RENEWAL OF HISTORIC AREAS

A Case Study of Hehuatang Historical and Cultural Block in Nanjing, China

ZHEHAO SONG[1], PENG TANG[2] and YACHENG SONG[3]
[1,2,3]*School of Architecture, Southeast University, Nanjing, 210096, China.*
[2]*Key Laboratory of Urban and Architectural Heritage Conservation (Southeast University), Ministry of Education, Nanjing, 210096, China.*
[1]*songzhehao1996@126.com, 0000-0003-3643-636X*
[2]*tangpeng@seu.edu.cn, 0000-0003-1658-6774*
[3]*song_yc@seu.edu.cn, 0000-0003-4605-0655*

Abstract. The conservation and renewal of historic areas are facing many complex and scattered problems, which are not suitable to be completed by a unified method. Designers tried to use the typo-morphology to analyse the morphology of each micro unit to carry out targeted conservation and renewal actions. However, any adjustment of spatial structure may affect the morphological characteristics of the whole block and each micro unit, designers need an efficient method to control the dynamic changes of the block in real-time. Based on the hierarchical structure of typo-morphology, a digital model of the historic areas was built. This model can be perfected as a morphological analysis tool and analyse the block's spatial morphological evolution during its conservation and renewal process. In the renewal design work of Hehuatang historical and cultural block in Nanjing, this method helps designers test each strategy's rationality and find a more scientific scheme to guide the further detailed design. The involvement of digital methods enables typo-morphology to assist design work more accurately and promote the working mode to gradually change from "experience-based artificial induction" to "data-based pattern extraction".

Keywords. Historic Area, Conservation and Renewal, Typo-Morphology, Hierarchical Structure, Digital Model

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 545-554. © 2023 and published by the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA), Hong Kong.

1. Introduction

1.1. TYPO-MORPHOLOGY METHOD IN THE CONSERVATION AND RENEWAL OF HISTORIC AREAS

Historic areas are the material carriers of the historical memory of a town. The conservation and renewal of historic areas are crucial to the improvement of the urban living environment and the inheritance of urban history and culture (Wang, 2022). However, historic areas often have a long history, so the spatial characteristics and property rights are very complex. Designers should pay more attention to the authenticity of the environment and the continuity of residents' lives in their actual work. For the above reasons, designers can't simply use a unified method to complete the design of the entire block but should carry out different conservation and renewal actions according to the morphological characteristics of each micro unit (Dong et al., 2021).

In this process, typo-morphology showed its value. Yao Sheng, Tian Yinsheng and others explored the morphological research methods of Chinese traditional settlements by using Conzen's theory of urban morphology (Yao et al., 2013). Fitrianty Wardhani and Samsul Bahri accomplish a comparison study between Bengkulu and Singapore with historical linkages in terms of identifying patterns and characteristics of its city morphology (Wardhani and Bahri, 2021). Han Dongqing and others paid more attention to practice and applied typological methods to the conservation and renewal of Nanjing Xiaoxihu historical and cultural block (Han, 2022). Designers can sort out the spatial structure of historic areas from the perspective of multi-scale hierarchical structure in the field of typo-morphology, and make a clear quantitative expression of the morphological characteristics of spatial elements to guide the design. However, the quantitative morphological analysis also brings a lot of computational work. Any adjustment of spatial structure in the process of these works may affect the morphological characteristics of the whole block and each micro unit, which makes designers need to constantly re-analysis. The conservation and renewal work needs an efficient method to control the dynamic changes of the block in real-time.

1.2. DIGITAL APPLICATION IN THE CONSERVATION AND RENEWAL OF HISTORIC AREAS

With the rapid development of information technology, scholars began to use digital methods to solve complex problems in the conservation and renewal of historic areas. Wu Jiayu, Lu Yutian and others used ridge regression and LightGBM with multi-source big geospatial data to explore whether urban morphological elements that affect the vitality of heritage and urban areas are consistent or have different spatial distributions and daily variations (Wu et al., 2022). Marra Adriana and Fabbrocino Giovanni built a crowd-based tool for indirect condition assessment and conservation of cultural heritage, this work has promoted the participation of diverse groups in the renewal design of historic areas (Marra and Fabbrocino, 2021). Wang Xiao, Tang Peng, and others develop a new method of feature analysis and generative design to regenerate the district in the study on Gunanjie Street, in Yixing, China. The proposed method generated the referable design schemes quantitatively and established generally accepted conservation plans and guidelines (Wang et al., 2019). Siew Leng Leong and Patrick Janssen proposed an innovative participatory design approach through a web application and assisted the co-design and co-decision process in heritage conservation (Leong and Janssen, 2022).

Through the above research, the digital method shows its great potential to enhance the scientific design, optimize the method flow and improve the work efficiency in the conservation and renewal of historic areas. This study built a digital model, which can calculate and store the analysis results of all micro units in the field of morphological typology. Based on this model, an analysis tool had been constructed and used in the conservation and renewal of the Hehuatang historical and cultural block in Nanjing. In this study, typo-morphology controls the dynamic changes of historic areas in real-time through digital methods. While the conservation and renewal work has received accurate support, the applicability of the theory has also been improved.

2. Research Method

2.1. THE HIERARCHICAL STRUCTURE AND MORPHOLOGICAL ANALYSIS OF HISTORIC AREA

The analysis of morphology always starts from the description of morphology (Han, 2013). The scale hierarchical structure in the field of typo-morphology provides a clear framework for the description of the morphology of historic areas. The spatial structure of Chinese historic areas is complex, but the spatial elements still show the characteristics of a hierarchical structure. This study follows Kropf's multi-level diagram and uses the hierarchical structure of " blocks - tissues - street and plots series - plots - areas and buildings " to recognize Chinese historic areas (Kropf, 2011) (Figure 1). Each historic area is composed of multiple simple tissues, a single street and the plots connected on both sides constitute a simple tissue. The plot also contains buildings and courtyards. Among them, tissues, plots, and buildings, as the most important three levels, are most closely related to the conservation and renewal of historic areas.

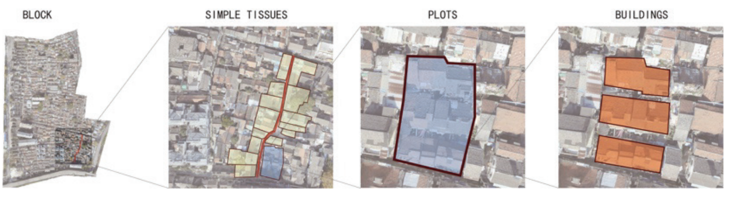


Figure 1. The hierarchical structure of Chinese historic areas

In this hierarchical structure, the research focuses on the characteristics of tissues, plots, and buildings in two aspects. On the one hand, the study analysed the morphological characteristics of space elements themselves, such as the length and width of roads, plot ratio of plots, building floor area ratio, etc. In this study, they are named Self Morphological Characteristics (Hereinafter referred to as SMC). SMC are helpful for designers to cognition and understand the spatial characteristics of renewal objects themselves, and respond with appropriate strategies. On the other hand, the study focuses on the connection between spatial elements and their relationship with the outside. This connection relationship is expressed quantitatively by the depth value, connectivity value and other values based on the space syntax (Hillier and Hanson, 1989). In this study, they are named External Connection Characteristics (Hereinafter referred to as ECC). ECC reflects the topological structure between the spatial elements in the block and can reflect the publicity and connection tightness of each spatial unit. It can help designers to give each spatial element corresponding functional positioning and development orientation. The specific morphological analysis framework based on the hierarchical structure is shown in Figure 2.

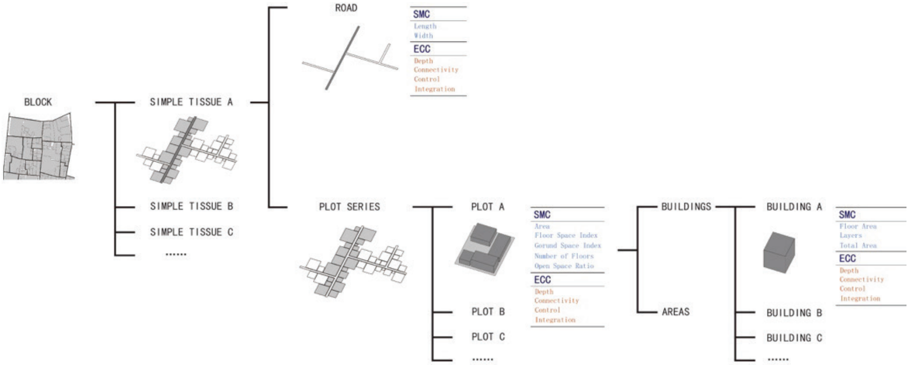


Figure 2. Morphological analysis framework based on the hierarchical structure

After obtaining the morphological analysis results corresponding to each spatial element, this study classifies the spatial elements according to the morphological characteristics. Classification can help to recognize the space through types and control the future design orientation. K-means clustering algorithm is introduced to determine the type, and all spatial elements at the same level are clustered according to their own SMC and ECC. The use of the clustering algorithm eliminates the human factors in manual classification and obtains more accurate type characteristics of the whole block.

Based on morphological typology, this study has formed a morphological cognition and analysis framework of "deconstruction-analysis-classification" of Chinese historic areas. This framework combed the relationship between spatial elements in the historic areas and presented the results of morphological analysis and type characteristics. In this framework, the connection between spatial elements is clear, and the characteristics are also expressed by quantitative values, which is very suitable for digital reproduction.

2.2. DIGITAL MODEL OF MULTISCALE HIERARCHICAL STRUCTURE

This study uses Java language programming to construct the corresponding digital model with the scale hierarchical structure as the basic framework. The classes in Java can be used to express things with the same attributes. Therefore, corresponding classes are defined in the digital model for the spatial elements at each level, and different objects of one class are used to represent multiple spatial elements of the same level. In addition to the morphological information of space elements at this level, the attributes of each class also include the objects of the space elements at the next level.

For example, the attributes of one plot object include the objects of buildings which are in this plot. This structure realizes the data linkage of all levels in the digital model (Figure 3).

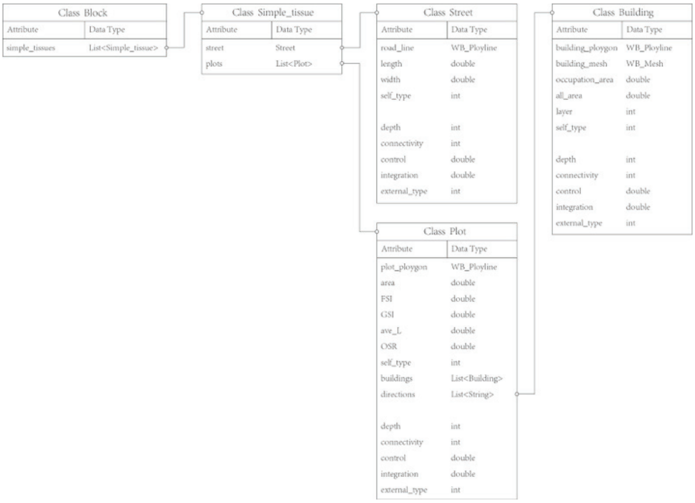


Figure 3. The structure of the digital model

The program cannot directly identify the historic areas' actual shape, so the geometric information of the spatial elements is extracted from a map file. The program reads the geometric information of each layer from the dxf file and translates it into corresponding class objects, which are stored in different data sets.

The morphological analysis and classification of spatial elements are completed in the digital model. In this study, the morphological analysis method is transformed into the corresponding geometric algorithm, and then into the Java program function. Each time the program runs, it can quickly calculate the morphological analysis results of all spatial elements at all levels. Then the program runs the K-means clustering algorithm and quickly gives the type characteristics of spatial elements at the same level according to the morphological analysis results.

In order to make full use of the real-time update and fast feedback characteristics of the program calculation, the functions of spatial structure adjustment and cycle update are also added to the digital model. The data sets of roads, plots and buildings in the mathematical model can be increased or decreased at any time. This function corresponds to the road network adjustment, plot division and building demolition in the process of conservation and renewal of historic areas. The digital model can be updated and calculated in real-time according to the adjusted spatial state, and new morphological analysis and classification results can be obtained. At this point, a digital model of historic areas has been constructed. It can store and link spatial elements information at multi-scale levels and update morphological analysis results in real-time according to changes in spatial structure (Figure 4).

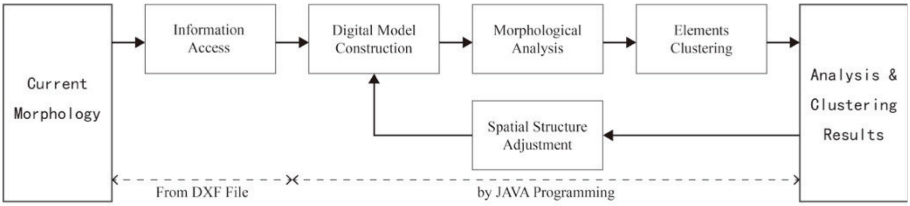


Figure 4. Operation process of the digital model

3. Analysis Tool Applied to Practice

3.1. MORPHOLOGICAL ANALYSIS TOOL BASED ON DIGITAL MODEL

After the completion of the digital model, a more direct medium is still needed to connect the model with the actual workflow, so that the analysis results can quickly assist the design. The digital model has an operation interface, thus becoming an interactive decision-making tool.

The centre of the tool interface is the general map of the block. The program draws the corresponding geometric figures according to the geometric information of each spatial element in the digital model and displays the morphological type of the spatial element through colours. The general map includes SMC and ECC modes. Each mode can be switched between the plot view or the building view, as well as the corresponding 2D or 3D perspective. In the tool interface, designers can view the detailed information of a space element according to the actual needs. The detailed data of a road, plot or building can be displayed in the right data bar to ensure that the morphological analysis results accurately guide the design (Figure 5).

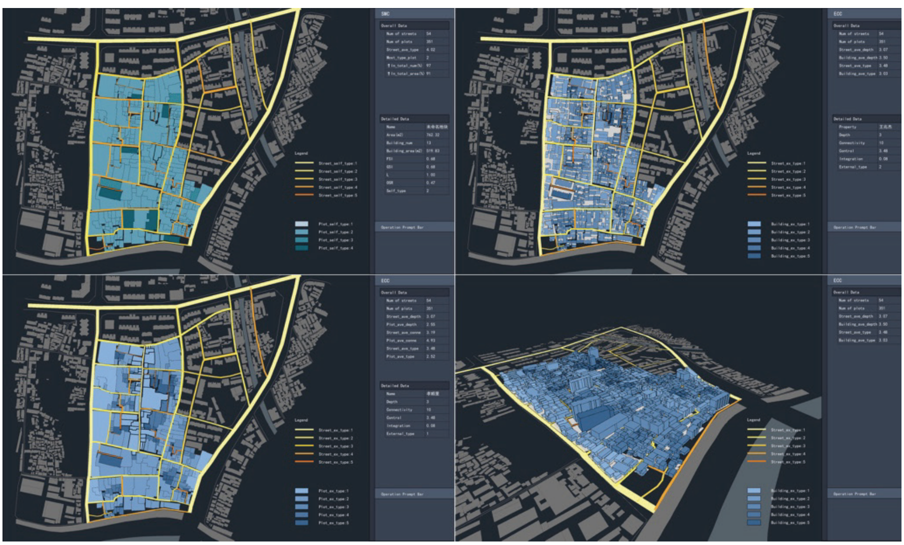


Figure 5. Multiple data display modes of the tool

The addition of the adjustment and update module is an important step for the tool to face the actual workflow. A buttoned operation panel is added on the left side of the tool interface, and the buttons are linked with the corresponding space adjustment and update methods in the digital model. The users can complete the adjustment of the morphology of the block through simple operations in the tool interface, including the increase or decrease of streets, merging, splitting, and adjusting the entry direction of plots, and the demolition of buildings. After each adjustment, the updated general analysis map can be obtained in real-time (Figure 6).

Figure 6. Complete tool interface with adjustment and update functions

On the one hand, the tool provides visual morphological analysis results, and users can view the real-time information of the whole block or a certain spatial element at any time to make more scientific design decisions. On the other hand, the tool highlights the characteristics of real-time updating of digital models. Users can obtain the analysis results in real-time during the design update process and conduct a rapid comparison to verify the rationality of the design strategy.

3.2. PRACTICE IN THE CONSERVATION AND RENEWAL OF HE-HUATANG

The method section introduces how to build a digital analysis tool based on the digital model, which is consistent with the actual workflow. This research is based on the actual work process of the conservation and renewal of Hehuatang historical and cultural block in Nanjing and verifies the effectiveness of the tool in practical work.

Located in Nanjing, Jiangsu Province, China, Hehuatang historical and cultural block has a history of hundreds of years and has always been an important residential area in the south of Nanjing. Hehuatang covers an area of about 12.57 hectares, and its interior is mostly one to three storey old houses. Due to its long history, its internal roads are narrow and its living environment is dilapidated and chaotic (Figure 7). In particular, the plots have become very fragmented due to the continuous redistribution of property rights, which makes the shape of the block very messy. Efficient shape analysis tools are urgently needed in the process of conservation and renewal.

In this study, several designers who participated in the renewal design of Hehuatang were selected to try out the tool based on the actual project design work. Firstly, the designers made operations such as adding roads, deleting roads, and resetting the road width in the tool interface. The road network of the block had been re-planned. The tool can provide new road network analysis results which were adjusted based on different design strategies so that designers can quickly compare and select multiple road planning schemes. After the new road network had been determined, the designers reorganized the morphology of all plots in the tool interface. Through the re-division and combination of the plots, the original plots divided based on property rights will be integrated into a form that is convenient for subsequent design updates. Finally, the designers determine the future functional orientation of each new plot based on its own morphological characteristics and get the new layout and function positioning map of the block. This map will be used as the working base map to guide the specific design scheme in the follow-up work (Figure 8).

Figure 7. Map and current environment of Hehuatang historical and cultural block

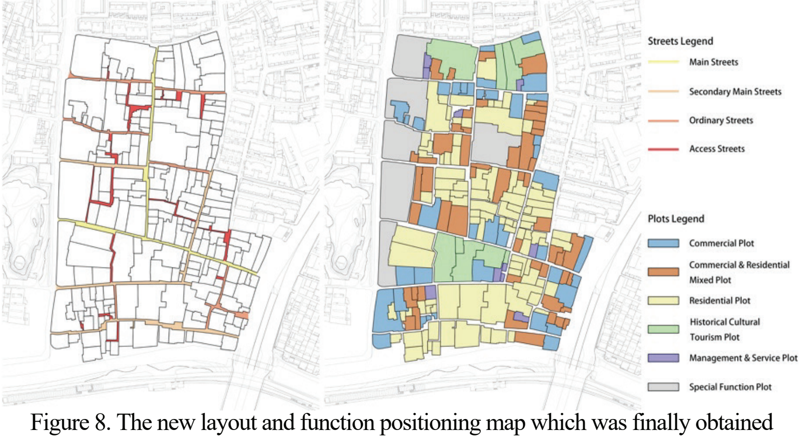


Figure 8. The new layout and function positioning map which was finally obtained

3.3. DISCUSSION ON APPLICATION PRACTICE

Through this practice, the application value of this morphological analysis tool has been confirmed. Designers can compare and choose road planning and plot division schemes more efficiently and scientifically.

The designers who participated in the practice also pointed out the limitations of the current method. Firstly, the conservation and renewal design of historic areas is not only based on the results of morphological analysis but also needs to consider more social and cultural issues. Therefore, the results of the morphological analysis can only be used as a partial reference for design. Secondly, the design is decided by the designer. However, through this method, designers can only make modifications according to their own preferences or morphological characteristics data. This process is still a subjective decision process with incomplete consideration. Finally, the essence of this research is still the analysis process before the design. The results which were provided do not involve the subsequent actual design process. In the follow-up work, the generative design method under the guidance of morphological analysis should be introduced. This research can be better applied to practice.

4. Conclusion

The conservation and renewal of historic areas is an eternal topic. The spatial structure in the historic area is fragmented and the social needs are complex. Designers need to carry out accurate morphological analysis of micro units, in order to complete the updated design more in line with the wishes of residents and other relevant users. Under this background, this study applies the theories of typo-morphology to the conservation and renewal of Chinese historic areas and constructs a morphological research process of "cognition-analysis-classification". Then the analysis framework is translated into a digital model, which can be expanded into a real-time morphological analysis tool. The involvement of digital methods enables the conservation and renewal of historic areas to obtain real-time and accurate morphological analysis references. The application field of digital methods has been expanded, and the theoretical research of typo-morphology and design practice are better integrated.

With the people-oriented design concept gradually becoming popular, the design scheme of historic areas is no longer based on the value judgment of a designer, but a more accurate result obtained after a comprehensive analysis of complex information. The morphological analysis must provide analysis reference for design in real-time, and data has become a powerful medium. The involvement of digital methods makes each step of the conservation and renewal of historic areas connected by data streams. The connection and interaction between morphological analysis and morphological design are formed. The transformation of the working mode from "experience-based artificial induction" to "data-based mode extraction" has become an important technical significance revealed by this study.

Acknowledgements

This work was supported by the National Natural Science Foundation of China under Grant 52178008.

References

Dong, Y., Han, D., & Huang, J. (2021). The Process-ness and Participation of Urban Design in the Small-scale and Gradual Conservation and Regeneration of the Xiaoxihu Historic Area in Nanjing. *Time+Architecture,* (01),51-55. doi:10.13717/j.cnki.ta.2021.01.010.

Han, D. (2013). Designing Cities:From Formal Understanding to Formal Design. *The Architect,* (04),60-65.

Han, D. (2022). Progresses of Interaction and Inclusion: The Practice of Conservation and Regeneration of Xiaoxihu Block in Nanjing. *Architectural Journal,* (1), 1–8. doi:10.19819/j.cnki.ISSN0529-1399.202201001.

Hillier, B., & Hanson, J. (1989). *The social logic of space*. Cambridge university press.

Kropf, K. (2011). Morphological investigations: Cutting into the substance of urban form. *Built Environment,* 37(4), 393-408, DOI:10.2148/benv.37.4.393.

Leong, S. L., & Janssen, P. (2022). Participatory Planning: Heritage Conservation Through Wang, J. (2022). On Professional Problems of Multi-scale Protection of Urban Architectural Heritage in China. *City Planning Review,* (06),7-24.

Co-design and Co-decision. In *27th International Conference on Computer-Aided Architectural Design Research in Asia: Post-Carbon, CAADRIA 2022* (pp. 505-514). The Association for Computer-Aided Architectural Design Research in Asia (CAADRIA), https://doi.org/10.52842/conf.caadria.2022.2.505

Marra, A., & Fabbrocino, G. (2021). Crowd-based tools for indirect condition assessment and conservation of cultural heritage. In *Digital Heritage. Progress in Cultural Heritage: Documentation, Preservation, and Protection: 8th International Conference, EuroMed 2020, Virtual Event, November 2–5, 2020, Revised Selected Papers* (pp. 38-50). Cham: Springer International Publishing.

Wang, J. (2022). On Professional Problems of Multi-scale Protection of Urban Architectural Heritage in China. *City Planning Review,* (06),7-24.

Wang, X., Tang, P., & Shi, X. (2019). Analysis and Conservation Methods of Traditional Architecture and Settlement Based on Knowledge Discovery and Digital Generation-A Case Study of Gunanjie Street in China. In *24th International Conference on Computer-Aided Architectural Design Research in Asia: Intelligent and Informed, CAADRIA 2019* (pp. 757-766). The Association for Computer-Aided Architectural Design Research in Asia (CAADRIA)

Wardhani, F., & Bahri, S. (2021). Comparative Study of the Patterns and Characteristics Urban Morphology of the Old City, Bengkulu, and Singapore that Has Relation to Historical Background. In *AUC 2019: Proceedings of the 15th International Asian Urbanization Conference, Vietnam* (pp. 553-570). Springer, Singapore, DOI:10.21163/gt_2020.151.34.

Wu, J., Lu, Y., Gao, H., & Wang, M. (2022). Cultivating historical heritage area vitality using urban morphology approach based on big data and machine learning. *Computers, environment and urban systems,* 91, 101716, https://doi.org/10.1016/j.compenvurbsys.2021.101716.

Yao, S., Tian, Y., & Chen, J. (2013). Theory of Urban Morphological Regionalization and its Application in Heritage Protection. *City Planning Review*, (11), 47–53+66.

THE MANY FACES OF THE METROPOLIS

Unsupervised Clustering Of Urban Environments In Mumbai Based On Visual Features As Captured In City-wide Street-view Imagery

ROHIT P. SANATANI[1] and TAKEHIKO NAGAKURA[2]
[1,2]*Massachusetts Institute of Technology (MIT)*
[1]*sanatani@mit.edu, 0000-0002-6273-5745*
[2]*takehiko@mit.edu, 0000-0002-3219-3930*

Abstract. The larger visual identity of a city is often a blend of smaller and distinct visual character zones. Despite the recent popularity of street-view imagery for visual analytics, its role in uncovering such urban visual clusters has been fairly limited. Taking Mumbai as a demonstrative case, we present what is arguably the first city-wide visual cluster analysis of an Indian metropolis. We use a Dense Prediction Transformer (DPT) for semantic segmentation of over 28000 Google Street View (GSV) images collected from over 7000 locations across the city. Unsupervised k-means clustering is carried out on the extracted semantic features (such as greenery, sky-view, built-density etc.) for the identification of distinct urban visual typologies. Through iterative analysis, 7 key visual clusters are identified, and Principal Component Analysis (PCA) is used to visualize the variance across them. The feature distributions of each cluster are then qualitatively and quantitatively analysed in order to examine their unique visual configurations. Spatial distributions of the clusters are visualized as well, thus mapping out the different ‘faces’ of the city. It is hoped that the methodology outlined in this work serves as a base for similar cluster-based inquiries into the visual dimension of other cities across the globe.

Keywords. Unsupervised Clustering, Urban Environments, Visual Features, Character Zones, Street-view Imagery

1. Introduction

Different neighborhoods within cities are known to possess unique visual identities. In Mumbai for example, the lush greenery of Powai, the vast seafront at the Marine Drive, and the dense congestion of Kalbadevi all exist in worlds of their own. Specific visual features play major roles in the way in which these localities are perceived, and also strongly influence the lasting ‘mental images’ that are formed in the minds of citizens and visitors (Nasar 1990). Moreover, the larger visual identity of a city is often shaped by these smaller visual clusters, each of which 'look' a certain

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 555-564.

way. In many cases, the visual appearance of a neighbourhood also impacts a variety of urban indicators, such as real-estate value, perceptions of safety (Naik et al. 2014), tourism potential (Kazemi et al. 2011) and the like.

While visually similar environments across different parts of a city are often well known to the long-time resident, there have been very few empirical approaches towards the clustering of distinct visual character zones in cities. Geolocated street-view imagery such as Google Street View (GSV) has been widely used for urban visual analytics research in the recent past (Biljecki et al. 2021), but its use for such tasks has been fairly limited. In an Indian context, the absence of street view imagery till recent times has been a significant impediment to such research directions.

Taking Mumbai as a demonstrative case, we present what is arguably the first city-wide visual character analysis of an Indian city. GSV imagery has only very recently been made available for multiple cities in India, thus providing valuable visual data for such an inquiry. This work applies state of the art computer vision models for the extraction of semantic visual features from street-level GSV images, and uses unsupervised clustering methods to uncover and analyse the different visual character zones that emerge from different parts of the metropolis.

2. State of the art: Visual analytics for urban environments

With growing availability of urban visual data over the past decade, empirical inquiries into the visual dimension of cities are becoming increasingly popular. Advances in computer vision have allowed for the extraction of common urban features from two-dimensional urban imagery. Semantic segmentation models such as DeepLab (Chen et al. 2017), PsPNet (Zhao et al. 2017) and more recently Dense Prediction Transformers (DPT) (Ranftl et al. 2021) have allowed for the rapid city-level cataloguing of semantic elements (such as greenery, sky, buildings, automobiles, people etc.) in the visual field. Datasets such as CityScapes (Cordts et al. 2014) have provided valuable training material in this regard. Moreover, instance segmentation (object detection) models trained on datasets such as ADEK20K (Zhou et al. 2017) or OpenImages (Kuznetsova et al. 2020) allow for occurrence/count-based extraction of such semantic elements.

Such methods, used in conjunction with geolocated street-view imagery collected at specific distance intervals along streets, have been used for urban infrastructure analysis, greenery mapping, health and well-being, morphological analysis, mobility, and real-estate analytics (Biljecki et al. 2021). Within the domain of urban perception, multiple studies have focused on the perceived quality of streetscapes (Zhang et al. 2018, Ye at al. 2019). Perceptual responses collected from human subjects have also been used to train supervised learning models that are capable of predicting specific perceptual attributes of urban environments, such as safety, beauty etc. (Naik et al. 2104, Dubey et al. 2016). Moreover, the impact of visual features on human physiological response have also been studied (Gorgul et al. 2019, Ojha et al. 2019).

Despite such advances in urban visual analytics, city-level unsupervised clustering of urban environments based on high-level visual features is relatively unexplored territory. Unsupervised methods have been used in the past for morphology-based urban clustering (Moosavi 2017), and also for uncovering visual features of interest for specific urban anchors such as tourist attractions (Sanatani et al. 2022). Applying similar methods to city-level visual data is thus a worthwhile endeavour. Specifically in an Indian context, the recent availability of street-view imagery at a high spatial resolution opens up significant methodological potential.

3. Methodological framework

The larger methodological framework for this work involves the following steps - (i) Visual data collection - downloading GSV imagery at specific intervals along the street network of the city, (ii) Visual feature extraction - the extraction of high-level semantic elements contained in the imagery, (iii) Unsupervised clustering of urban locations based on visual features, and (iv) Cluster analysis - analysing the visual character and geographical distribution of the different visual clusters.

3.1. VISUAL DATA COLLECTION

For city-level collection of visual data, the street network configuration for Mumbai was downloaded from the Open Street Maps database using the osmnx network analysis library for Python. Download locations were then generated for every 250 m intervals along the street network. The panoIDs (street view image IDs) were then programmatically queried for each of these locations using the Google Street View API. A list of all valid panoIDs, along with their latitude and longitude values were stored as a .csv file. PanoIDs for a total of 7104 locations were collected across the city. Four separate street view images (corresponding to heading 0, 90, 270 and 360) were then programmatically downloaded for each panoID, using the API. For all images, pitch was maintained at 0, ensuring that the camera axis was parallel to the horizontal plane. A total of (7104 x 4 = 28416) images were downloaded **(Figure 1)**.

Figure 1. Street view query locations (left) and representative samples (right)

3.2. VISUAL FEATURE EXTRACTION

For extraction of semantic visual features from the downloaded street-view imagery, a pre-trained Dense Prediction Transformer (DPT) (Ranftl et al. 2021) was used. A transformer model was preferred to traditional Convolutional Neural Networks (CNNs) because of significant improvements in prediction accuracy. The DPT had been trained on the ADE20K dataset (Zhou et al. 2017), and was able to segment each image pixel into one of 150 semantic classes, such as tree, sky, building, automobile, person, road etc. Given the large volume of city-wide visual data, processing time was a significant factor. The model was run on a Python environment in a Linux based host machine with GPU, and took around 10.5 hours to complete. For each image, the percentage of pixels belonging to each class was computed and compiled as a .csv file. **Figure 2** below depicts representative street view imagery, and corresponding segmentation using the DPT model.

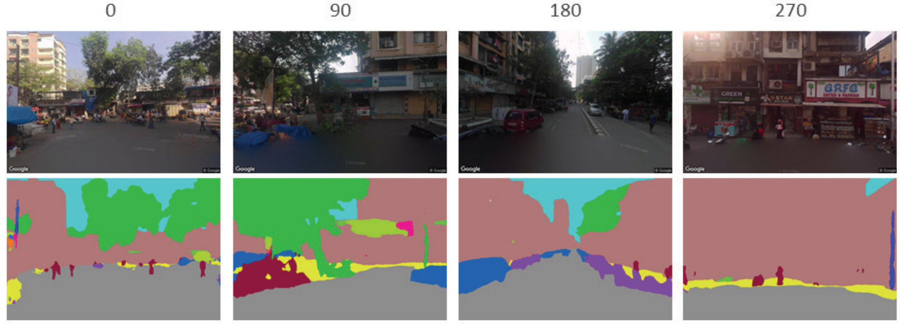


Figure 2. Representative street-view imagery (top) and segmented results (bottom).

3.2.1. Visual feature mapping

Figure 3. Visual feature maps for key semantic classes - greenery (left), built (centre) and sky (right)

The extraction of geolocated visual content information allowed for the mapping of visual features across Mumbai for exploratory analysis. For a given location (panoID), the mean feature scores for each class were computed across the 4 images collected from that point. These scores were saved as a .csv file and served as the primary dataset for all further visualisation and analysis. Visualisation was carried out in QGIS, to examine the spatial distribution of semantic features. **Figure 3** shows feature maps for key semantic classes such as greenery, built density and sky view.

3.2.2. Principal component analysis

For exploratory analysis, it was necessary to not only visualise the spatial distribution of specific features, but also examine the distribution of data points within the feature-space. For this, each data point (location) was converted into a 150-dimensional vector, corresponding to the feature scores for the semantic classes. Principal Component Analysis (PCA) was then carried out on the feature matrix, and three principal components extracted for the dataset. This lower dimensional feature-space was then visualised, and the visual character of datapoints from different parts of the space qualitatively evaluated. It was apparent from this exercise, that the principal components were closely aligned with greenery, built density and sky-view. **Figure 4** shows the 3-dimensional principal component feature space, and representative image samples from different parts of the same after clustering.

3.3. UNSUPERVISED CLUSTERING BASED ON VISUAL FEATURES

The primary aim of this body of research was to identify visual character zones across the city, based on the high-level visual features extracted from each location. An unsupervised k-means clustering approach was taken forward in this regard. Clustering was carried out using the scikit-learn library. Choosing an appropriate cluster count was a critical step in this regard. A low cluster count would club together visually diverse environments within a single cluster, while a high cluster count would result in multiple clusters with similar visual character. Different cluster counts between 5 and 20 were iteratively experimented with, and the qualitative visual similarities/differences between them in the 3D feature space (after PCA) were evaluated. This also allowed for the identification and removal of invalid images (such as indoor environments). Based on this exercise, a cluster count of 7 was selected, as it corresponded to an optimum visual diversity between clusters.

3.4. CLUSTER ANALYSIS

Cluster analysis was carried out on two levels. Firstly, the visual character of each cluster was both quantitatively and qualitatively evaluated. To do this, the cluster centroids were computed for each cluster, and a K-Nearest Neighbours (KNN) algorithm applied to identify the $k = 5$ samples from the database which were nearest to these centroids. These samples within a cluster were thus the most representative of the visual character that the cluster embodied. Quantitative evaluation entailed an examination of the average feature distribution for these representative samples, as well as the distribution across the entire cluster. Qualitative evaluation entailed visualising the images for the representative samples, and comparing the visual character of these samples across clusters.

On a second level, the geographical distribution of these clusters across the urban landscape of the city was examined. The data points were mapped out spatially, and colour-coded by cluster. The pattern of cluster distribution was analysed qualitatively and quantitatively. For qualitative evaluation, the correlation between the cluster distribution and specific neighbourhoods/quarters or functional zones of the city were analysed. For quantitative evaluation, the degrees of dispersion of these clusters were looked into using quadrat analysis (Rogers 1969). The geographic extent of the map was uniformly divided into discrete bins, and the number of samples in each bin computed. The variance-mean ratio (vmr) for each cluster provided a quantitative metric for relative dispersion. Clusters which were widely dispersed across bins (low vmr) represented commonly occurring character zones that are to be found across the city, while clusters which were concentrated in fewer bins (high vmr) represented visually homogenous 'pockets' which could be easily localised on the ground.

4. Results - The many faces of Mumbai

4.1.1. Visual feature distributions

Based on an analysis of the visual feature distributions, clusters corresponding to 7 distinct visual character zones were identified. In line with the principal components identified earlier, most of the variance between these zones were primarily a result of varying proportions of visible built density, sky view, greenery and automobiles/humans (congestion/crowd). The character zones were assigned labels A through G, and the feature distributions analysed for the most representative samples (cluster centroids).

Figure 4 presents a visual as well as quantitative summary of the different clusters. It is worth noting that clusters A and G are the largest, and are thus representative of the most commonly occurring visual character in the city. This is in line with Mumbai's popular visual identity comprising dense buildings and traffic, but with visible urban greenery along streets. Cluster A (dense buildings with significant greenery) blends into cluster F within the PCA feature space, which comprises lush green avenues with buildings partially visible through vegetation. Cluster G (dense buildings with little greenery) blends into cluster C, which comprises locations where buildings dominate the visual field, with almost no greenery.

In counterpoint to the greenery-built axis lies the sky-view axis, with cluster B occupying the other end of the PCA feature space. This cluster comprises of wide thoroughfares where the visual field is dominated by the sky, road and automobiles. Cluster D comprises samples which closely resemble cluster B, but gradually blend into cluster F, where much of the sky view is replaced by greenery. Along the sky-built axis, the overarching dominance of the sky reduces as we move away from cluster B, and buildings begin to slowly frame the visual field (cluster E). As we move further along the PCA space, the sky diminishes further and samples emerge from cluster G and finally C.

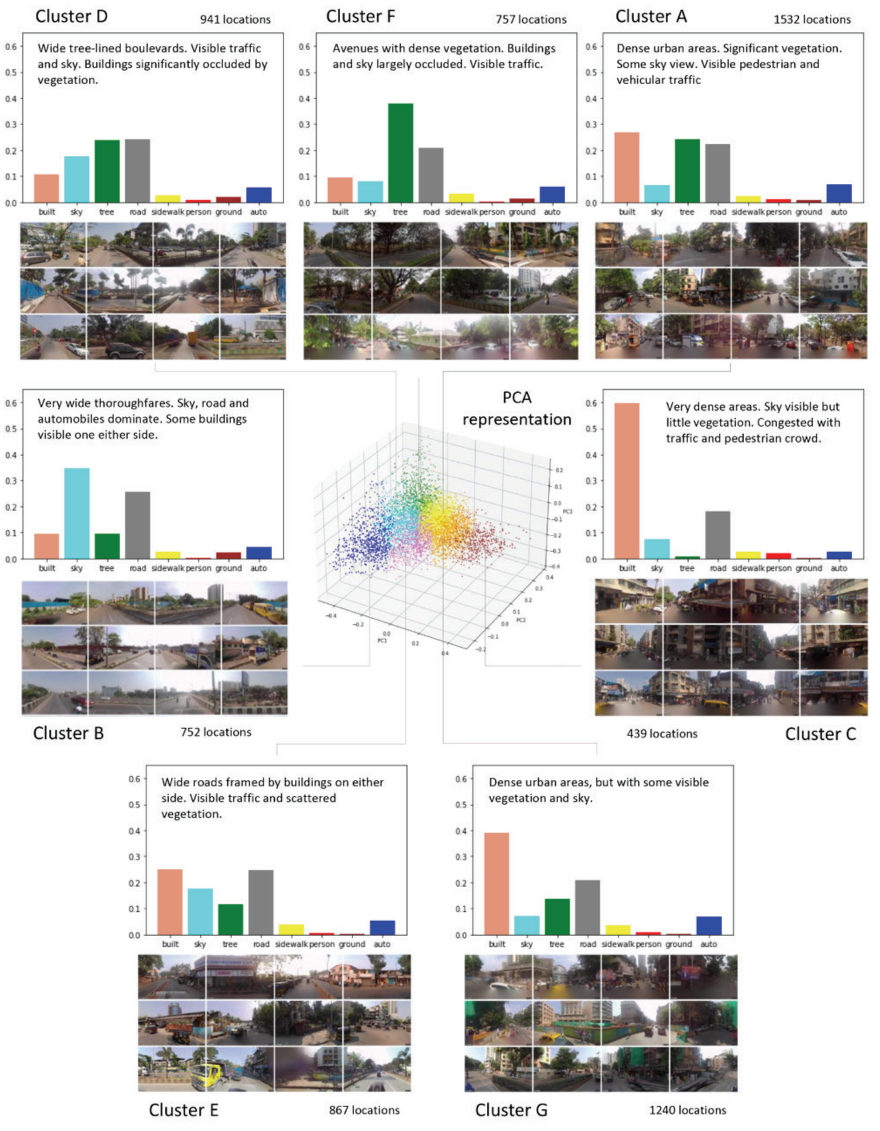


Figure 4. Summary of visual clusters with most representative samples and feature distributions

4.1.2. Spatial distribution

As discussed earlier, it was important for us to evaluate not only the visual feature distributions of the clusters, but also the spatial distribution of the same. Such an analysis allows for the identification of visual character zones that are spatially localised as well. **Figure 5 (left)** depicts the locational distribution of the clusters.

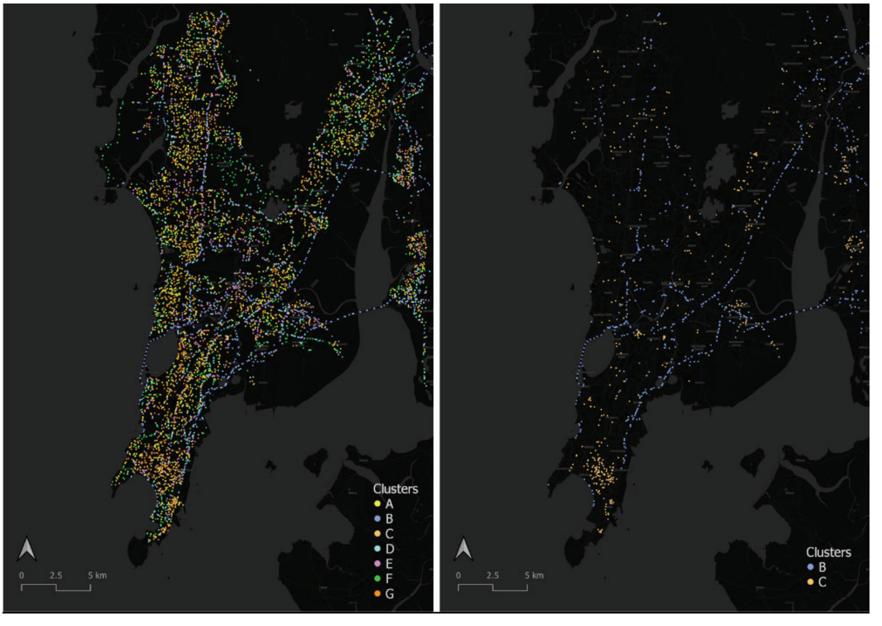


Figure 5. Spatial distribution of all clusters (left) and clusters B and C in isolation (right)

We see that the largest clusters (A and G) are spatially dispersed across the city, with no notable geographical clustering (vmr = 7.78 and 8.42 respectively). This reinforces our previous observation of them being representative of the overall visual identity of the city. Samples from these clusters occur commonly across different neighbourhoods and different functional zones, thus dominating popular mental images of the city. Clusters D and E are similarly dispersed even though they represent fewer samples. Moreover, there are certain localities which only comprise cluster F (dense greenery) and no other visual character - the Arey colony and Madh areas being notable examples. Most such zones however lie along the outskirts of the metropolis.

The visual character zones which indeed exhibit strong spatial patterns are Clusters B and C (Figure 5, right). Cluster B, as discussed earlier, comprises wide thoroughfares with sky views. While physically dispersed across Mumbai (vmr = 7.14), these locations correspond to the major highways that run through the city. Multiple samples trace the trajectory of the Eastern and Western express highways - two major north-south thoroughfares through the city. Most of the Mumbai-Satara highway, connecting to Navi (New) Mumbai, the Bandra-Worli sealink, and the iconic Marine Drive fall within this cluster as well. The samples of cluster C (dense buildings with little/no greenery) on the other hand, exhibit strong spatial clustering (vmr = 16.01), a major concentration lying in Lower Mumbai, in and around the neighbourhoods of Kalbadevi, Bhulshwar and Chor Bazaar. It is worth noting that the 'chawl' building typology (high density tenement housing) which is unique to the city, commonly occurs in these areas. Another set of samples emerge from Dharavi, one of the largest informal settlements in the country. As a matter of fact, an overwhelming majority of the samples in this cluster emerge from dense informal settlements across the city which, in many cases, also correlate with lower income inhabitants. Visual character often has socio-economic dimensions, and this cluster is perhaps a representative example.

5. Reflections: Towards new paradigms of urban visual analytics

While many of the findings from this study empirically validate common knowledge for those familiar with the city, the demonstrated analytical methodology can indeed be applied to any city across the world. Despite recent developments in urban visual analytics, the use of unsupervised clustering for visual character analysis is a novel approach. Moreover, the added value of employing unsupervised methods in this context lie in the fact that such methods rely purely on visual feature data to unravel patterns amongst sampled locations in a bottom-up manner - patterns which are often extremely valuable to designers and planners. This allows for meaningful qualitative analyses of many complex perceptual and affective attributes, the structured and quantitative representation of which is extremely challenging within supervised-learning frameworks.

There are however a number of key limitations of this methodology that are worth touching upon. While GSV imagery provides access to city-wide visual data, it remains largely restricted to streets with vehicular access. Pedestrianised zones, large urban plazas, and very narrow lanes in dense neighbourhoods are thus excluded from such datasets, despite being major contributors to urban visual identity. Moreover, such images only capture a location at a specific time of day. Temporal variations in visual character, which form an important part of the Indian urban experience, are thus not captured. Night-time character is also entirely excluded from such analyses. On the computer-vision front, the models used for visual feature extraction are trained to only identify predefined high-level semantic classes. Nuanced visual qualities outside such a predefined vocabulary are often missed.

Having said that however, such a methodology is extremely useful for multiple research directions. As discussed earlier, visual character often correlates with urban indicators such as real-estate price, safety and tourism potential, and such an approach can be taken forward for focused inquiries into such correlations. The methodology is also valuable from an experiential standpoint. While every city has a unique overall 'image', such an image is seldom homogenous. In most cases, a diverse set of spaces and places come together in the city, distributed in unique proportions and unique configurations. The larger visual identity of the city is thus rooted in each of these micro identities. The methods demonstrated through this work allow us to empirically uncover and analyse the nuances of such micro-identities - thus systematically unravelling the rich and diverse faces of the metropolis.

References

Biljecki, F., & Ito, K. (2021). *Street view imagery in urban analytics and GIS: A review.* Landscape and Urban Planning, 215, 104217. https://doi.org/10.1016/j.landurbplan.2021.104217

Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2017*). Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 40(4), 834–848.

Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., Franke, U., Roth, S., & Schiele, B. (2016). *The cityscapes dataset for semantic urban scene understanding*. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 3213–3223.

Dubey, A., Naik, N., Parikh, D., Raskar, R., & Hidalgo, C. A. (2016). *Deep learning the city: Quantifying urban perception at a global scale*. European Conference on Computer Vision, 196–212.

Gorgul, E., Chen, C., Wu, K. K., & Guo, Y. (2019). *Measuring street enclosure and its influence to human physiology through wearable sensors*. In Adjunct Proceedings of the 2019 ACM Int. Joint Conference on Pervasive and Ubiquitous Computing and Proceed. of the 2019 ACM International Symposium on Wearable Computers (pp. 65-68).

Kazemi, M., Pour, S., Saadat, F., & Bitaraf, F. (2011). *The effect of tourist's mental image on the perceived value of coastal cities in the Caspian Sea*.

Kuznetsova, A., Rom, H., Alldrin, N., Uijlings, J., Krasin, I., Pont-Tuset, J., Kamali, S., Popov, S., Malloci, M., Kolesnikov, A., & others. (2020). *The open images dataset v4*. International Journal of Computer Vision, 128(7), 1956–1981.

Moosavi, V. (2017). *Urban morphology meets deep learning: Exploring urban forms in one million cities, town and villages across the planet*. ArXiv e-prints, page. ArXiv Preprint ArXiv:1709.02939.

Naik, N., Philipoom, J., Raskar, R., & Hidalgo, C. (2014). *Streetscore-predicting the perceived safety of one million streetscapes*. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 779–785.

Nasar, J. L. (1990). *The evaluative image of the city*. Journal of the American Planning Association, 56(1), 41–53.

Ranftl, R., Bochkovskiy, A., & Koltun, V. (2021). *Vision transformers for dense prediction*. Proceedings of the IEEE/CVF International Conference on Computer Vision, 12179–12188.

Rogers, A. (1969). *Quadrat analysis of urban dispersion: 1. Theoretical techniques.* Environment and Planning A, 1(1), 47–80. https://doi.org/10.1068/a010047

Sanatani, R., Nagakura, T., & Tsai, D. (2022). *The Tourist's Image of the City: A data-driven comparative analysis of the features and topics of tourist interest across three global cities.* Proceedins of the Iberoamerican Society of Digital Graphics, 2022.

Ye, Y., Zeng, W., Shen, Q., Zhang, X., & Lu, Y. (2019). *The visual quality of streets: A human-centred continuous measurement based on machine learning algorithms and street view images.* Environment and Planning B: Urban Analytics and City Science, 46(8), 1439–1457. http://dx.doi.org/10.1177/2399808319828734

Zhang, F., Zhou, B., Liu, L., Liu, Y., Fung, H. H., Lin, H., & Ratti, C. (2018). *Measuring human perceptions of a large-scale urban region using machine learning.* Landscape and Urban Planning, 180, 148–160. https://doi.org/10.1016/j.landurbplan.2018.08.020

Zhou, B., Zhao, H., Puig, X., Fidler, S., Barriuso, A., & Torralba, A. (2017). *Scene parsing through ade20k dataset*. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 633–641.

MULTISOURCE ANALYSIS OF BIG DATA ON STREET VITALITY USING GIS MAPPING AND DEEP LEARNING

A Case Study of Ding Shu, China

JINZE LI[1] and PENG TANG[2]
[1,2]*School of Architecture, Southeast University, Nanjing, China.*
[1]*lijinze0106@gmail.com, 0000-0001-6725-3318*
[2]*tangpeng@seu.edu.cn, 0000-0003-1658-6774*

Abstract. Urban vitality is the driving force behind sustainable urban development. As the most frequently used public space in cities, the enhancement of street vitality is of great significance for improving human-centred habitats. Based on multi-source big data, this study uses spatial and statistical analysis methods to explore the impact factors of street vitality. Through the quantitative evaluation of these factors, we propose corresponding strategies to enhance the vitality of the street. Firstly, the spatial elements of streets are extracted using deep learning algorithm based on the acquired street view images. Further, the impact factors of street vitality are demonstrated using statistical methods by combining multi-source data. We established an evaluation system based on the impact factors of street vitality, which can quantify and predict street vitality. In this way, we can propose vitality enhancement strategy for the street with lower vitality in a targeted approach. The feasibility of the process is demonstrated by using Ding Shu as an example. This study provides a basic framework for a people-centred approach to enhance street vitality based on big data. It also contributes to causal inference in urban problems.

Keywords. Multi-source Data, Street Vitality, Deep Learning, Spatial Analysis, Statistical Analysis, Causal Inference, People-centred City

1. Introduction

In this new era, with the problems such as the decline of urban space quality and environmental degradation emerging (Chen et al., 2019), urban development worldwide is beginning to shift from incremental development to built-up area redevelopment that advances broader social well-being. The development and renewal of cities are about reshaping the urban form and providing better places and spaces for residents and their social activities (Von Schönfeld and Bertolini, 2017). In this context, constructing a human-centred smart city is an effective way to improve the quality of urban space and implement sustainable development. Data-driven can be applied to assess and improve urban spaces from a multi-dimensional perspective. As the most

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 565-574.

typical public space (Mehta, 2013), streets are essential in everyday life and enable daily travel, recreation, and socializing. Street vitality is an essential indicator to evaluate the liveability of urban space. Street vitality has two attributes (Li et al., 2022): external representation and internal composition. Improving street vitality is decisive in improving urban residents' quality of life and making the urban living environment more liveable (Mehta, 2019).

Street vitality has been discussed since the 1960s as an essential issue in urban research. In traditional studies, questionnaires and field research methods are often used to gather information about spatial elements and human behavioural activities (Mehta, 2019; Handy et al., 2002; Gehl, 1987). With the emergence of geographic information and computer science, the accuracy of street vitality assessment has improved considerably. Multi-source data has provided an opportunity to measure how people use and quantitatively experience street space (De Nadai et al., 2016; Sulis et al., 2018). Increasingly, Street View Image (SVI) has been widely accepted as an effective way to quantify street space (Kelly et al., 2013; Shen et al., 2018; Ye et al., 2019).

Existing studies have revealed that street vitality often co-presents with the combination of street accessibility, road density, and functional diversity according to statistical analyse (Marcus, 2010; Oliveira, 2013; Ye and van Nes, 2014). Most researchers mainly focused on the correlation between elements (Ye et al., 2019; Song et al., 2020; Cong et al., 2022; Jiang, 2022). However, correlation does not prove the causality between them (Pearl, 2000). Hence, there needs to be more research on the driving mechanism of street vitality. In response, this study explores the causal relationship between the representations of street vitality and the intrinsic constituent elements through statistical analysis.

The study focuses on Ding Shu Town, Yixing, Jiangsu Province, China, as a case study area to test a data-driven approach to extract the spatial elements of streets and evaluate their vitality from multiple dimensions (fig. 1). The specific study covers an area of 1532 streets, which were first analysed through GIS to assess the morphological characteristics of the streets. By requesting Baidu Map API, 35000 street view images in the study area were obtained. By comparing three neural network models (Dilated ResNet-105, Ademxapp Model A1, multi-scale Context Aggregation Net) pre-trained on Cityscapes Data sets, the improved Dilated ResNet-105 is finally used for feature recognition and semantic segmentation of the street view images. The 3D spatial elements of the streets, such as the amount of greenery, sidewalk space, and pedestrians, are extracted. Further, with the help of GIS spatial analysis and SPSS statistical analysis software, the distribution of people in space is used as the external representation of vitality, and the spatial elements of different dimensions are used as the internal composition to explore the internal action mechanism of street spatial vitality.

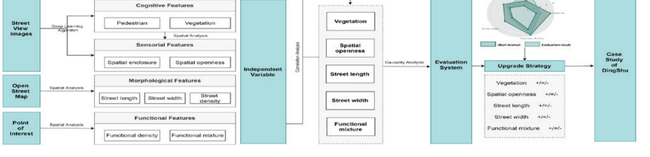


Figure 1. Research framework

2. Study Area and Data

2.1. STUDY AREA

Ding Shu is a typical Chinese town in Jiangsu Province, China, producing world-renowned Yixing clay teapots. After a long historical evolution, Ding Shu has developed a diverse street system. The street space presents a variety of forms and functions, which is an essential support for creating and maintaining street vitality. Due to the complexity of connotations and the diversity of external forms of the built environment in Ding Shu, it is a typical example to research street vitality. The study revolved around the central area of Ding Shu, which covers 1532 streets.

2.2. DATA SOURCE

2.2.1. Street Network and Buildings

The morphological data of street networks and buildings were mainly obtained from an open-source map, OpenStreetMap (OSM). Considering that there may be missing data in remote areas, Baidu map data was used for additional proofreading.

The original street network data were cleaned, topologically checked, and processed. And a total of 1532 streets were obtained.

2.2.2. Point of Interest (POI)

The POI (Point of Interest) data was collected from Baidu Map. Similar to Google Maps, Baidu Map is one of China's most widely used internet map service providers. Moreover, it is also the fastest-updated map in the China region.

A total of 12,920 POIs of 15 categories for the day of 10 May 2022 in Ding Shu were collected by sending requests via Python to the Baidu Map API. Each POI data contains the name, latitude, longitude, address, category, and other information of the geographical entity.

2.2.3. Street View Image (SVI)

Street View Image (SVI) has recently been widely accepted as an effective way to quantify the built environment in the street (Kelly et al., 2013; Shen et al., 2018; Ye et al., 2019).

Based on the street network, the sampling points were generated at 20m intervals according to the centre lines. A total of 17,720 sampling points were in the central area of Ding Shu. The coordinate of each point can be obtained in the GIS platform, and panorama images with a 180-degree view were collected by sending Uniform Resource Locator (URL) parameters to the Baidu API through a GET request. The acquired images have unique IDs and latitude and longitude coordinates. Furthermore, the size was set to 960×720. After data cleaning, 16,250 valid SVIs can be used in this study (fig. 2).

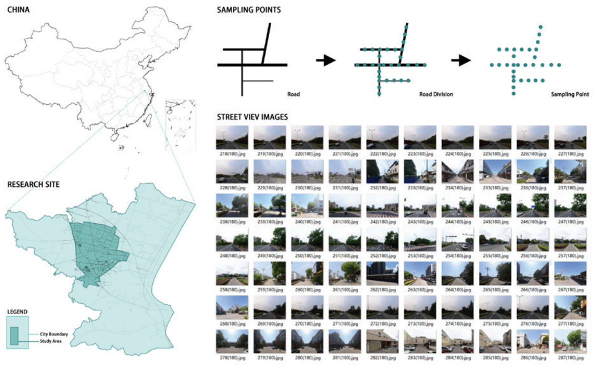


Figure 2. Study area and Baidu Street View Image acquisition

3. Methodology

3.1. SPATIAL ELEMENT EXTRACTION

Image semantic segmentation method based on Convolutional Neural Networks (CNN) was used to segment the acquired streetscape images to extract the spatial elements of streets. Three pre-trained neural network models for semantic segmentation of urban street scenes based on the Cityscapes dataset, which is a collection of 25,000 annotated images for semantic understanding of urban street scenes, were compared, including Dilated ResNet-105, Ademxapp Model A1, and Multi-scale Context Aggregation Net. As we can see from Figure 3, The Ademxapp Model has weaknesses in recognition of the sky, while the Multi-scale Context Aggregation Model is not accurate in the recognition of distant people. So, Dilated ResNet-105 was chosen to apply segmentation of SIV based on its accuracy in recognising street scenes in Ding Shu.

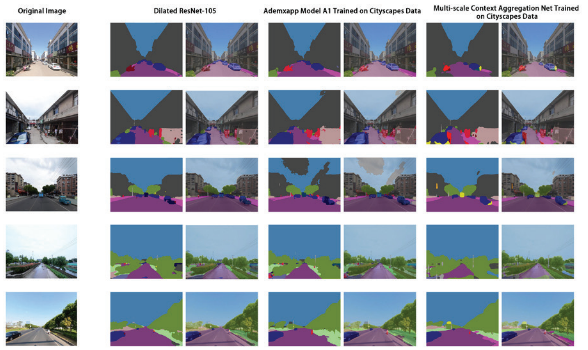


Figure 3. Comparison of neural network recognition of street view images

The images were segmented into 13 elements closely related to human perception and behaviour, and nine elements directly related to human spatial perception were selected for further analysis. The environmental elements extracted in this study include "roads, pavements, buildings, walls, plants, sky, people, cyclists, cars and others". The percentage of these elements can be easily calculated by the count of pixels (fig. 4).

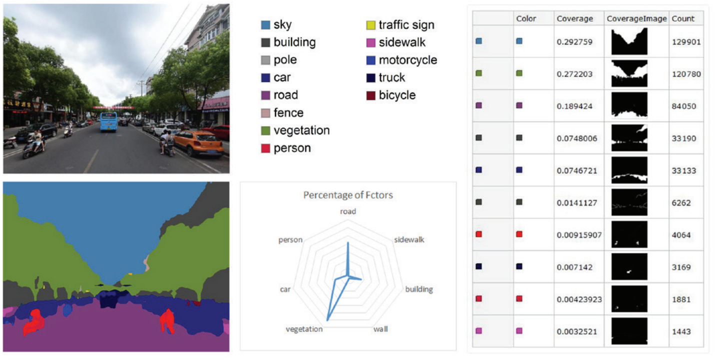


Figure 4. Semantic segmentation of street view images

3.2. PARAMETER SYNTHESIS OF STREET VITALITY

3.2.1. Dependent Variable

Jacobs believes street vitality comes from enough people on the street (Jacobs, 1961). So, the dependent variable is people in the street, which represents street vitality.

The count of people is calculated from the street view images, which can be mapped into each sampling point.

3.2.2. Independent Variable

The dependent variable has four dimensions, including a cognitive dimension, a perceptual dimension, a morphological dimension and a functional dimension (tab. 1).

The cognitive dimension features focus on the spatial elements of the street that are visually visible to the human eye, with statistics and analysis of the percentage of pavements and green views that are directly visible after semantic segmentation of the streetscape images.

The sensorial dimension features focus on the spatial perception of people in the street space, analysing the spatial enclosure and spatial openness of the streetscape images and other elements that affect people's spatial perception.

The morphological dimension features focus on the impact of morphological elements such as road length, road width and road network density, which are directly related to human use of the road, on the public nature of the street space from a macro perspective.

The functional dimension features focus on the density of available services and the mixture of functions from the point of the service functions they can provide.

Main	Secondary	Indicators and Formula	
Cognitive Features	Pedestrian	$Pedestrian = pixr\_sw$	$pixr\_sw$ The pixel ratio of the pavements
	Vegetation	$Vegetation = pixr\_gre$	$pixr\_gre$ The pixel ratio of the plants
Sensorial Features	Spatial enclosure	$SpatilEnclosure = \frac{pixr\_ver}{pixr\_hor}$	$pixr\_ver$ The pixel ratio of vertical elements, including buildings, walls and plants $pixr\_hor$ The pixel ratio of horizontal elements, including roads and pavements
	Spatial openness	$SpatialOpenness = pixr\_sky$	$pixr\_sky$ The pixel ratio of the sky
Morphological Features	Street length	$StreetLength = L_i$	L_i The length of a single street segment
	Street width	$StreetWidth = pixr\_rd + pixr\_sw$	$pixr\_rd$ The pixel ratio of the roads
	Street density	$StreetDensity = \frac{\sum_{i=1}^{n} L_i}{S_A}$	L_i The length of a single street segment S_A The buffer area of 300 m on both sides of the road section
Functional Features	Functional density	$FunctionalDensity = \frac{sum\_fuc}{L_i}$	sum_fuc The total number of POIs within 55 m on both sides of the street segment
	Functional mixture	$Number\ of\ categories = -\sum_{x=1}^{n} \frac{p\_num_x}{sum\_fuc} \log \frac{p\_num_x}{sum\_fuc}$	p_num_x Number of POIs of category x n The number of category of POIs

Table 1. Independent variables and data sources

4. Results and Discussion

4.1. DISTRIBUTION CHARACTERISTICS OF STREET

As shown in Figure 5, streets in Ding Shu are mainly in the form of small blocks and a dense road network. Intuitively, from a morphological and functional point, the street density and functional density of DingShu gradually decrease around the centre of Ding Shan.

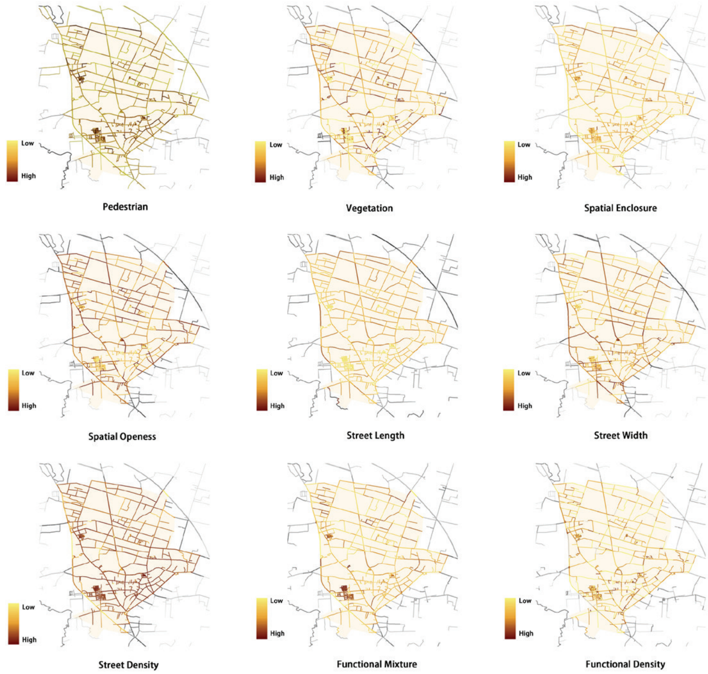


Figure 5. Evaluation of independent variables in the study area

4.2. CORRELATION AND CAUSALITY ANALYSIS OF STREET VITALITY

4.2.1. Correlation Relationship

To verify the relationships between spatial characteristics of street and street vitality, Correlation Analysis was used to examine the correlation between the independent variables and the dependent variable.

As shown in Table 2, significant differences have been observed between spatial characteristics of street and street vitality for spatial openness, vegetation, street length, street width, and functional mixture values. The spatial openness, street length and street width have negative correlation with the street vitality, while the vegetation, functional mixture have positive correlation with the street vitality.

Table 2. Correlation regression

	Vegetation	Pedestrian	Spatial Enclosure	Spatial Openness	Street Width	Street Length	Functional Density	Functional Mixture
Street Vitality	.185** (.003)	-.001 (.965)	.035 (.177)	-.169*** (.000)	-.347** (.009)	-.162** (.008)	-.012 (.639)	.218*** (.000)

After excluding irrelevant variables, Regression Analysis was performed on the independent and dependent variables. The Durbin-Watson test was used to test autocorrelation between independent variables. The result indicated that there is no correlation between sample data.

4.2.2. Causality Inference

Causality inference is generally used to explain the direction and degree of influence of the independent variable on the dependent variable. Using methods from econometrics, we selected an instrumental variable (IV) to effectively address the errors associated with the endogeneity issue. In general, geographic space is often chosen as the instrumental variable that is stochastically independent at a particular level of analysis but closely related to human behaviour and its social consequences. In this study, the distance of the sampling point from the nearest bus stop (Dist_Bus) was used as an instrumental variable in the street vitality analysis.

Two-stage Least Squares (2SLS) was used to explore the two-way influence of the independent and dependent variables through two linear regressions. The first stage of the regression estimates the independent variables with a two-way effect, and the second stage analyses the related questions. The result of 2SLS Regression Analysis in this study can be seen in Table 3. The F-statistic for the weak instrumental variables test considering heteroskedasticity was 38.16, indicating that the instrumental variables were valid (tab. 3).

Table 3. Instrumental variable regression

Explanatory Variables	Street Vitality	Vegetation	Spatial Openness	Street Width	Street Length	Functional Mixture
Dist_Bus		.241**	.354*	.606*	.638*	.259**
Control variable	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.629					
F	38.16	38.16	38.16	38.16	38.16	38.16
N	1519	1519	1519	1519	1519	1519

The study shows that the vitality of the street is closely related to the appropriate spatial openness, higher spatial greenness, street length and width, and higher functional mixture. However, spatial openness and street scale are not responsible for street vitality directly. The result demonstrates that increased functional mixture and greenness will significantly improve street vitality.

A wind rose was created diagram to evaluate the vitality of any street based on the indicators proven to influence its vitality. The individual indicators are normalised and mapped between 0-10 to facilitate comparisons between indicators. The top 20% of the most vibrant streets were selected by ranking the street vitality to extract their values for each impact indicator. Further, based on the extracted indicators, we established indicator intervals for the ideal state as shown in Figure 6.

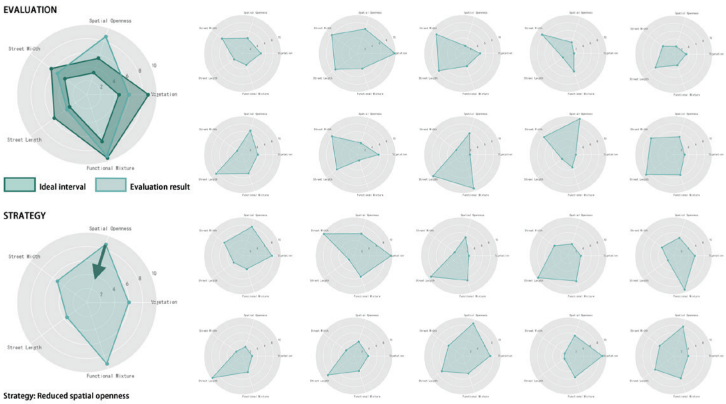


Figure 6. Multidimensional evaluation and optimization strategies for street vitality

In this way, it is possible to evaluate and predict the vitality of any street. And depending on the deviation of the actual state from the ideal state, a strategy for optimising the vitality of any street can be proposed in a targeted approach.

5. Conclusion

This paper examines the causal relationship between street elements and street vitality from an architectural perspective by using multi-source big data, providing a basis for decision-making at the micro level to rejuvenate the street and provide a liveable human environment. This result provides a way to evaluate street vitality and explores the intrinsic mechanism of street space vitality. However, the limitations of this study are as follows.

- The count of people in the street view images is used to represent street vitality. However, people's change in the temporal dimension were not considered.
- Not all possibilities for influencing street vitality were considered in selecting the dependent variable, and more variables may require further research.
- Future work will systematically discuss the relationship between endogenous issues such as street grades, street types, and street vitality. In this way, we can avoid differences in street vitality due to the classification of streets themselves.

References

Chen, M., Gong, Y., Lu, D., & Ye, C. (2019). Build a people-oriented urbanization: China’s new-type urbanization dream and Anhui model. Land use policy, 80, 1-9. https://doi.org/10.1016/j.landusepol.2018.09.031

Cong, C., Kwak, Y., & Deal, B. (2022). Incorporating active transportation modes in large scale urban modeling to inform sustainable urban development. Computers, Environment and Urban Systems, 91, 101726. https://doi.org/10.1016/j.compenvurbsys.2021.101726

De Nadai, M., Staiano, J., Larcher, R., Sebe, N., Quercia, D., & Lepri, B. (2016, April). The death and life of great Italian cities: a mobile phone data perspective. In Proceedings of the 25th international conference on world wide web (pp. 413-423).
environment affects physical activity: views from urban planning. American journal of preventive medicine, 23(2), 64-73. https://doi.org/10.1016/s0749-3797(02)00475-0
Gehl, J. (1987). Life between buildings (Vol. 23). New York: Van Nostrand Reinhold.
Google Street View to audit the built environment: inter-rater reliability results. Annals of behavioral medicine: a publication of the Society of Behavioral Medicine, 45 Suppl 1(Suppl 1), S108–S112. https://doi.org/10.1007/s12160-012-9419-9.
Handy, S. L., Boarnet, M. G., Ewing, R., & Killingsworth, R. E. (2002). How the built
Jacobs J. (1961). The Death and Life of Great American Cities. Random House, New York, NY, USA.
Jiang, Y., Han, Y., Liu, M., & Ye, Y. (2022). Street vitality and built environment features: A data-informed approach from fourteen Chinese cities. Sustainable Cities and Society, 79, 103724. http://dx.doi.org/10.1016/j.scs.2022.103724
Kelly, C. M., Wilson, J. S., Baker, E. A., Miller, D. K., & Schootman, M. (2013). Using
Li, Y., Yabuki, N., & Fukuda, T. (2022). Exploring the association between street built environment and street vitality using deep learning methods. Sustainable Cities and Society, 79, 103656. https://doi.org/10.1016/j.scs.2021.103656
Marcus, L. (2010). Spatial capital. Journal of Space Syntax, 1(1), (pp. 30-40).
Mehta, V. (2013). The street: a quintessential social public space. Routledge. https://doi.org/10.4324/9780203067635
Mehta, V. (2019). Streets and social life in cities: a taxonomy of sociability. Urban Design International, 24(1), 16-37. https://doi.org/10.1057/s41289-018-0069-9
Neuberg, L. (2003). CAUSALITY: MODELS, REASONING, AND INFERENCE, by Judea Pearl, Cambridge University Press, 2000. Econometric Theory, 19(4), 675-685. https://10.1017/S0266466603004109
Oliveira, V. (2013). Morpho: a methodology for assessing urban form. Urban morphology, 17(1), 21-33.
Shen, Q., Zeng, W., Ye, Y., Arisona, S. M., Schubiger, S., Burkhard, R., & Qu, H. (2018). StreetVizor: Visual Exploration of Human-Scale Urban Forms Based on Street Views. IEEE transactions on visualization and computer graphics, 24(1), 1004–1013. https://doi.org/10.1109/TVCG.2017.2744159
Song, X. P., Richards, D. R., He, P., & Tan, P. Y. (2020). Does geo-located social media reflect the visit frequency of urban parks? A city-wide analysis using the count and content of photographs. Landscape and Urban Planning, 203, 103908. https://doi.org/10.1016/j.landurbplan.2020.103908
Sulis, P., Manley, E., Zhong, C., & Batty, M. (2018). Using mobility data as proxy for measuring urban vitality. Journal of Spatial Information Science, 16, 137-162. https://doi.org/10.5311/JOSIS.2018.16.384
Von Schönfeld, K. C., & Bertolini, L. (2017). Urban streets: Epitomes of planning challenges and opportunities at the interface of public space and mobility. Cities, 68, 48-55. https://doi.org/10.1016/j.cities.2017.04.012
Ye, Y., & Van Nes, A. (2014). Quantitative tools in urban morphology: Combining space syntax, spacematrix and mixed-use index in a GIS framework. Urban morphology, 18(2), 97-118.
Ye, Y., Richards, D., Lu, Y., Song, X., Zhuang, Y., Zeng, W., & Zhong, T. (2019). Measuring daily accessed street greenery: A human-scale approach for informing better urban planning practices. Landscape and Urban Planning, 191, 103434. https://doi.org/10.1016/j.landurbplan.2018.08.028

GEOGRAPHIC INFORMATION SYSTEM BASED ANALYSIS ON WALK-ABILITY OF COMMERCIAL STREETS AT GROWING STAGE

YOONJUNG CHOI[1]and HYUNSOO LEE[2]
[1,2]*Yonsei University, Seoul, South Korea.*
1choi_2021@yonsei.ac.kr, 0000-0002-1796-6835
[2]*hyunsl@yonsei.ac.kr, 0000-0003-0023-88270*

Abstract. From the perspective of human-centered urban planning, walkability is a crucial concept for enhancing the quality of the neighborhood environment experienced in day-to-day life. Commercial facilities have the greatest impact on a walkable neighborhood environment. However, few studies analyzed the walkable environment's characteristics in consideration of local businesses' economic growth. This study aims to classify commercial areas according to vitality level and to analyze the correlation between store density and walkability factors through a case study on Seongsu district, Seoul, a commercial district where small businesses are growing. First, a Geographic Information System (GIS) based hotspot analysis is performed using the commercial area vitality index to select a target area for the case study. Second, through the Seongsu district case study, the walkability features of the cluster at the street level are evaluated and compared based on 3D (density, diversity, design). The results show that store density is correlated with walkability factors in growing commercial areas, and that there are distinct spatial differences depending on the factors. Based on this study's results, it is possible to propose a combination of a multi-use main street, a commercial street close to life, and a specialized street adjacent to green spaces.

Keywords: Walkability, Commercial Area, GIS, Hotspot Analysis, Density Based Clustering, Multi-source Data

1. Introduction

As the urban planning paradigm evolves from quantitative growth to a human-centered approach, several studies on walkability are conducted as a crucial concept for improving the quality of the neighborhood environment experienced in day-to-day life. Based on the previous studies, Forsyth (2015) classified walkability into three frameworks: environmental conditions, outcomes, and proxies, in order to reduce the ambiguity of terms. According to Wang and Yang (2019, p. 43), walkability is "a measure of whether the built environment of a neighborhood encourages people to walk." Consequently, the essential aspects for walkability analysis involve defining the conditions for a walkable environment and developing quantitative and qualitative evaluation methods for those conditions.

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 575-584.

As variables affecting neighborhood walkability, the 3Ds (density, diversity, design) suggested by Cervero and Kockelman (1997) are the predominantly used framework in previous studies. This concept is in line with Jacobs' (1961) assertion that the concentration of various destinations and the small-sized blocks create vitality in the streets. Commercial facilities have the greatest impact on a neighborhood environment's walkability. Providing various goods and services that people need nearby is one of the most useful ways to enliven streets in urban design. Mehta (2007) emphasized the role played by private business owners in creating walkable streets and encouraging social interaction. Through a study of metropolitan Washington, Leinberger and Alfonzo (2012) discovered that the performance in various economic indicators, such as sales and real estate prices, was better in clustered walkable districts than in stand-alone walkable places.

Self-reporting surveys, observational surveys, Geographic Information System (GIS) analysis using big data, and image vision analysis using Google Street View have been developed for measuring walkability. A number of studies have utilized GIS that evaluates the built environment using objective data. Connectivity, proximity, residential density, and land mix are the main factors evaluated through GIS analysis (Tong et al., 2016). As the need for a more detailed measurement and application of walkability increases, research has recently been conducted to subdivide the scope of analysis into three stages based on spatial scales; macro-, meso-, and micro-scale. In order to build a walkable environment, it is important to consider the factors of the street level as well as the city-level factors (Koo et al., 2022). Interaction with pedestrians is more emphasized at the micro-scale, such as buildings and street furniture, rather than the macro-scale.

Since the vitality of commercial streets plays an important role in creating a pedestrian-friendly neighborhood environment, decisions should be made based on clear measuring tools when planning the layout of commercial streets at the city level. A number of studies have highlighted the importance of commercial facilities in creating a walkable environment; however, few studies have analyzed walkability in light of the economic growth of local businesses. Most studies did not consider the difference in vitalization levels of commercial areas and selected the study site at the researcher's discretion. This study aims to conduct a case study to develop a model that classifies a commercial area's growth stage based on the level of the commercial area's vitality and analyzes the walkability characteristics of each growth stage. The research is undertaken as outlined below: First, a GIS based hotspot analysis is performed using commercial area vitality index to select a target area for the case study. Second, through the Seongsu district case study, the walkability features of the cluster at the street level are evaluated and compared based on 3Ds.

2. Methods

2.1. STUDY AREA

Seoul, a metropolitan city with an area of 605.96km2, has three Central Business Districts in Jongno, Gangnam, and Yeouido and seven subcenters. There are 1,211,053 businesses in Seoul, with 41.14% of them being restaurants, retail, and daily service businesses. In this study, classifying the spatial hierarchy according to the level of vitality of commercial areas in Seoul, the study area where small businesses are growing was identified. While considering the actual distribution of commercial facilities compared to administrative boundaries, the entire area of Seoul was divided into a 400 m by 400 m grid, and the map consisting of a total of 5156 cells was used as the study unit, excluding areas where data did not exist.

2.2. CALCULATING COMMERCIAL AREA VITALIZATION

2.2.1. Assessing Commercial Area Vitality Index (CVI)

The Commercial Area Vitality Index(CVI) was used to select the case study area. Lee et al. (2020) derived the CVI by integrating multiple indicators of growth and decline of commercial districts based on the Multiple Decline Index (Park and Kim, 2010). This study introduced the CVI calculation method by modifying it into a research-appropriate indicator, as in formula (1). The weight calculation using factor analysis was performed using the SPSS 26.0 tool.

$$CVI_j = \sum_{k=1}^{k} w_k \cdot z_{jk} \qquad \begin{aligned} CVI_j &= Total\ CVI\ of\ cell\ j \\ w_k &= Weight\ of\ index\ k \\ z_{jk} &= \text{Z-score of index } k \text{ of cell j} \end{aligned} \tag{1}$$

Based on a literature review, five indicators of commercial vitality were selected: commercial facilities density, the land value, the number of usage conversion, sales, and the de facto population. Commercial facilities density and the number of usage conversions from residential to commercial indicates the concentration of commercial investment. The land value represents the value obtained through land use (O'Sullivan, 1996). Sales show how much consumption occurs. This study used de facto population data, rather than floating population, to include both the resident and working populations.

Using a raw dataset provided by the Seoul Metropolitan Government (SMG), the indicators were collected as of September 2022 via publicly accessible open portal. Restaurants, retail, and life services were selected for the Point-of-Interest (POI) dataset; however, businesses that do not operate stores for general consumers, such as e-commerce and door-to-door sales, were excluded. The building dataset comprised location, usage, site area, and total floor area information. The land prices per 1m2 stated annually by the autonomous district was indicated in the land value dataset. Sales data was estimated using credit card consumption information in census boundary units. The de facto population is all the population existing in a specific area and at a specific point in time, estimated by SMG and KT using public big data and telecommunication data. This study refined and used de facto population data between 9 AM and 10 PM on weekdays.

2.2.2. Hotspot Analysis with Getis-Ord Gi Algorithm*

The Getis-Ord Gi* algorithm of ArcGIS Pro 3.0.2 was used to perform the hotspot analysis for the entire area of Seoul divided into 5156 grid cells. Getis-Ord Gi* is a spatial statistic that determines whether a specific spatial unit's attribute data is clustered with high or low values based on a spatial weight matrix. The concentration of statistically significant values can be determined by calculating the z-score of the attribute data for each spatial unit at the entire site. If the z-score is positive, it can be concluded that high attribute data are spatially clustered, and vice versa if it is negative. This is expressed as shown in formula (2).

$$G_i^* = \frac{\sum_{j=1}^{n} w_{i,j} x_j - \bar{X} \sum_{j=1}^{n} w_{i,j}}{S \sqrt{\frac{\left[n \sum_{j=1}^{n} w_{i,j}^2 - \left(\sum_{j=1}^{n} w_{i,j}\right)^2\right]}{n-1}}} \quad (2)$$

xj = the attribute value for feature j

$w_{i,j}$=the spatial weight between feature i and j

n = the total number of features

The CVI indicators for each grid cell were calculated and converted into z-scores. The total sum that reflects the weight of indicators was obtained. The CVI and hotspot analysis results were derived by converting this sum back into a z-score. Basic statistics for the indicators and weight used to calculate CVI are shown in Table 1.

Variables	N	Mean	SD	Communality	Weight
The Number of Stores/cell	5156	52	77.45366	0.746	0.2998
Land value/cell	5156	576133213	625332001	0.527	0.2118
The Number of Conversions/cell	5156	1	2.019241	0.493	0.1982
Sales/cell	5156	8421831934	21609023906	0.391	0.1572
De facto population/cell	5156	128470	254468.325	0.331	0.1330
Sum				2.488	1.0000

Table 1. Basic statistics for indicators and weight used to calculate CVI

Based on the results of the CVI hotspot analysis, seven intervals were identified where spatial boundaries were clearly revealed. The growth stages of commercial areas were classified for the top four sections above the average, and the three sections below the average were classified as cold spots and excluded from the analysis. The four types are as follows: 1) Developed area, which is highly revitalized and serves as the center of Seoul; 2) Growing area that serves as subcenters; 3) Neighborhood area, commercial area in densely populated residential areas that provide goods and services to residents; and 4) Stagnant area with low residential and commercial density. It was defined as a cold spot that is challenging to analyze due to insufficient data. There were three cores in the developed area and five in the growing area (figure 1). Table 2 shows the number of cells corresponding to each type and representative district.

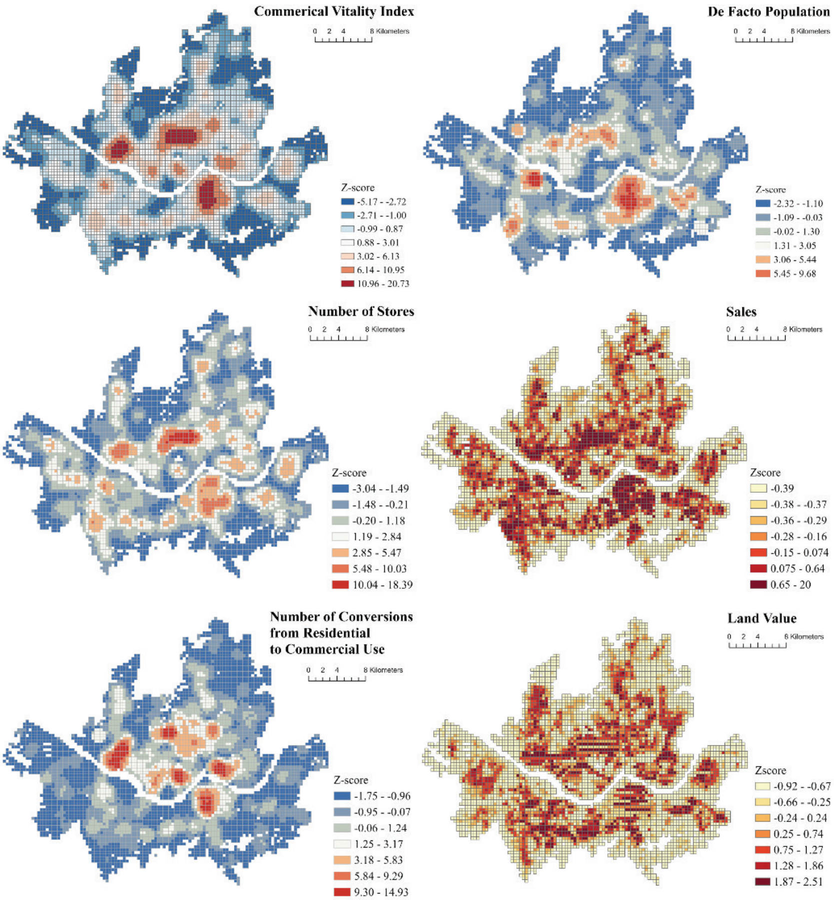


Figure 1. Spatial distribution of commercial growth stage and CVI indicators.

CVI	Growth Stage	N	Description (Examples)
≤20.734	Developed Area	127	Mature commercial area including central business districts.
≤10.946	Growing Area	242	Activated area with mixed-use of residential and commercial
≤6.134	Neighborhood Area	517	Commercial areas included in large residential concentrations
≤3.011	Stagnant Area	808	Low-density areas between neighborhood area
≤0.865	(Cold spot)	2409	Lack of commercial business data
Total		5156	-

Table 2. Classification of commercial district Growth Stage

2.3. ANALYZING STREET LEVEL WALKABILITY FACTORS

This study focused on analyzing the walkability factors of the growing stage commercial district, which are distinct from the downtown's. According to the analysis on commercial area vitality, Seongsu district, one of the growing commercial areas in Seoul is selected for the case study area.

To measure Street Level Walkability (SLW), the study area was divided into smaller clusters than the neighborhood level. Using POI data from the selected study area, density-based clustering analysis was conducted using the DBSCAN method. It identifies clusters and noise by detecting features existing within a given distance from the core. A cluster of point data has a higher density than the surrounding area, whereas the noise area has a lower density than that of other clusters. This study calculated valid clusters from 10 features within a radius of 50 meters from the core. Also, 3Ds were quantified to analyze the walkability features of each street cluster. This study analyzed the physical environment at street level where small businesses are concentrated, and assumed that factors in micro-levels are not considered. The proxies representing each feature and their descriptions are as follows.

Density

To determine the concentration of urban activities within the study area, the average floor area ratio (FAR) of buildings was considered. Providing various destinations within a short travel distance, the density of buildings induces various activities among city residents (Jacobs, 1961).

Diversity

The Shannon Entropy Index was employed to measure the diversity of commercial businesses on the streets. As a tool for evaluating land-use mix (LUM) in the neighborhood walking environment, it was widely used (Frank and Pivo, 1994; Ewing and Cervero, 2010; Song et al., 2013). Using the entropy index, this study measured diversity by classifying land-use types into residential, commercial, and industrial. The formula for calculating the entropy index of the diversity of businesses is expressed as shown below in formula (3):

$$LUM = -\sum_{i=1}^{n} \frac{p_i \ln(p_i)}{\ln(n)} \tag{3}$$

Design

Hillier and Hanson (1984) proposed space syntax as a quantitative method for analyzing the spatial configuration concerned with the network between the spaces. Integration (INT), which is one of the space syntax attributes, is used to determine the degree of integration between a specific axial line is and other axial lines in the entire axial map. Koohsari et al. (2016) adopted the space syntax to develop an alternative walkability index. They found a correlation between conventional full walkability and space syntax walkability. DepthMapXnet0.35 software was used for axial map analysis. Because this study focuses on pedestrians on two-lane or more roads where sidewalks and roads are separated, axial lines are drawn on each sidewalk on both sides of the road rather than the road's centerline—then connected with crosswalks.

3. Results

3.1. CLASSIFYING COMMERCIAL STREETS CLUSTERS

The case study was conducted in the Seongsu district, located on the Han River's northeast side. The area is about 0.82km2; it consists of an industrial area with small factories, a general residential area with a mixture of single and multi-unit houses, and a commercial area developed around three subway stations. Seongsu district is a urban industrial area where shoes, leather, and printing industires are concentrated. Since the 2000s, it has suffered stagnation because the light industry which was the mainstream of Seongsu industrial district has been dispersed to the outskirts of Seoul or other countries. Since then, more than 50 social enterprises and non-profit organizations entered the area, as well as shops and restaurants renovated from factories. As restaurants and retail shops developed in places where factories had left, it changed from a residential, industrial area to a commercial area (Kim and Lee, 2016). Because of the unique characteristics created by old factories and modern commercial facilities, it has recently attracted global brands to invest in the area.

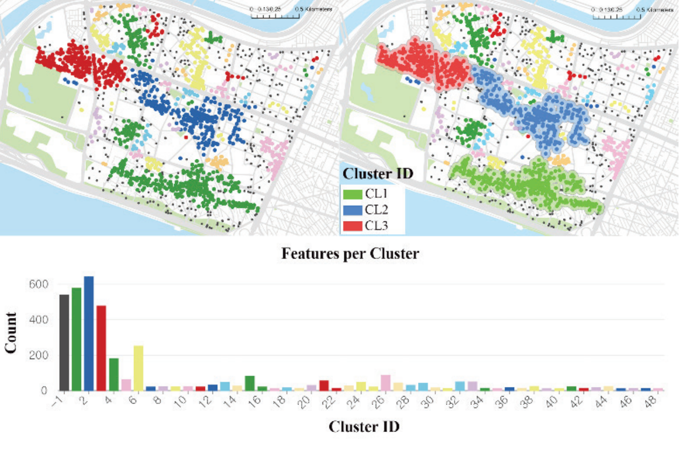


Figure 2. The detected clusters and the study unit with a 40 m buffer zone.

The top 3 clusters containing the most features were selected through clustering analysis in ArcGIS, and a 40 m radius buffer was applied to limit the study unit. Figure 2 shows the cluster analysis results and features per cluster count chart.

3.2. ANALYSIS ON SPATIAL HETEROGENEITY OF SLW

Cluster CL1, which contains 577 features, was centered on a residential area that was located south of the target area. The stores were widely distributed along the narrow two-lane road that cuts through the low-rise residential area. The LUM was 0.1023, and the average FAR was 46.41%, which was the lowest among the three clusters. Presumably, this is because most the buildings in this cluster are small-scale residential buildings. Conversely, the average INT was 0.7784, which was the second highest value among clusters, indicating relatively high accessibility. These results are listed in Table 3 and Figure 3.

Cluster ID	Number of POI	Diversity		Density		Design	
		LUM	N	Avg. FAR	N	Avg. INT	N
Total	3818	0.1119	4929	63.85	5236	0.7991	706
CL1	577	0.1023	1017	46.41	1030	0.7784	99
CL2	641	0.1697	513	73.11	529	0.9836	72
CL3	477	0.1380	512	92.92	519	0.6921	66

Table 3. Description of SLW value per cluster

CL2, which contains the most features (641), formed a cluster along the south side of the 6-lane road adjacent to the Seongsu subway station and the side streets. Among all the three clusters, this cluster was analyzed to have the highest LUM, which represents diversity and INT, which indicates a structure with high accessibility. Also, the average FAR was higher than that of the total area. This area has the highest proportion of industrial-use buildings at 21.21%, which means that this area is the center of the largest number of workplaces as well as stores. The presence of this main street supports Jacobs' (1961) argument that the vitality of the streets increases as pedestrians with various purposes are concentrated within a certain area.

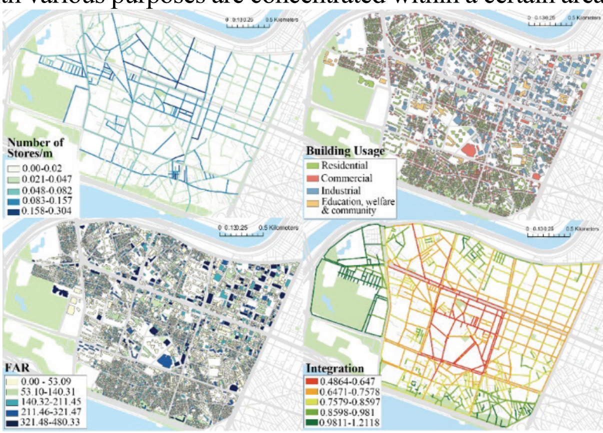


Figure 3. Results of the analysis on SLW factors: (a) the number of stores per meter; (b), (c) the distribution of building usage and FAR; and (d) the axial map indicating integration value.

Finally, CL3, which contained 477 features, had the characteristics of being distributed around Seoul Forest Park, located at the west of the target area and alleys adjacent to it. It shows the highest average FAR, while LUM occupied a second place among all clusters. However, average integration, which means accessibility, was lower than that of the other two clusters, indicating heterogeneous characteristics in the spatial hierarchy. In other words, it is assumed that distinct factors, such as the existence of parks, acted as reasons for the formation of clusters in locations with low accessibility.

4. Conclusion

Considering the commercial area vitality, this study suggests a model that measures the walkability of the commercial streets at various stages and analyzed the walkability characteristics of newly growing commercial areas. First, using the CVI that applied to previous studies, commercial areas in Seoul were divided into four categories of areas: developed, growing, neighborhood, and stagnant. Second, as a commercial area in the growth stage, a case study was conducted on the walkability factors of the Seongsu district. Using POI data, street-level clusters were derived, and the walkability characteristics of each cluster were compared in three aspects: diversity, density, and design.

According to the results of a case study on the Seongsu District, the streets where stores are concentrated showed distinct characteristics according to 3Ds even within one neighborhood. Specifically, it was analyzed that it consists of the main street with the highest diversity, density, and accessibility, a neighborhood commercial district formed around low-rise residential areas and supports daily needs, and a specialized alley commercial district that utilizes green spaces. The three types of commercial streets are balanced and constitute the walkability characteristics of the entire district.

This study identified the walkable characteristics of several streets where stores are concentrated within a region in detail from the perspective of urban renewal. The physical characteristics of streets with high small-business density are distinguishing, and the primary uses of streets are in complementary relationships. As an old mixed-use area of the city, to increase the vitality of the city, it is necessary to play the role of a main street with high density, diversity, and accessibility and a living street to supplement it. This planning approach becomes the basis for realizing a human-centered urban design by expanding the range of choices for street users. Utilizing the framework applied in this study, walkability features can be evaluated for areas with different commercial growth stages. Also, it showed that street clusters with different densities, diversity, and spatial structures coexist in one commercial area, and further research is needed on the interaction brought about by this combination.

However, this study has limitations. First, it has not been verified whether the case of the Seongsu district can be generalized to other regions. More cases are needed for this to settle into a standardized model. Second, it is necessary to analyze the characteristics of individual commercial streets from a user's point of view by including environmental variables that can be recognized in walking situations, such as storefronts and street furniture. Based on this study's evaluating process, future studies will investigate districts in different growth stages and conduct a comparative analysis by considering the meso-scale to the micro-scale walkability factors together.

Acknowledgments

This research was supported by the BK21 FOUR (Fostering Outstanding Universities for Research), which is funded by the Ministry of Education (MOE) of Korea and the

National Research Foundation (NRF) of Korea.

References

Cervero, R., & Kockelman, K. (1997). Travel demand and the 3Ds: Density, diversity, and design. Transportation Research Part D: Transport and Environment, 2(3), 199-219, https://doi.org/10.1016/S1361-9209(97)00009-6.

DepthmapX development team. (2017). depthmapX (Version 0.6.0) [Computer software]. Retrieved from https://github.com/SpaceGroupUCL/depthmapX/

Ewing, R., & Cervero, R. (2010). Travel and the built environment: A meta-analysis. Journal of the American planning association, 76(3), 265-294, DOI:10.1080/01944361003766766.

Forsyth, A. (2015). What is a walkable place? The walkability debate in urban design. Urban Design International, 20(4), 274–292, https://doi.org/10.1057/udi.2015.22.

Frank, L. D., & Pivo, G. (1994). Impacts of mixed use and density on utilization of three modes of travel: single-occupant vehicle, transit, and walking. Transportation research record, 1466, 44-52.

Hillier, B., Hanson, J. (1984). The Social Logic of Space. Cambridge: Cambridge University Press.

Jacobs, J. (1961). The Death and Life of Great American Cities. New York: Random House.

Kim, S.,& Lee, H. (2016). A case study on the gentrification process and characteristics of seongsu-dong. The Journal of Culture Contents, (7), 81-105.

Koo., et al. (2022). How are Neighborhood and Street-Level Walkability Factors Associated with Walking Behaviors? A Big data Approach Using Street View Images. Environment and Behavior, 54(1), 211–241, https://doi.org/10.1177/00139165211014609.

Koohsari, M. J., Owen, N., Cerin, E., Giles-Corti, B., & Sugiyama, T. (2016). Walkability and walking for transport: characterizing the built environment using space syntax. International Journal of Behavioral Nutrition and Physical Activity, 13(1), 1-9.

Lee, D., Lee., &Cheon, S. (2020). Comparison of Survival Rate in Food Service Industry between Growing Commercial Districts and Declining Commercial Districts of Seoul. The Korea Spatial Planning Review, 105, 65–84, DOI : 10.15793/kspr.2020.105..005.

Leinberger, C., & Alfonzo, M. (2012). Walk this way: The economic promise of walkable places in metropolitan Washington, DC.

Mehta, V. (2007). Lively Streets: Determining Environmental Characteristics to Support Social Behavior. Journal of Planning Education and Research, 27, 165
–187, https://doi.org/10.1177/0739456X07307947.

O'Sullivan, A. (1996). Urban economics. Chicago: Irwin.

Park, B., &Kim, J. (2010). A study on the dynamic decline types of local cities using multiple decline index. Journal of the Korean Regional Science Association, 26(2), 3-17, https://doi.org/10.5389/KSAE.2019.61.5.057.

Park, S., Choi, K., &Lee, J. (2017). Operationalization of Path Walkability for Sustainable Transportation. International Journal of Sustainable Transportation, 11(7), 471–485, https://doi.org/10.1080/15568318.2016.1226996.

Song, Y., Merlin, L., & Rodriguez, D. (2013). Comparing measures of urban land use mix. Computers, Environment and Urban Systems, 42, 1-13, DOI:10.1016/j.compenvurbsys.2013.08.001.

Tong, X., Wang, Y., & Chan, E. H. (2016). International research trends and methods for walkability and their enlightenment in China. Procedia Environmental Sciences, 36, 130-137, https://doi.org/10.1016/j.proenv.2016.09.023.

Wang, H.,& Yang, Y. (2019). Neighborhood walkability: A review and bibliometric analysis. *Cities*, 93, 43-61, https://doi.org/10.1016/j.cities.2019.04.015.

DEVELOPMENT OF A METHOD FOR ASSESSING THE VIEW INDEX OF PLANTS OF INTEREST USING DEEP LEARNING

ANQI HU[1], NOBUYOSHI YABUKI[2] and TOMOHIRO FUKUDA[3]
[1,2,3] Division of Sustainable Energy and Environmental Engineering, Osaka University Graduate School of Engineering.
1anqihu1028@gmail.com, 0000-0002-8568-1030
[2]yabuki@see.eng.osaka-u.ac.jp, 0000-0002-2944-4540
[3]fukuda.tomohiro.see.eng@osaka-u.ac.jp, 0000-0002-4271-4445

Abstract. Urban planning often overlooks the diversity of plant species and the perspectives of pedestrians. This study introduces the View Index of Plants of Interest (VIPI) as a new index for evaluating street plants from a pedestrian perspective. VIPI uses image classification and semantic segmentation techniques and was applied to four popular ornamental street plants: cherry trees, maple trees, magnolia trees, and ginkgo trees. The model used achieved a high level of accuracy with a mean intersection over union (mIoU) of 81.06. And the VIPI satisfaction criteria were used to evaluate several cases. The results provide valuable insights for urban planners and policymakers, allowing for a more detailed and accurate evaluation of urban plants from a pedestrian perspective and can guide urban greening actions. Additionally, this study demonstrates the potential of utilizing computer science techniques to inform urban planning and design decisions.

Keywords. Deep learning, Image Classification, Semantic Segmentation, View Index of Plants of Interest (VIPI), Street view Images, Urban Green Space.

1. Introduction

Urban greenery plays a crucial role in landscape design, as it offers numerous benefits to the urban environment beyond aesthetics. These benefits include air and water purification, noise reduction, and climate regulation (Wolf, 2005; Virtudes, 2016). Advances in computing technology have facilitated a more comprehensive understanding and evaluation of the urban environment from a human perspective, which has become a central focus in urban planning research.

One widely-used method for evaluating urban greenery is the Green View Index (GVI), which calculates the proportion of greenery along streets and estimates regional greenery planting (Ki & Lee, 2021). However, it would be beneficial to develop more advanced models that consider different plant species, as each species has a unique impact on humans and the environment (Elsadek & Fujii, 2014).

In light of this, this study proposes a new evaluation index from a human perspective for street plants, the View Index of Plants of Interest (VIPI). The VIPI index allows for a more detailed study of the Green View Index (GVI) (Aoki et al., 1985) by refining the study's target population from all plants to specific plants. By extending the ponderous plants in general to plant categories, the VIPI index enables a more comprehensive analysis. The VIPI index was calculated using techniques of image classification and semantic segmentation. Four popular ornamental street plants, cherry trees, maple trees, magnolias, and ginkgo trees, were selected for this study. Since there is no publicly available related dataset for these plants, a dataset of these street plants was created.

In addition, an anonymous experimental survey was conducted within the school and nearby neighborhoods to better assess the VIPI of each area in view. The results were used as a reference for developing VIPI assessment criteria. This criterion allows for better evaluation of VIPI across time and space and thus makes recommendations accordingly. This study also used the criterion to assess VIPI in several areas of Japan. The VIPI results are also used to guide urban greening actions by providing a more detailed assessment of each urban plant from a pedestrian perspective, thus providing valuable insights for urban planners. New computer science techniques are used to gain insight into the urban environment's physical world to assess urban plants' performance, thus providing better information for urban planning and design.

2. Literature review

2.1. DEVELOPMENT OF GREENING INDICATORS

Urban greening is a vital aspect of sustainable urban development and has been studied using various indicators. These indicators can be divided into two types: horizontal and vertical. Horizontal indicators, such as the NDVI, measure urban greenery using satellite imagery (Aryal et al., 2022) but do not account for scattered street greenery. Vertical metrics, such as the GVI, measure urban greenery from a human perspective. Subsequently, many new studies have been developed based on the GVI, such as the Panoramic View Green View Index (Xia et al., 2021) and the Enhanced Green View Index (Puppala et al., 2022).

While these studies provide valuable insights, they lump all plants into one category and fail to account for the variation in attractiveness and impact of different plant species (Pratiwi et al., 2019). This study proposes a new metric, the VIPI, that focuses on a single plant as the object of study and provides a more precise and accurate way to measure the value of urban plants, providing valuable insights for urban planners and guiding urban greening initiatives.

2.2. DATA SET OF PLANTS

This study aims to create a dataset of street scenes with plants observed from a human perspective in an urban environment, addressing the need for such datasets. The dataset is based on the Cityscapes dataset. Photographs were selected to capture the most visually appealing period of the selected plants, such as when most maple leaves are red. Plant selection criteria include widespread distribution and popularity in urban landscapes, aesthetic appeal to the public, and potential for use as a design technique or to add value. Four plant species were selected: cherry, maple, ginkgo, and magnolia. These species are widely distributed in Asian cities, have annual festivals that bring economic benefits, and have positively impacted the environment and human well-being (Asgarzadeh et al., 2014; Guan et al., 2017; Pratiwi et al., 2019). The study also included steps for selecting plants for the landscape, following the book "Landscape Architecture" (Laurie, 1975, pp149-148). It states that the criteria for plant selection include wide distribution and popularity in the urban landscape, aesthetic appeal to the public, and potential for use as a design technique or to add value.

3. Methodology

3.1. RESEARCH FRAMEWORK

Figure 1 displays the research framework for this study. The preliminary phase includes three steps: creating an AI dataset, selecting a model, and developing VIPI formulas and satisfaction criteria. The primary process, model deployment, is the fourth step. The dataset creation step begins with collecting photographic data, including photos taken by the authors and open-source data, focusing primarily on cherry trees, maple trees, ginkgo trees, and magnolias. The raw data is labeled to create an artificial intelligence (AI) dataset. The second step involves the model design, including image classification, semantic segmentation, model selection, and testing. The VIPI calculation formula is developed in the third step, and the VIPI satisfaction criteria are set. Finally, in the fourth step, the model is used to input images from test locations, using the image classification model learned from the raw dataset and the semantic segmentation model learned from the AI dataset to evaluate local VIPI values.

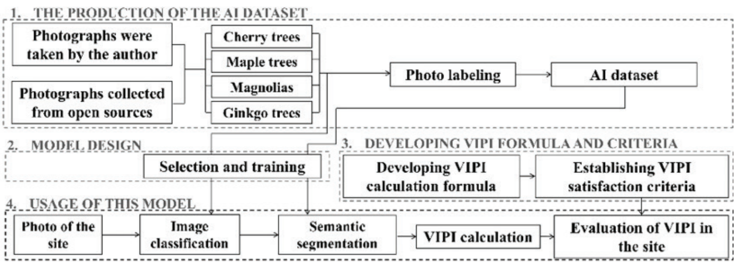


Figure 1. Research Framework.

3.2. THE PRODUCTION OF THE AI DATASET

3.2.1. Dataset photo collection

The dataset for this study was created using the Cityscapes dataset as a reference. It consists primarily of street scenes with the subject plants, with some close-ups to improve model accuracy later. The photographs were chosen to show the plants at their most visually appealing stage, such as when most maple leaves are red. The complexity of the plants, including single and multiple overlapping plants, was also considered. Photographs were collected in a variety of urban landscapes. Table 1 lists the number of photographs and sources of photographic data for each plant.

	Cherry	Maple	Magnolias	Ginkgo	Total
Photographs were taken by the author (sheets)	170	160	40	55	425
Photographs collected from open sources (sheets)	50	60	180	165	455
Total (sheets)	220	220	220	220	880

Table 1. Urban landscape plant dataset.

3.2.2. Photo labeling

The next step in creating the dataset was manually labeling the plants of interest in the raw photographs using the segment.ai segmentation platform. This step was done by manually segmenting and annotating as much detail as possible, as shown in Figure 2.

Figure 2. Process image of the labeled data.

3.3. MODEL DESIGN

The model design for this study focuses on the selection of appropriate models. Specifically, an image classification model and a semantic segmentation model were used. While a semantic segmentation model using scene parsing alone can label plants, incorporating an image classification model can significantly improve the model's accuracy. In addition, converting the multi-classification task to a binary classification task reduces the GPU memory required for computation.

3.3.1. Image classification models

Image classification uses image processing to differentiate between classes of targets based on their features in the image information. Deep neural networks, like convolutional neural networks and visual transformers, have proven effective when large models can be trained with rich labeled data (Kolesnikov et al., 2020). In this study, the use of an image classification model as a first step in the analysis system significantly optimizes the model structure. It improves the results of the semantic segmentation at a later stage. The tree species were classified using the EfficientNet network structure (Tan & Le, 2019) as a baseline network. EfficientNet achieves a state-of-the-art maximum accuracy of 84.3% on ImageNet. An accuracy of 97.9% was obtained using EfficientNet-b4 to run the dataset for this study.

3.3.2. Semantic segmentation models

This study employs semantic segmentation for scene parsing, a widely studied technique for assigning labels to image pixels in computer vision. Recent advances in deep learning have led to increased use of this method for automated landscape evaluation (Tang & Long, 2019). In this study, a semantic segmentation approach is utilized for scene parsing. Several models were tested on the dataset, and the highest accuracy was achieved by PSPNet (Zhao et al., 2017), DANet (Xue et al., 2019), and ISANet (Huang et al., 2019), with mIoU values of 81.66, 81.09, and 80.03, respectively. However, as the study's goal was to calculate the VIPI, three methods for calculating the number of vegetation pixels of interest were proposed: utilizing only the best semantic segmentation model, utilizing the intersection of two or more models, and utilizing the average of all models. 20% of the 880 collected photographs were used as a test set to evaluate the accuracy of these methods."

This study uses the mean square error (MSE) as the model comparison method, and the results can be calculated by Equation (1).

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2 \qquad (1)$$

Where n: the amount of data. i: natural number from 1. $\hat{y}_i$: predicted values. y_i: correct values. Table 2 shows that DANet has good results for both MSE and mIoU.

Model	ISA Net	PSP Net	DA Net	ISANet ∩ PSPNet	ISANe t∩ DANet	PSPNe t∩ DANet	Mean
MSE	235.6 6	337.09	148.74	339.21	244.43	340.24	218.82

Table 2. MSE results for the three methods

Therefore, DANet was used as a model for semantic segmentation. The DANet model is trained on the computer by relying on pre-trained weights. Based on the analytical choices above, EfficientNet-b4 was chosen for the image classification part of the model, and DANet was chosen for the image segmentation part. The model structure diagram is shown in Figure 3.

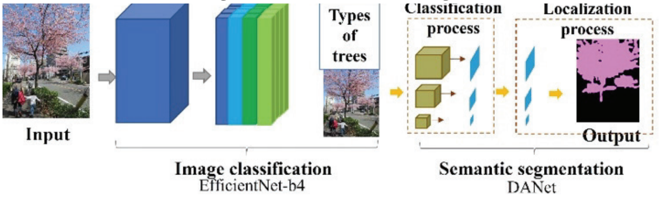


Figure 3. Structure of the model.

3.4. DEVELOPING VIPI FORMULA AND SATISFACTION CRITERIA

3.4.1. Developing VIPI calculation formula

The image segmentation results presented in Section 3.3 demonstrate using the VIPI formula, defined in Equation (2).

$$VIPI = \frac{\sum_{i=1}^{n} Area_{v\_i}}{\sum_{i=1}^{n} Area_{t\_i}} \times 100\ (\%) \qquad (2)$$

Where $Area_{v\_i}$ represents the total number of pixels of the target plant at the intersection in the image taken along the horizontal direction, with i ranging from 1 to n, where n is the total number of photos taken in the test area. $Area_{t\_i}$ represents the total number of pixels in that image. The VIPI of a photograph is calculated as the ratio of the interested interest category's pixel value to the input photograph's total pixel value. The average value of the VIPI obtained from all images of the test area is used to determine the visibility value of the plant of interest at that location.

3.4.2. Establishing VIPI satisfaction criteria

This study aimed to objectively assess the visual impact of planting in various regions and timeframes through a comparative analysis of the VIPI. A sample of 10 participants (5 with urban design education and five from the targeted community) were surveyed to investigate their visual preferences for plants in urban environments. The results showed that the participants valued diversity in plant species, proximity to plants, and the dynamic beauty of landscape plants that change with the seasons. These criteria of accessibility, comfort, pleasurability, and diversity were used to evaluate the visual effects of ornamentals.

In the subsequent questionnaire, participants were asked to assign indicator weights to physical satisfaction when experiencing plants. They performed pairwise comparisons on the criteria of accessibility, comfort, pleasurability, and diversity and provided a scoring matrix. The scores were then mathematically averaged, and an Analytic Hierarchy Process (AHP) was conducted. Table 3 showcases the results of the AHP analysis, including the consistency ratio (CR) and other relevant parameters. The model passed consistency judgment (CR = 0.0002 < 0.1), and the resulting weights were deemed usable.

Satisfaction indicators	Accessibility	Comfort	Pleasurabil-ity	Diversity
Weight	16.73%	32.52%	32.53%	18.22%
λmax = 4.007	CI = 0.0002	RI = 0.89	CR = 0.0002.	

Table 3. Indicator weightings and consistent judgments result from using AHP for physical satisfaction evaluation.

An additional sample of 10 participants was recruited to score 40 street photos on accessibility, comfort, pleasurability, and diversity using a Likert scale. The above weights were used to calculate the overall satisfaction score for the test photos. A scatter plot was created to investigate the correlation between the satisfaction scores and VIPI, with an R-squared value of 0.087. The scatter plot and the formula for the linear fit of the data were presented in Equation (3).

$$VIPI = 0.249 + 0.052 \times Satisfaction\ score \qquad (3)$$

Through clustering and analysis of the data, it was determined that the data were best grouped into three categories with a Sum of Squared Errors (SSE) of 31.948. Figure 4 presents both the scatter plot and the clustering results. Based on this analysis, the satisfaction criteria for VIPI were developed: 42%-100% for very satisfied (V), 37%-42% for satisfied (S), and 0-37% for dissatisfied (D).

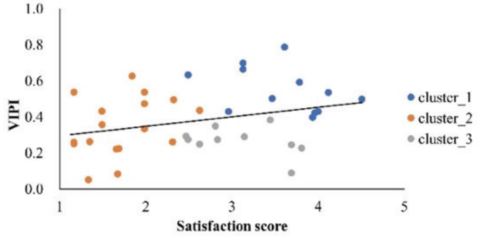


Figure 4. Scatterplot of VIPI and satisfaction.

3.5. THE USAGE OF THIS MODEL

This study used street view images from Google Earth and photographs as data sources. All photographs were precisely adjusted in distance and angle to follow the "Osaka Prefectural Green Vision Survey Guide guidelines." To take pictures of cherry blossoms, picked 24 famous places in Japan during the best time to see them using Google Earth. Cherry blossom snapshots taken on the streets of Suita were also included in the collection.

4. Results

The accuracy of the model system tested in the plant database of this study was determined to be 81.06% using the mean intersection over union (mIoU) metric. The IoU values for each plant are shown in Table 4.

Table 4. Per-plant results on this study's dataset

	Cherry	Maple	Magnolias	Ginkgo	mIoU (%)
IoU (%)	82.84	83.66	79.06	78.68	81.06

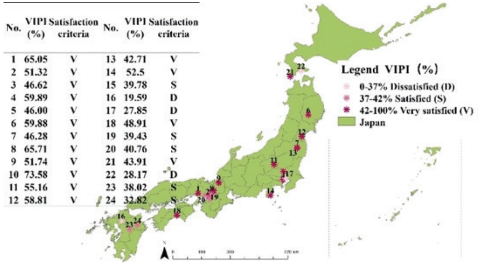


Figure 5. VIPI results for 24 famous cherry blossom viewing areas in Japan.

The results of using the model are presented in the following section. The Google Street View data of the 24 famous cherry blossom viewing sites in Japan were entered into the model system. The VIPI results for each area were obtained and combined with the geographic coordinates of the area. These results were then imported into ArcGIS Map geographic data processing software to produce Figure 5.

Statistical analysis shows that 21 sites have reached a very satisfactory standard of VIPI. Three sites are in areas that do not meet the standard, and it is necessary to improve the visibility of cherry blossoms in these areas. Five areas still meet the satisfactory standard but still strengthen the improvement.

In order to evaluate the operability and accuracy of the system with non-street view images, six photos of cherry blossoms on the street were taken. Figure 6 displays the VIPI results, the geographical location of cherry blossoms, and the changes in visibility. In addition, the VIPI satisfaction of each location was evaluated. The advantage of analyzing photos taken by individuals is that the shooting points are more flexible, making it possible to track changes in VIPI for successive street plants; as can be seen, Site 1 and Site 4 have a significant change in VIPI.

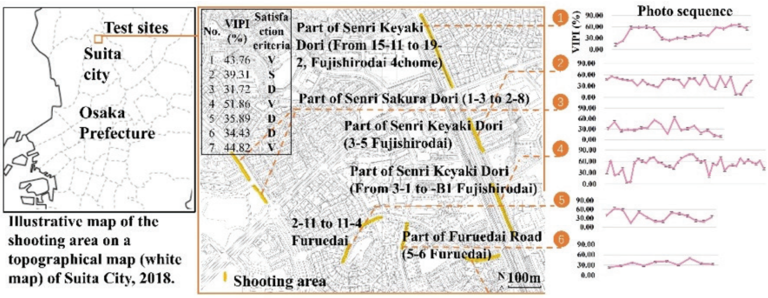


Figure 6. VIPI statistics for streets in Suita, Osaka Prefecture, Japan.

5. Discussion

Based on the results of this study, the VIPI index is a valuable tool for assessing interest in viewing different types of plants in an urban environment. The proposed system for calculating VIPI using image classification and semantic segmentation models is efficient and accurate. VIPI is used in the context of urban planning and design. For example, this study uses their VIPI data to inform the placement of new green spaces after learning about 24 famous cherry blossom viewing spots. To make a viewing path attractive for pedestrians, it should have a VIPI of at least 37%. If an existing green space attracts few visitors, try changing the viewing path and increasing the distance to the plants to improve VIPI satisfaction. It is also possible to increase the plant species richness and the design ingenuity of the overall landscape environment.

In addition, VIPI information can be used to analyze how visitors move around a site and how this relates to the VIPI scores of different areas. For example, pedestrian activity at seven observation points in the Suita area was recorded and analyzed with VIPI. The results are expected to shed new light on the design of the streetscape. Of course, the recording of pedestrian activity must be carried out under legal conditions.

However, several limitations need to be addressed in future research, such as optimizing the plant dataset and increasing the variety of photo data. One potential area for further exploration is using VIPI data to improve the attractiveness of urban green spaces. By identifying areas with substandard VIPI, urban planners and designers can take specific actions to improve the visibility and accessibility of certain plant species. For example, they can use VIPI data to identify the most popular plant species for pedestrians and prioritize planting those plants in areas with low VIPI scores. Another potential area of exploration is the use of VIPI data in the context of urban green space planning and design. VIPI data can be used to inform decisions about the layout and design of green spaces, such as the placement of different plant species, the use of different types of vegetation in different areas, and the use of different types of surface materials. In addition, by combining VIPI data with other data sources such as environmental, social, and economic data, urban planners and designers can gain a more comprehensive understanding of the impact of urban greening on different aspects of urban life.

Finally, it is worth noting that although the study focuses on cherry blossom viewing sites in Japan, the VIPI approach can be applied to any other sites and plant species if the corresponding plant data sets are provided.

6. Conclusion

This study makes several significant contributions to urban landscape plant analysis. First, it refines the plant categorization scheme. It proposes the View Index of Plants of Interest (VIPI) as a more specific indicator of greenery than previously used statistical indicators such as GVI, PVGVI, and EGVI, which treat all types of greenery as a single category. Second, it expands the dataset of urban landscape plants to include a variety of species, such as cherry trees, maple trees, magnolias, and ginkgo trees and is collected from a human-centered perspective. Third, it demonstrates the high accuracy of the proposed framework, which utilizes deep learning models such as EfficientNet-b4 and DANet, in detecting vegetation in urban scenes. Finally, it shows that VIPI criteria can be used to evaluate greenery at different times and spaces. This research provides valuable insights for urban planners, policymakers, and sociologists to create healthier and more livable urban environments using VIPI as a base data.

Acknowledgments

This work was supported by JST SPRING, Grant Number JPMJSP2138.

References

Aoki, Y., Yasuoka, Y., & Naito, M. (1985). Assessing the impression of street‐side greenery. Landscape Research, 10(1), 9–13. https://doi.org/10.1080/01426398508706131

Aryal, J., Sitaula, C., & Aryal, S. (2022). NDVI Threshold-Based Urban Green Space Mapping from Sentinel-2A at the Local Governmental Area (LGA) Level of Victoria, Australia. Land, 11(3), 351. https://doi.org/10.3390/land11030351

Asgarzadeh, M., Koga, T., Hirate, K., Farvid, M., & Lusk, A. (2014). Investigating oppressiveness and spaciousness about building, trees, sky and ground surface: A study in Tokyo. Landscape and Urban Planning, 131, 36–41. https://doi.org/10.1016/j.landurbplan.2014.07.011

Elsadek, M., & Fujii, E. (2014). People's Psycho-physiological Responses to Plantscape Colors Stimuli: A Pilot Study. International Journal of Psychology and Behavioral Sciences, 10, DOI:10.5923/j.ijpbs.20140402.02.

Guan, H., Wei, H., He, X., Ren, Z., & An, B. (2017). The tree-species-specific effect of forest bathing on perceived anxiety alleviation of young adults in urban forests. Annals of Forest Research, 0(0). https://doi.org/10.15287/afr.2017.897

Huang, L., Yuan, Y., Guo, J., Zhang, C., Chen, X., & Wang, J. (2019). Interlaced Sparse Self-Attention for Semantic Segmentation (arXiv:1907.12273). https://doi.org/10.48550/arXiv.1907.12273.

Ki, D., & Lee, S. (2021). Analyzing the effects of Green View Index of neighborhood streets on walking time using Google Street View and deep learning. Landscape and Urban Planning, 205, 103920. https://doi.org/10.1016/j.landurbplan.2020.103920

Kolesnikov, A., Beyer, L., Zhai, X., Puigcerver, J., Yung, J., Gelly, S., & Houlsby, N. (2020). Big Transfer (BiT): General Visual Representation Learning. In A. Vedaldi, H. Bischof, T. Brox, & J.-M. Frahm (Eds.), Computer Vision – ECCV 2020 (pp. 491–507). Springer International Publishing. https://doi.org/10.1007/978-3-030-58558-7_29

Laurie, M. (1975). *Introduction to landscape architecture.* American Elsevier Pub. Co.

Pratiwi, P. I., Xiang, Q., & Furuya, K. (2019). Physiological and Psychological Effects of Viewing Urban Parks in Different Seasons in Adults. International Journal of Environmental Research and Public Health, 16(21), 4279. https://doi.org/10.3390/ijerph16214279

Puppala, H., Tamvada, J. P., Kim, B., & Peddinti, P. R. T. (2022). Enhanced green view index. MethodsX, p. 9, 101824. https://doi.org/10.1016/j.mex.2022.101824

Tan, M., & Le, Q. (2019, May). Efficientnet: Rethinking model scaling for convolutional neural networks. In International conference on machine learning (pp. 6105–6114). PMLR, https://doi.org/10.48550/arXiv.1905.11946.

Tang, J., & Long, Y. (2019). Measuring visual quality of street space and its temporal variation: Methodology and its application in the Hutong area in Beijing. Landscape and Urban Planning, 191, 103436. https://doi.org/10.1016/j.landurbplan.2018.09.015

Virtudes, A. (2016). Benefits of Greenery in Contemporary City. IOP Conference Series: Earth and Environmental Science, p. 44, 032020. https://doi.org/10.1088/1755-1315/44/3/032020

Wolf, K. L. (2005). Business district streetscapes, trees, and consumer response. Journal of Forestry, 103(8), 396–400.

Xia, Y., Yabuki, N., & Fukuda, T. (2021). Development of a system for assessing the quality of urban street-level greenery using street view images and deep learning. Urban Forestry & Urban Greening, 59, 126995. https://doi.org/10.1016/j.ufug.2021.126995

Xue, H., Liu, C., Wan, F., Jiao, J., Ji, X., & Ye, Q. (2019). DANet: Divergent Activation for Weakly Supervised Object Localization. 2019 IEEE/CVF International Conference on Computer Vision (ICCV), 6588–6597. https://doi.org/10.1109/ICCV.2019.00669

Zhao, H., Shi, J., Qi, X., Wang, X., & Jia, J. (2017). Pyramid Scene Parsing Network (arXiv:1612.01105). arXiv, https://doi.org/10.48550/arXiv.1612.01105.

GENERATIVE PRE-TRAINED TRANSFORMERS FOR 15-MINUTE CITY DESIGN

RUTVIK DESHPANDE[1], SAYJEL VIJAY PATEL[2], CAMIEL WEIJENBERG[3], MACIEJ NISZTUK[4], MIRIAM CORCUERA[5], JIANXI LUO[6], QIHAO ZHU[7]
[1,2,3,4,5]*Digital Blue Foam,*
[6,7]*Singapore University of Technology and Design,*
[1]*rutvik@digitalbluefoam.com, 0000-0002-3837-4492*
[23]*{sayjel|camiel}@digitalbluefoam.com*
[4]*maciej@digitalbluefoam.com, 0000-0001-6520-5128*
[5]*miriam@digitalbluefoam.com, 0000-0002-1100-8775*
[6]*luo@sutd.edu.sg, 0000-0001-5892-8432*
[7]*qihao_zhu@mymail.sutd.edu.sg, 0000-0002-5401-6679*

Abstract. Cities globally are adopting "The 15-Minute City" as an urban response to various crises, including the Covid-19 Pandemic and climate change. However, the challenge of linking location-specific requirements with potential design solutions hinders its effective implementation. To bridge this gap, this paper introduces a novel urban 15 Minute City concept generation tool that applies an artificial intelligence (AI) method called a pre-trained language model (PLM). The PLM model was fine-tuned with structured examples based on 15-Minute City principles. Using a PLM, the tool maps 15-Minute City concepts to a location and project specific prompt, automatically generating neighbourhood design concepts in the form of natural language.

Keywords. 15-Minute City, Neighbourhood Design, Data-driven Design, Urban Design, Natural Language Generation, Generative Pre-trained Transformer.

1. Introduction

The 15-Minute City is a residential urban concept in which most basic amenities are available by either walking or cycling from residents' homes, resulting in pedestrian friendly neighbourhoods, reducing the CO2 emissions, and improving the quality of life (Moreno et al., 2021). Despite growing design and planning interest in the 15-Minute City, the ability to explore and validate 15-Minute City designs is difficult for several reasons: First, it needs to be calibrated to location-specific requirements like climate and demographics; Second, domain specific expertise is required as part of the design team; and third, comparative evaluation of designs with respect 15-Minute City criteria is required. This paper presents recent advances in Artificial Intelligence (AI) towards 15-Minute City Design as part of a solution to address these challenges. In this

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 595-604.

paper, we describe a data-driven approach to retrieve data from C40 CITIES database (C40 | Reinventing Cities, 2022), fine-tune a natural language model and generate neighbourhood design solutions based on given conditions. The generated solutions were validated based on various metrics and the approach was then applied to real-world design challenges. The goal of the research is to explore the potential Pre-Trained Language Model (PLM) to create recommendation engines for neighbourhood design ideation in the form of natural language.

2. Related Work

The use of Natural Language Processing (NLP) in urban design is becoming more eminent. The widely available textual data in form of user generated social media reviews, thoughts, perceptions, and descriptions have created new opportunities big data analysis at a city scale (Chen & Hsu, 2020; Ataman et al., 2022; Aman et al., 2022; Chen et al., 2017). Research work has been done to review the application of NLP in urban design and studies, the wide range of applications include urban governance and management, public health, land use and functional zones, mobility, and urban design (Cai, 2021). Transformers are neural networks that recently showed notable success in handling various machine learning tasks such as Natural Language Processing (NLP), and computer vision tasks (Vaswani et al., 2017; Phuong & Hutter, 2022).

Recent advancements in the field of generative AI have been impressive and researchers in the design field have become increasingly interested. For example, generative pre-trained transformers (GPT) have been used for generating early phase design concepts by problem-driven reasoning and analogy-driven reasoning in the form of natural language (Zhu & Luo, 2022). This research was developed further to generate and evaluate light-weighted flying car concepts inspired by nature by fine-tuning large PLMs (Zhu et al., 2022). The similar concepts could be applied to the domain of urban design, the work focuses on making PLMs available to designers through a recommendation system to augment designers' creativity and knowledge base. This paper describes the use of GPT-3 as a virtual urban design expert who can come up with surprising, practical, and non-trivial design ideas for a 15-Minute City.

3. Methodology

The method to develop a recommendation system based on 15-Minute City using PLM consists of four parts: [1] collecting structured set of examples based on 15-Minute City; [2] clean the data and create input-output pairs of text; [3] fine-tuning the base PLM with new data; [4] a web-based user interface that allows the interaction with the PLM.

3.1. DATA COLLECTION

The dataset used to fine-tune the GPT-3 model was collected from winning entries of "Reinventing Cities" competition conducted by C40 Cities (C40 | Reinventing Cities, 2022). The goal of this competition was to stimulate sustainable development through innovative solutions to environmental and urban challenges. The teams included urban designers, architects, environmental experts, and real estate developers with domain expertise. The evaluation criteria of this competition align with the principles of the 15

Minute City framework, as both focus on zero-carbon, sustainable, and resilient neighbourhood design solutions that address a wide range of urban and climatic challenges. For example, the design proposals address urban challenges like social inclusion and community engagement, green space, and high-quality architecture and urban design, emphasizing on creating livable and equitable neighbourhoods. This competition had 33 excellent winning solutions based on site requirements from around the world. The textual dataset, shown in Table 1, has a diverse set of information, which can be used for different fine-tuning strategies. The competition brief for each city includes textual information like, site description, expected program, plot area, type of property transfer intended, site context and redevelopment goals, climate risks and environmental challenges, and planning rules and regulations, for neighbourhoods from cities around the world. The proposed solutions of the winners included solution description, key components, and presentation of the site.

Table 1. Data composition of each neighbourhood brief in the fine-tuning dataset

Components	**Description**
Neighbourhood location	The name of neighbourhood, city and country
Climate	The climatic conditions and environmental challenges
Expected program	The expected program or plan decided by the local authorities
Solution description	The brief introduction or description of proposed solution
Key components	The key components of design solution addressing the challenges
Presentation of the site	The site presentation to showcase relevance of solution to specifics of the site

3.2. IMPLEMENTATION

3.2.1. Prompt Structure

The textual data was reformulated and customised to create different input-output pairs, and based on these 4 main experiments were conducted. The input text consisted of the combination of site location, climatic conditions, expected program, and the total plot area, which was kept the same for all experiments and output text was different for different experiments as shown in Table 2. The input components were added into a text paragraph, to have similar structure across all data points.

Table 2. Data composition of each neighbourhood brief in the fine-tuning dataset

Experiment	**Output text**	**Types of Data**
Experiment I	Key Components	Paragraph
Experiment II	Solution Description	Paragraph
Experiment III	Presentation of the site	Paragraph
Experiment IV	Major Innovations	Keywords

3.2.2. Fine-Tuning

The base GPT-3 model was fine-tuned for different tasks to better understand the context and constraints. The structured dataset was used for fine-tuning the base model, with the API offered by OpenAI (OpenAI API, 2022). The data used for training the fine-tuned model, had to be in specified JSON Lines (JSONL) format, where each line is an input-output or prompt-completion pair corresponding to an example used for training.

3.2.3. User Interface

A simple user interface was created to facilitate this approach of generating design solutions. The user can input different parameters as text and click on generate to get novel and feasible neighbourhood design solutions. The user interface is shown in Figure 1.

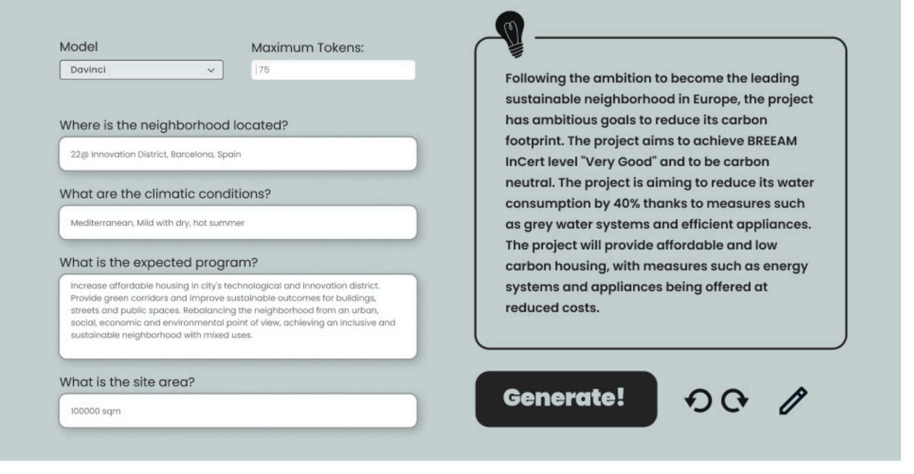


Figure 1. Digital Blue Foam's web-based User Interface for generating design proposals based on 15-Minute City principles

4. Results

GPT-3 offers 4 different main models, namely Davinci, Curie, Babbage, and Ada. Out of these four, three models were explored for experimentation, Davinci, Babbage, and Ada. Table 3 has the key components and solutions generated from different models for the same input prompt.

Table 3. Key components generated by different models with the same parameters for the given input prompt

Input Prompt
The site is located in 22@ Innovation District, Barcelona, Spain. The climate here is Mediterranean, Mild with dry, hot summer. The expected program is to increase affordable housing in city's technological and innovation district. Provide green corridors and improve sustainable outcomes for buildings, streets, and public spaces.

Rebalancing the neighborhood from an urban, social, economic and environmental point of view, achieving an inclusive and sustainable neighborhood with mixed uses. The site area is 100000 sqm.
Davinci Model Generation
The project is a neighborhood of 800 affordable housing units, including social housing and affordable housing. The project promotes positive energy and carbon neutrality. The project promotes biodiversity and climate resilience by creating green areas and plant-covered facades. The project promotes social inclusiveness through the creation of a new neighborhood center. The project promotes sustainable mobility by creating a large electric vehicle parking and charging station and promoting walking and cycling. The project provides new services and jobs and provides training programs to local residents.
Babbage Model Generation
The project promotes sustainable mobility and energy efficiency through a LEED certified building and a shared energy grid. The project reduces carbon emissions by 85% and achieves a net zero emissions thanks to energy production from renewable sources (90% wind- power and 40% solar panels). The project also contributes to the urban heat island index by reducing the surface area of the site by 10%. The project aims to optimize energy performance and reduce CO2 emissions by 80%. The project will be particularly efficient in terms of water.
Ada Model Generation
The project will increase the number of affordable housing units in the area and will also provide green corridors. Green walls will be installed on the building, which will promote the use of sustainable construction and green spaces. Greening the neighbourhood will contribute to carbon offset by offsetting the demand for green building materials. Greening the neighbourhood will promote the preservation of biodiversity, reduce the demand of invasive species, including native species. Biodiversity specialists will be in charge of managing and preserving the vegetation.

Table 4 consists of the results from the final experiment. Given the input prompts, GPT-3 generated major innovations as keywords, which could be used for further design ideation and conceptual phase.

Table 4. Major innovations extracted from proposals as keywords

Input Prompt	**Major Innovations**
The site is located in Tafelsig, Cape Town, South Africa. The climate here is Mediterranean, Mild with cool, dry summer. The expected program is significant mixed commercial and residential development of variable scale, density / intensity and typology.It should also include the necessary community use / civic elements and a (tenure-blind) gap / affordable housing component, together with the associated public realm improvements.The site area is 36000 sqm.	Living lab, affordable & inclusive housing, Zinc-battery Housing
The site is located in Balvanera Sur, Buenos Aires, Argentina. The climate here is humid subtropical, Mild with no dry season, hot summer. The expected program is to regenerate public space, streets and underutilized infrastructure throughout the site in a way that improves scenery, provides places where people can interact, improves the environment and promotes healthy lifestyles. Attention could also be given to how the city can help private land owners improve degraded buildings and improve their sustainability. The site area is 530000 sqm.	Solar panels, biosand filters, aquaponics system

4.1. EVALUATION CRITERIA

The rubrics used for the evaluation for generated neighbourhood design solutions are shown in Table 5. The solutions were evaluated based on feasibility and novelty ratings, going from 1-5. This evaluation rubric is inspired by the work of Data-Driven Innovation Lab at the Singapore University of Technology and Design, which developed this rating methodology to evaluate bio-inspired design ideas using pre-trained language models (Zhu et al., 2022). Based on these rating rubrics, a survey was conducted in the team of 4 human evaluators, who have worked in the field of urban design and planning. On average, 7 out of the 10 solutions generated using each of the base models made sense and could be designed. The analysis is presented in Figure 2 by base models. Comparing this, the results generated by Ada and Babbage base models got a good feasibility score but a lower novelty score, while the ones generated from the Davinci model were both more feasible and novel. Based on the generated solutions, it seems that the model understands the intent of the text.

Table 5. The Feasibility and novelty ranking rubrics used for the evaluation

Feasibility Rank	**Description**
1	The solutions make no sense from the 15-Minute City perspective
2	The solution makes little sense with today's technology, but could be possible in the future
3	The solution makes sense, but efforts are needed to work out a design roadmap
4	The solution makes good sense, and a design roadmap can be easily established to realise it.
5	The solution makes perfect sense and there are existing resources or methods to realise it.

Novelty Rank	**Description**
1	Neighbourhood solution exists and is commonly seen in the cities.
2	Neighbourhood solution exists but is uncommon in terms of a 15-Minute City.
3	New components and metrics are proposed for a 15-Minute City, but similar approaches or technology can be commonly seen in related urban contexts.
4	New features are proposed for a 15-Minute City, and similar approaches or technology can be rarely found in related urban contexts.
5	New features are proposed, and no similar design strategies, approaches or technology can be found nowadays.

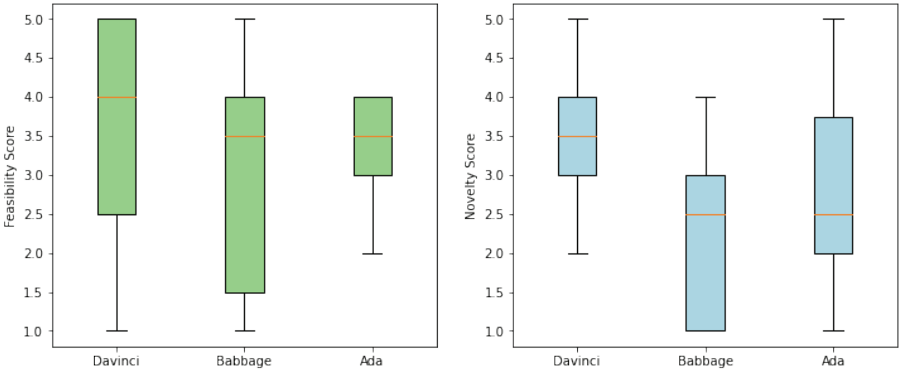


Figure 2. Human evaluation results based on the predefined rubrics for feasibility scores and novelty scores

The "Reinventing Cities" competition had specific guidelines and challenges that needed to be addressed. The solutions were evaluated based on their feasibility, specificity to site and 10 climate challenges (C40 | Reinventing Cities, 2022). Figure 3 categorises the GPT-generated neighbourhood design solutions based on the "10 challenges for climate" listed by C40 Cities in its competition (C40 | Reinventing Cities, 2022). The analysis shows a different distribution of categories for solutions generated by different base models.

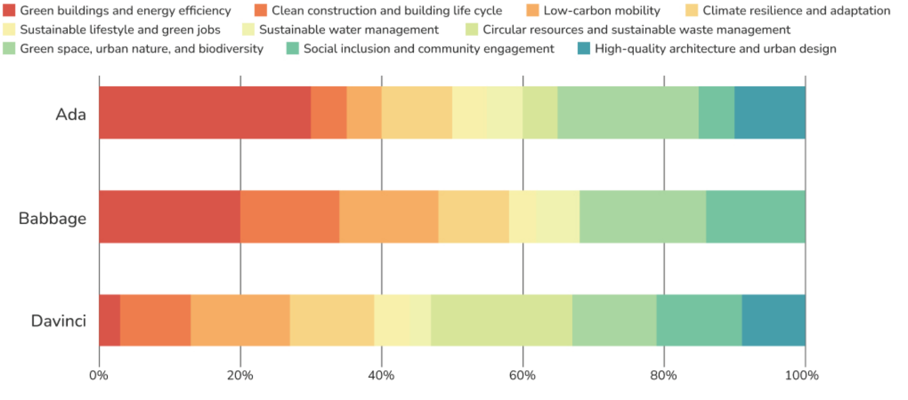


Figure 3. Categorical distribution of challenges for climate in generated design solutions

5. Case Study

This method was to generate key components and solutions for a site brief from Reinventing Cities: Student Competition (Students Reinventing Cities | Reinventing Cities, 2022) and the results were compared to the proposed winning entry. The site chosen for the case study was Gran San Blas Sur locality in Amposta, Madrid, Spain. The Davinci model was used as it was the best model, in terms of novelty and feasibility ranks. The generated solutions as shown in Figure 4 cannot be semantically distinguished from the solutions written by humans as shown in Table 6.

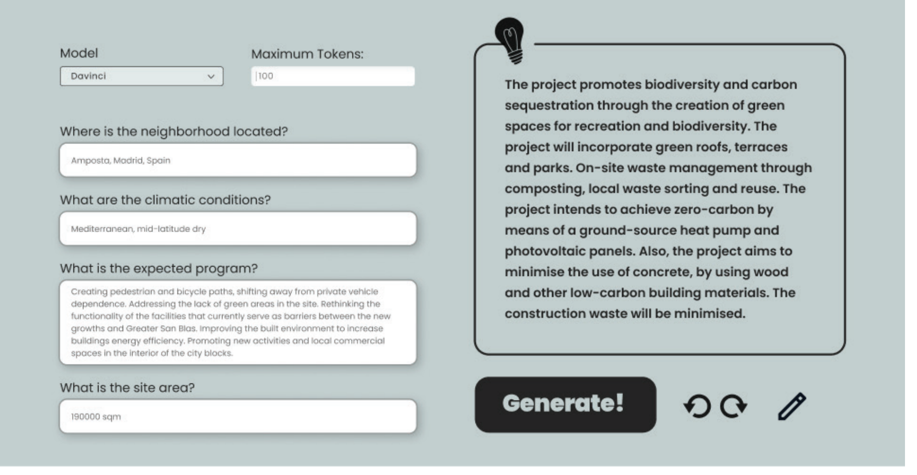


Figure 4. The GPT-3 generated design solution based on given inputs in the user interface

Table 6. The key components of winning solution proposed by students' team

Proposed Solution by Winners
The project has a safe bike and pedestrian infrastructure, encouraging car and bike sharing. Create inner-block car-free zones which require a well-connected public transport system. Improvement of the building envelope to increase energy efficiency, exchange of heating devices and other technical equipment, solar thermal energy and green roofs. The aim is to reduce overall energy consumption of the buildings in the neighbourhood by 80%. Development of green corridors and walls, roof gardens, spaces for community agriculture as well as bio-swales and rain gardens. Small-scale manufacturing opportunities for neighbourhood revitalization and growth of the local business sector. Opening of public facilities and services to improve playgrounds and sports.

6. Discussion

6.1. CHALLENGES

One of the main challenges faced during the fine-tuning of the PLM was finding a well-structured customised dataset. This structured real-world dataset had to be related to the 15-Minute City. Because PLMs like GPT-3 are capable of few-short learning, it only needed a few innovative, rational, and validated examples of neighbourhood design proposals. One of the other challenges was the generated samples are often semantically similar due to lack of diversity in the dataset. In terms of evaluation of results there is no defined statistical assessment approach regarding the performance and the quality of solution ideas generated.

6.2. NEXT STEPS

This paper explores a novel approach to generating design solutions for given constraints and requirements, though it has certain limitations. Firstly, the dataset chosen for fine-tuning GPT-3 was small. Even though the PLMs are capable of few-shot learning, having a set of around 200 quality examples would improve results considerably. The generated neighbourhood design solutions explored did match certain principles of the 15 Minute City framework, but a more comprehensive and real-world data repository is required for the model to follow the framework. Second, the generated design solutions were evaluated by human experts based on a limited set of metrics, which could lead to ideological bias and preconceptions. Future research may include using machine-based auto-evaluation techniques, such as using GPT-3 or advanced Large Language Models (LLMs) to evaluate generated concepts using frameworks set by urban design experts. LLMs are deep learning algorithms trained on massive textual datasets that interpret, translate, summarise, and generate text in the form of natural language. Third, this paper only explores GPT-3 as the PLM, even though it is good at generalization, it may not have the specialized domain knowledge required for tasks like urban design concept generation. Other PLMs can be explored too, like; Bidirectional Encoder Representations from Transformers, BERT (Devlin et al., 2019), which understands the sentiments, context, and meaning of the text better; or the T5 model developed by Google (Raffel et al., 2020), which can perform well in question-answering tasks. This can be taken further by training custom transformers with raw texts from scientific papers, journals, articles, and blogs written and documented by subject matter experts.

7. Conclusion

This paper introduces a new methodology of data-driven urban design using natural language generation. While this research paper explores the generative aspect of urban designing using the GPT-3, the next step would involve evaluating the generated solutions based on the 15-Minute City framework and logic developed by urban design experts. The transformer model could be trained on the large corpus of textual data that would include publications, journals, articles, and blogs to learn from expert methodology.

References

Aman, J., Matisziw, T. C., Kim, J. B., & Luo, D. (2022). Sensing the City: Leveraging Geotagged Social Media Posts and Street View Imagery to Model Urban Streetscapes Using Deep Neural Networks. 595–604. https://doi.org/10.52842/conf.caadria.2022.1.595

Ataman, C., Tunçer, B., & Perrault, S. (2022). Asynchronous Digital Participation in Urban Design Processes: Qualitative Data Exploration and Analysis With Natural Language Processing. 383–392. https://doi.org/10.52842/conf.caadria.2022.1.383

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D., Wu, J., Winter, C., … Amodei, D. (2020). Language Models are Few-Shot Learners. *Advances in Neural Information Processing Systems, 33, 1877–1901.* https://papers.nips.cc/paper/2020/hash/1457c0d6bfcb4967418bfb8ac142f64a-Abstract.html

C40 | Reinventing Cities. (2022). https://www.c40reinventingcities.org/en/

Cai, M. (2021). Natural language processing for urban research: A systematic review. Heliyon, 7(3), e06322. https://doi.org/10.1016/j.heliyon.2021.e06322

Chen, H., & Hsu, P. H. (2020). Data Mining as a User-oriented Tool in Participatory Urban Design. In L. C. Werner, & D. Koering (Eds.), Anthropologic - Architecture and Fabrication in the cognitive age (pp. 11-18). (Proceedings of the International Conference on Education and Research in Computer Aided Architectural Design in Europe; Vol. 1). Education and research in Computer Aided Architectural Design in Europe.

Chen, N.C., Zhang, Y., Stephens, M., Nagakura, T., & Larson, K. (2017). Urban Data Mining with Natural Language Processing: Social Media as Complementary Tool for Urban Decision Making.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), 4171–4186. https://doi.org/10.18653/v1/N19-1423

Moreno, C., Allam, Z., Chabaud, D., Gall, C., & Pratlong, F. (2021). Introducing the "15-Minute City": Sustainability, Resilience and Place Identity in Future Post-Pandemic Cities. Smart Cities, 4(1), Article 1. https://doi.org/10.3390/smartcities4010006

OpenAI API. (2022). https://beta.openai.com

Phuong, M., & Hutter, M. (2022). Formal Algorithms for Transformers (arXiv:2207.09238). arXiv. http://arxiv.org/abs/2207.09238

Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W., & Liu, P. J. (2020). Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer. Journal of Machine Learning Research 21, https://doi.org/10.48550/arXiv.1910.10683.

Romanov, A., Volchek, D., Chirkin, A., Mouromtsev, D., Sender, A., & Dergachev, A. (2018). Implementing a Natural Language Processing Approach for an Online Exercise in Urban Design. 15.

Students Reinventing Cities | Reinventing Cities. (2022). https://www.c40reinventingcities.org/en/students/

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is All you Need. *Advances in Neural Information Processing Systems, 30*. https://proceedings.neurips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html

Zhu, Q., & Luo, J. (2022). Generative Pre-Trained Transformer for Design Concept Generation: An Exploration. *In Proceedings of the Design Society (Vol. 2, p. 1834)*. https://doi.org/10.1017/pds.2022.185

Zhu, Q., Zhang, X., & Luo, J. (2022, November 11). Generative Pre-Trained Transformers for Biologically Inspired Design. *ASME 2022 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference.* https://doi.org/10.1115/DETC2022-90366

BOTTOM-UP APPROACH FOR CREATING AN URBAN DIGITAL TWIN PLATFORM AND USE CASES

A City Energy System Dataset Visualisation And Query

PRADEEP ALVA[1], MARTIN MOSTEIRO-ROMERO[2], WANYU PEI[3], ANDREA BARTOLINI[4], CHAO YUAN[5], RUDI STOUFFS[6]
[1,2,3,5,6]*National University of Singapore.*
[1,2,4]*Future Resilient Systems (FRS), Singapore-ETH Centre.*
[1]*akipaa@nus.edu.sg, 0000-0001-7997-1894*
[2]*mosteiro@nus.edu.sg, 0000-0001-7160-0298*
[3]*peiwanyu@u.nus.edu, 0000-0003-2822-6418*
[4]*andrea.bartolini@sec.ethz.ch, 0000-0001-8011-5454*
[5]*akiyuan@nus.edu.sg, 0000-0001-5404-5050*
[6]*stouffs@nus.edu.sg, 0000-0002-4200-5833*

Abstract. Smart city initiatives have been a driving force for city-level dataset collection and the development of data-driven applications that benefit effective city management. There is a need to demonstrate use cases for effective city management using the available dataset. Urban Digital Twin (UDT) is a 3D city model that can integrate multi-disciplines and improve systems operability on a digital platform. However, UDTs are developed within organisations, and there is only limited availability of authoritative open 3D datasets to explore the potential of UDT concepts. This paper reports a methodology for creating a UDT platform for visualising and querying city energy data. We demonstrate a bottom-up approach to constructing an integrated 3D city dataset and create a query system for rapid access and navigation of the 3D city dataset through a visualisation platform using *Cesium Ion*. Various use cases are explored based on the dataset, such as building material stock management, energy demand simulation, electric vehicles (EV) demand and flexibility, and estimation of greenhouse gas (GHG) emissions. These use cases can help decision-makers and stakeholders involved in city planning and management. Furthermore, it provides a guideline for developers willing to create UDT applications for smart city initiatives.

Keywords. Energy Modelling, City Dataset, Urban Analytics, Building Stock Management, Decarbonisation

1. Introduction

Smart city initiatives have been a driving force for city-level dataset collection and the

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 605-614.

development of data-driven applications that can help in effective city management. Smart cities have humans, technology and institutions as their three elements, with the environment, energy, transportation, safety, healthcare and education as fundamental disciplines. City sustainability, infrastructure, quality of life and service to the inhabitants are enhanced by smart city initiatives. A city-scale digital twin is one of the smart city initiatives that can integrate all disciplines mentioned above and improve systems' operability on a digital platform. Urban Digital Twin (UDT) is a 3D geospatial data model of a city consisting of physical assets, multimodal sensor data and bi-directional automated dataflow. Furthermore, planning and decision support is provided by UDTs to cities in terms of their administration, infrastructure, economic development, or citizen engagement (see Figure 1).

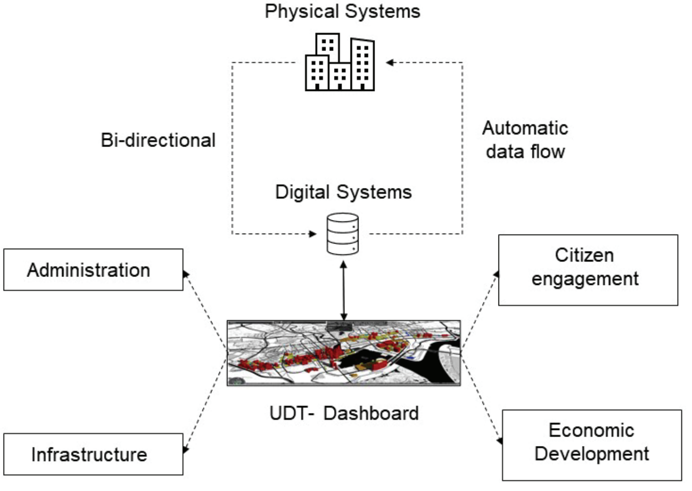


Figure 1: UDT framework

Urban Digital Twins can be used for various purposes (Alva et al., 2022a), such as archaeological restoration (Trento et al., 2020), urban flood simulation (Kikuchi et al., 2022), energy system resilience evaluation (Alva et al., 2022b), civic engagement, sustainability planning (Corrado et al., 2022) etc. UDTs tend to integrate more than one domain into their processes depending on the stakeholder's requirements.

Particularly, UDTs have been emerging for energy optimisation and operation purposes (Wang et al., 2022). Several research efforts and visualisation platforms based on the CityGML format serve as precedents to the evolution of UDTs for city energy systems (Rosser et al., 2019, Ishimaru et al. 2020). However, UDT use cases and applications that explore decarbonisation initiatives for cities are limited, as the applications restrict themselves to building level or below the urban scale. Limited access to sophisticated 3D data at city scale is one of the major reasons.

Most importantly, there is a need for methodologies which can help develop datasets to deliver UDT platforms for decision-making. More use cases along with their bottom-up methodologies have to be tested for developing UDTs for comprehensive decision-making. To address this issue, our research introduces a bottom-up approach to data modelling and processing to create an integrated 3D city dataset for the city of Singapore. This paper reports a methodology for creating such a UDT platform for visualising and querying city energy data. We demonstrate the development of four variously different use cases based on the created open city dataset, namely: building material stock management with urban-scale material flow analysis (MFA), energy demand simulation, electric vehicle (EV) demand flexibility, and estimation of greenhouse gas (GHG) emissions. In the following section, we elaborate on our methodology for creating the integrated 3D city dataset in four steps. We further explain our process of creating the UDT platform, querying the input based on the variety of planning areas, built years, and building typologies on the platform, and presenting the results in terms of city energy-usage patterns. In Section 3, we introduce various use cases that we developed based on the UDT platform created. In Section 4, we discuss future uses of the dataset and user experiences to support decision-makers and stakeholders, followed by a conclusion section.

2. Methodology

We describe our methodology for creating the integrated 3D city dataset in four steps, namely: Data Collection, Processing, Classification and Visualisation (see Figure 2). Our initial steps involve extracting and processing OpenStreetMap (OSM) data and combining them with the available building information to create a raw dataset. Specifically, building information such as building typology, year of completion, and building materials are retrieved along with the building footprint geometry.

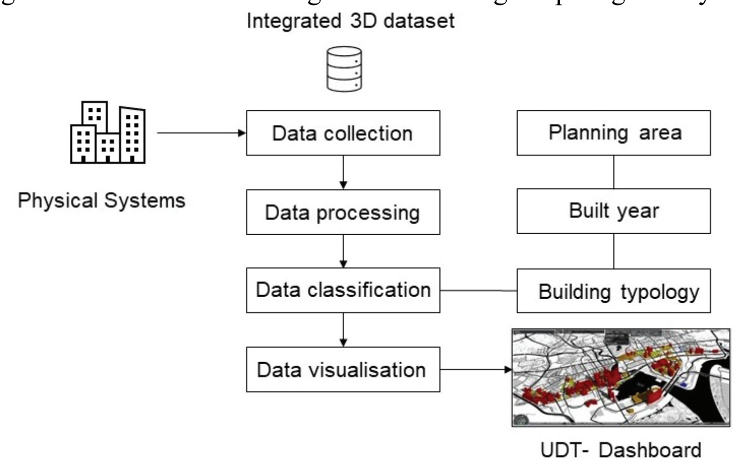


Figure 2: The four-step methodology for creating the bottom-up integrated 3D city dataset.

The data is further processed by additional information on yearly building energy consumption and building systems from an openly-available local dataset. After processing this raw data, it is classified in terms of Planning Area, Building Typology and Building Materials. This classification is used as a filter to input variables for users across the dataset. Subsequently, a query system is created for rapid access and navigation of the 3D city dataset through a visualisation platform using *Cesium Ion*. All the use cases and applications developed can be accessed in the visualisation dashboard along with the data queried from the open dataset (see Section 3).

2.1. DATA COLLECTION

Data Collection is the first step in creating the integrated 3D city dataset. We have screened many open data sources available in the Singapore context before assembling the dataset. Free 2D geographic database OpenStreetMap (OSM) and satellite building data with Digital Elevation Map (DEM) constitute our 3D geometry base with features representing building footprint along with the height. We combine OSM and satellite building data to create a building footprint geometry shapefile.

Planning Areas are defined based on the Master Plan 2019 Land Use layer and Subzone boundary available on Singapore's open data portal (Data.gov.sg). Building information such as building typology, year of completion and yearly building energy consumption are collected from various openly-available local datasets. Benchmarking Report 2021 from Building Energy Submission System (BESS) has 565 commercial buildings listed in a spreadsheet with building information as mentioned. Similarly, Housing Development Board (HDB) property information available on the Singapore open data portal has 12,496 residential buildings.

2.2. DATA PROCESSING

Raw data collected from various sources mentioned in Section 2.1 are processed to provide concise information for the use cases and applications. Along with 2D Building footprints and heights, OpenStreetMap provides a shapefile with metadata for building properties that include the building name, house number, street address and postal code. We orderly concatenate the columns representing building properties to create a complete property address for each feature. We use the newly created property address column to spatially join the attribute of building footprint geometry shapefile with BESS and HDB property information as mentioned in Section 2.1 using QGIS — a free and open-source cross-platform desktop geographic information system application.

2.3. DATA CLASSIFICATION

Planning area, building year and typology are the categories used to classify the collected data. These categories are further used to query the building stock for an assessment of energy consumption in the UDT dashboard. Additionally, this forms the hierarchy and order in the dataset, making it easier to navigate during user input queries.

There are 55 planning areas defined in the Master Plan 2019 Subzone Boundary (No Sea), along with large regional boundaries and smaller subzones (Data.gov.sg). The Join attributes by location — a data management tool in QGIS with a geometric predicate of 'intersects' — is used to transfer the planning area information to the base building footprint geometry shapefile. Each building is now classified into different Subzones, Planning Areas and Regions.

Building year information collected is categorised into several periods: before 1970, 1970-1979, 1980-1989, 1990-1999, 2000-2009, and 2010-present. Buildings are classified into their respective building typologies according to building function, such as academic, agriculture, civic, commercial, healthcare, industrial, infrastructure, mixed, Mass Rapid Transit (MRT), office, parking, port/airport, recreation, religious, reserve, residential, transportation, and utility. Building typology classification adds another layer of filter for the queries and furthermore helps in building applications such as building material stock management (refer Section 3.1) and energy demand simulation using City Energy Analyst (CEA) (refer Section 3.2).

2.4. DATA VISUALISATION

The UDT dashboard is developed as a web map application using *Cesium Ion* (Alva et al., 2022b). All the features collected with its information are converted to the 3D Tiles format for rapid streaming on the web. A query system with an input scroll bar of three categories specified in Section 2.3 is created for quick access and navigation to the dataset within the visualisation platform.

3. Use Cases based on the UDT Platform

3.1. BUILDING MATERIAL STOCK MANAGEMENT

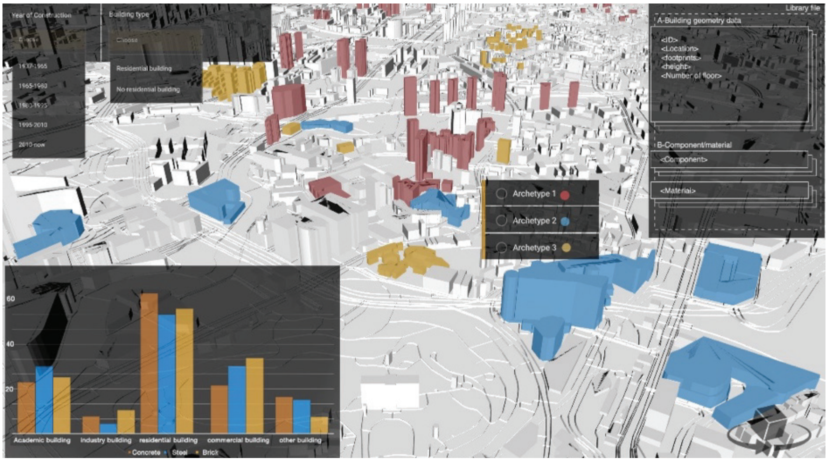


Figure 3:Quantifying building material stock by developing archetype databases.

Building material stock management is the first use case proposed based on the UDT platform. Material flow process and their spatial-temporal characteristics are essential to reducing the inflow of raw materials and the outflow of demolition waste. Material flow analysis (MFA) is an effective method for estimating and illustrating the construction material flows and stocks at the regional and urban level and developing a spatio-temporal model assessing the environmental impacts such as embodied primary energy, greenhouse gas emissions and water requirements of the whole building material system. Developing a Building Stock Model (BSM) is an essential and primary step of MFA. A bottom-up modelling approach is widely used to form the BSM, providing a detailed description of (individual) buildings' materials, and offers a better spatial differentiation than a top-down approach. However, the bottom-up approach requires a large amount of various data. Alternatively, archetype buildings can be defined as a virtual model representing the stock's most typical buildings.

Several parameters of buildings involved in the 3D city dataset can be used to identify the archetypes, such as the "year of construction" and "building type" because many characteristics (*e.g.,* material intensity) correlate to these two parameters. Each archetype is described by some characteristics related to material quantification, such as the number of floors, area of the floor, and window-to-wall ratio (WWR). After developing the archetypes, the mass/volume of each and the total building material can be calculated. The archetype-based building material stock model developed can be used to conduct MFA, which describes the potential environmental impact of the dynamics of material flow associated with the construction activities (*e.g.,* construction, renovation, demolition, etc.) of buildings at the urban scale and compare alternative scenarios. The UDT platform can represent and visualise the material stock and flows to identify geospatial constraints and determine hotspot areas with high potential for post-renovation emission reductions or with high concentrations of construction materials (see Figure 3).

3.2. ENERGY DEMAND SIMULATION USING CITY ENERGY ANALYST

In order to understand the distribution of urban-scale energy demands and GHG emissions, a detailed uncover of building energy demand patterns is required. However, as discussed in section 2.1, energy demand data is only available for a relatively small subset of the total building stock. Furthermore, this data is generally limited to annual energy use intensities, which limits their applicability for the assessment of specific energy and GHG emission reduction interventions.

Urban building energy modelling serves to estimate both the energy demands of buildings for which no energy consumption data is available, and future demands under different development scenarios thereby assessing the potential effects of building and system interventions. The UDT's dataset is used as an input to a building energy demand model developed using the City Energy Analyst (CEA), an open-source platform for building energy simulation and supply system optimisation.

The inputs to the CEA model comprise not only building geometries and typologies (which can be extracted directly from the dataset) but also building properties, building systems and operating parameters, as well as weather data. The BESS data provides information about a small subset of commercial buildings, namely: the type, age, and efficiency of the cooling systems; percentage of conditioned floor area and gross floor area; and average monthly building occupancy rate. All other parameters, and those for buildings not included in the BESS database, must be estimated. CEA includes building archetypes characterised by building typology, which are used as an initial estimate for the urban-scale building stock. For buildings with available energy consumption data, the model may be calibrated as done in a district-scale demonstrator for this UDT (Mosteiro et al., 2022).

Occupancy patterns in different areas of the city may be further estimated based on openly-available data on parking spot occupancy in different areas of the city (see section 3.3), while weather patterns may be extracted from openly available weather datasets. By incorporating such real-time data into the modelling workflow, the model can be updated in order to provide increasingly reliable energy demand estimates at the urban scale.

3.3. ELECTRIC VEHICLE (EV) DEMAND AND FLEXIBILITY POTENTIAL BASED ON CAR PARKING BEHAVIOURAL PATTERNS

Studying energy systems-related questions requires actual measurable data (such as the energy consumption pattern of a building), which theoretically can always be measured, but also projections of how ongoing trends which did not entirely unfold yet could impact such systems. An example of such trends is the ongoing uptake of electric vehicles (EVs) in substitution of conventional internal combustion engine-powered ones, which is expected to significantly impact urban electricity consumption and the distribution infrastructure. Hence, an additional use case for the UDT is to visualise scenarios of EV energy demand and potentially attainable flexibility.

EV-related electricity demand has been shown to be driven by the parking habits of users, with the typical behaviour of EV owners arriving at the parking facility, parking their vehicle and eventually plugging it in for a recharge. Hence, in the absence of actual EV data usage (such as is the case for Singapore), data that can describe parking usage patterns can be leveraged to model how EV loads could unfold once EV adoption becomes widespread. Furthermore, modelling can also be used to estimate the flexibility of such EV-related electricity demand: where flexibility in this context means shaping the electricity demand curve by shifting the load in time in order to provide different services to the local grid (Crozier et al., 2020).

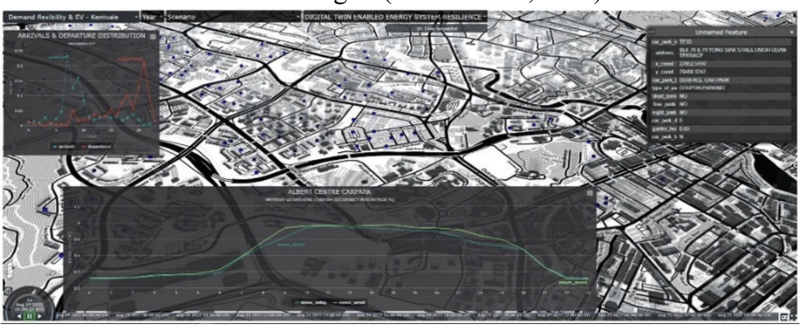


Figure 4:UDT dashboard showing the EV demand and flexibility potential based on car parking.

Towards such a goal, several datasets openly released by the Singapore government can be used in different ways towards building load and flexibility models. The most important among such data sources is the live availability of car park slots around Singapore, which is provided in the form of an open API (Data.gov.sg). Once queried, such API returns the live status of available and total slots for a large number of (approximately 1700) public car parks around Singapore, which is updated approximately every 30 minutes. This allows to infer realistic patterns of occupancy, departures and arrivals, which, combined with other datasets (typical daily mileage of vehicles, GPS location of car parks, etc.), allows to model scenarios for EV demand and attainable load flexibility under the assumption of a given EV penetration level (see Figure 4).

The actual real data and the models' outputs can be visualised in the UDT dashboard. Such options are valuable to different stakeholders, but especially to city and system planners. The shifting of mobility demand to electricity transmission and distribution infrastructure is expected to have significant impact and will often require upgrades to such infrastructure. Hence, the UDT platform and the embedded modelling can provide insights in what-if scenarios to understand better the implications of future trends, leading to more informed urban infrastructure planning.

3.4. ESTIMATION OF GREENHOUSE GAS (GHG) EMISSIONS

A greenhouse gas (GHG) inventory is created as a use case to calculate embodied and operational emissions based on the building energy use in the UDT platform. The GHG inventory must account for the release and usage of seven major GHGs: carbon dioxide (CO_2), methane (CH_4), perfluorocarbons (PFCs), hydrofluorocarbons (HFCs), sulfur hexafluoride (SF_6), nitrous oxide (N_2O), and nitrogen trifluoride (NF_3). In Singapore, GHG emissions in the atmosphere, which are released directly from both industrial processes and product use, and fuel combustion activities within the facility boundary and under the operational control of the corporations, are documented in accordance with the measurement and reporting requirements. The GHG emissions are further classified as reckonable and non-reckonable under the Carbon Pricing Act (CPA) established in 2019. However, we do consider indirect emissions from building electricity consumption which is excluded from CPA.

The Singapore Energy Market Authority specifies electricity grid emission factors (GEF) annually for the estimation of operational CO_2 emissions based on building energy-use intensities (EUI) (EMA, 2022). The Operating Margin (OM) GEF indicator measures the average CO_2 emitted per unit of net electricity generation in the system by all grid-connected power units. Generation technologies from combined cycle power plants, waste-to-energy, embedded co-generation plants and solar are included in the OM GEF calculation. Similarly, Upstream Fugitive Methane Emission Factor (UFME) is an indicator of emitted methane. The OM GEF and UFME for 2021 are 0.4057 kg CO_2/kWh and 0.0021 kg CH_4/kWh, respectively. We use these values to calculate the global warming potential (GWP) per kWh in CO_2 equivalents (CO_2e), assuming 25 kg CO_2-eq/kg CH_4. Furthermore, we use all the GHG emissions calculations based on their GWP to generate graphs for each building in the UDT dashboard (see Figure 5).

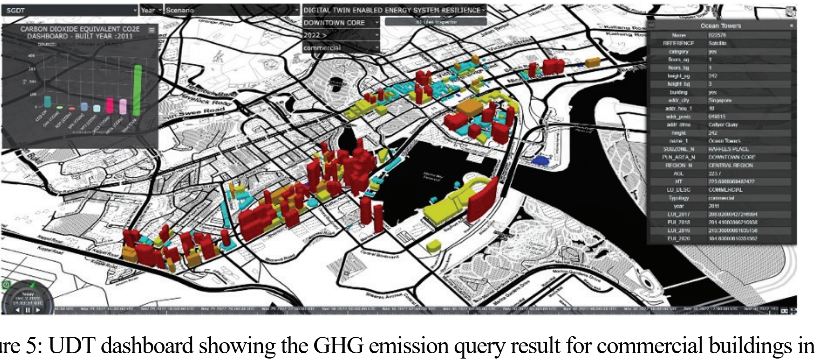


Figure 5: UDT dashboard showing the GHG emission query result for commercial buildings in the

Downtown planning area of Singapore. Various operational GHG emissions are displayed with their respective Carbon dioxide equivalent (CO2e) depicted in the bar.

4. Discussion

With the methodology introduced in Section 2, we developed an integrated 3D dataset for Singapore with 119,000 features representing its buildings and infrastructure. It is further classified based on planning area, building year and building typology. Such 3D datasets created particularly for energy systems can further be used to train machine learning models and study city consumption patterns. For instance, yearly building energy-use intensities can help forecast energy demands for assessing net-zero energy goals for cities. As explained, additional datasets related to the mobility sector (car park occupancy in this case) can help define scenarios regarding future trends in electricity demand, assess the impact on existing infrastructure, and prepare for potential future shocks to further enhance its resilience.

User Experience (UX) with real-time and historical data interaction is what sets UDTs apart from other visualisation platforms. Beyond visualisation, UDTs offer the latest empirical data to support stakeholders in the decision-making process and generate what-if scenarios for further analysis. The use cases presented in the paper provide scenarios that a user can explore and make decisions based on the displayed analysis results in the dashboard.

5. Conclusion

In conclusion, it is an overwhelming task to provide rapid access and navigation for a large 3D city dataset. Therefore, developing applications that can demonstrate use cases for effective city management using such dataset is important. However, large city datasets are usually not openly accessible, and bottom-up approaches to creating such datasets are limited. In this paper, we elaborate on the process of creating the UDT platform, querying the input based on three categories of planning area, built year, and building typologies on the platform, and presenting the results in terms of city energy-usage patterns. The use cases and application development proposed can help decision-makers and stakeholders involved in city planning and management in decarbonisation purposes. Furthermore, it can guide developers willing to create open 3D datasets and UDT applications for smart city initiatives.

Acknowledgements

This work is an outcome of the Future Resilient Systems project "Digital Twin Enabled System Resilience" at the Singapore-ETH Centre (SEC) supported by the National Research Foundation, Prime Minister's Office, Singapore under its Campus for Research Excellence and Technological Enterprise (CREATE) programme.

References

Alva, P., Biljecki, F., and Stouffs, R. (2022). Use Cases For District-Scale Urban Digital Twins. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. https://doi.org/10.5194/isprs-archives-XLVIII-4-W4-2022-5-2022

Alva P., Mosteiro-Romero M., Miller C., and Stouffs R. (2022). Digital Twin-based Resilience Evaluation of District-scale Archetypes: A COVID-19 Scenario Case Study Using a University Campus Pilot. In *POST-CARBON, Proceedings of the 27th CAADRIA Conference*, Sydney, Australia, 9-15th April 2022.

Corrado, C. R., DeLong, S. M., Holt, E. G., Hua, E. Y., and Tolk, A. (2022). Combining Green Metrics and Digital Twins for Sustainability Planning and Governance of Smart Buildings and Cities. *Sustainability* (Basel, Switzerland), 14(20), 12988. https://doi.org/10.3390/su142012988

Crozier, C., Morstyn, T., and McCulloch, M. (2020). The Opportunity for Smart Charging to Mitigate the Impact of Electric Vehicles on Transmission and Distribution Systems, *Appl. Energy*, 268, 114973, Jun. 2020, https://10.1016/j.apenergy.2020.114973

Deng, T., Zhang, K., and Shen, Z.-J. (2021). A Systematic Review of a Digital Twin City: A New Pattern of Urban Governance Toward Smart Cities. *Journal of Management Science and Engineering*, 6(2), 125-134. http://dx.doi.org/10.1016/j.jmse.2021.03.003

Ishimaru, N.; Kurokawa, C.; Tanaka, Y.; Oishi, T.; Akahoshi, K.; Kutzner, T.; Kolbe, T. H. (2020). CityGML Urban Planning ADE for i-Urban Revitalization. Open Geospatial Consortium 2020. Taken from: https://portal.ogc.org/files/?artifact_id=92113.

Ivanov, S., Nikolskaya, K., Radchenko, G., Sokolinsky, L., and Zymbler, M. (2020). Digital Twin of City: Concept Overview. In *2020 Global Smart Industry Conference (GloSIC)*, 78-186, https://10.1109/GloSIC50886.2020.9267879.

Kikuchi, N., Fukuda, T., and Yabuki, N. (2022). How a Flooded City Can Be Visualised from Both the Air and the Ground with the City Digital Twin Approach, System Integration of Flood Simulation and Augmented Reality with Drones. In *POST-CARBON, Proceedings of the 27th CAADRIA Conference*, Sydney, Australia, 9-15th April 2022.

Lee, A., Lee, K.-W., Kim, K.-H., and Shin, S.-W. (2022). A Geospatial Platform to Manage Large-Scale Individual Mobility for an Urban Digital Twin Platform. *Remote Sensing* (Basel, Switzerland), 14(3), 723. https://doi.org/10.3390/rs14030723

Mosteiro-Romero, M., Alva, P., Stouffs, R., and Miller, C. (2022). Towards occupant-driven district energy system operation: A digital twin platform for energy resilience and occupant well-being. In *PLEA 2022: Will Cities Survive?,* November 23-25 2022, Santiago, Chile. https://doi.org/10.5194/isprs-archives-XLVIII-4-W4-2022-5-2022

Rosser, J. F., Long, G., Zakhary, S., Boyd, D. S., Mao, Y., & Robinson, D. (2019). Modelling Urban Housing Stocks for Building Energy Simulation using CityGML EnergyADE. ISPRS International Journal of Geo-Information, 8(4), 163. https://doi.org/10.3390/ijgi8040163

Ruohomaki, T., Airaksinen, E., Huuska, P., Kesaniemi, O., Martikka, M., and Suomisto, J. (2018). *Smart City Platform Enabling Digital Twin*, 155-161.

Schrotter, G., and Hürzeler, C. (2020). The Digital Twin of the City of Zurich for Urban Planning. *PFG 88*, 99-112. https://doi.org/10.1007/s41064-020-00092-2

Shahat, E., Hyun, C. T., and Yeom, C. (2021). City Digital Twin Potentials: A Review and Research Agenda. *Sustainability* (Basel, Switzerland), 13(6), 3386. https://doi.org/10.3390/su13063386

Singapore open data portal. Data.gov.sg. Retrieved November 4, 2022, from https://data.gov.sg/.

The CEA Team (2022). City Energy Analyst v3.32.0. Zenodo. doi: 10.5281/zenodo.7317134

Trento, A., Kieferle, J. B., and Wössner, U. (2020). A Decision Making Tool for Supporting Strategies of Archaeological Restoration-Case Study of Ostia, Maritime 'Portus' of the Imperial Rome. In *Proceedings 38th eCAADe, Berlin: Anthropologic-Architecture and Fabrication in the cognitive age*, 1, 107.

Wang, W., Guo, H., Li, X., Tang, S., Xia, J., and Lv, Z. (2022). Deep Learning for Assessment of Environmental Satisfaction using BIM Big Data in Energy Efficient Building Digital Twins. *Sustainable Energy Technologies and Assessments*, 50, 101897.

ASSESSMENT OF THE VISUAL IMPACT OF A TALL BUILDING ON A UNESCO LISTED HISTORIC URBAN LANDSCAPE

KLARA CZYŃSKA[1]
[1] West Pomeranian University of Technology in Szczecin
[1]klara.czynska@zut.edu.pl, https://orcid.org/0000-0003-3855-6736

Abstract. This paper presents a methodology designed to study the visual impact of a tall building located in the vicinity of a protected area of exceptional cultural scenic value listed by the UNESCO World Heritage. Digital analyses of the city were carried out using the following methods: a) the Visual Impact Size (VIS), which allows to determine both the visual impact field and the domination of architectural objects in space; b) simulations of selected views from the human eyesight level. The proposed cityscape survey methodology is presented based on the example of the city of Toruń (Poland) in relation to the location of a planned high-rising building. The study used a 3D Digital Surface Model of the city (DSM). All simulations were carried out using proprietary software (C++).

Keywords. Digital Cityscape Analysis, Historical Cityscape, Heritage Protection, Unesco, Viewshed, 3D Isovist, Vis Method, Tall Buildings.

1. Introduction

This article describes a methodology to study the visual impact of a tall building located close to a protected area of outstanding cultural scenic value. These advantages often limit the possibility of erecting tall buildings in historic cities. This is because tall buildings are sizable, make a significant visual impact and can be seen over a distance of many kilometres (Van der Hoeven, Nijhuis, 2012). Moreover, tall buildings may adversely interact with historic buildings (Ashrafi, Kloos, Neugebauer, 2021). The use of 3D city models and computational methods offers the possibility to enhance the protection of cultural heritage in cities (Rubinowicz, 2020).

The aim of the research is to present a method that helps to identify how a new tall building fits into the historic cityscape. The study used the Visual Impact Size (VIS) method to investigate the visual impact field of the building (Czyńska, 2015). It also involved a series of simulations for selected panoramic exposures. A Digital Surface Model (DSM), a 3D model of the city, with high accuracy was used (Rubinowicz, 2018). The aforementioned methodology is presented using the example of the city of Toruń, Poland, which planned to develop a new high-rising building. The city centre, together with its wider buffer area, has been listed by the UNESCO World Heritage. The study identifies the visual relationship between the high-rising building and the Old Town complex. The study provides a guideline for the determining of the height, form, and the colour scheme for the new building.

2. Materials and Methods

2.1. VISUAL IMPACT SIZE (VIS)

The study of the visual impact of a tall building on the cityscape includes: a) visual impact analyses for the planned investment using the VIS method, b) identification of strategic vistas, c) simulations of the impact the investment has on the cityscape of Toruń within key vistas. A detailed description of the research has been included in the study report (Czyńska, Marzęcki, Rubinowicz 2020). Only selected observations and research results are presented in this paper.

The VIS method played a pivotal role for the research. The aim of the method is to provide an objective presentation of the extent and strength of visual impact (Czyńska, 2015). The basics of the method stem from "viewshed" (sometimes called 3d isovist), that is an area from which a target point is visible (Caha 2017). The novelty of the method is that together with the visibility range, the method calculates the degree of domination of a facility in space, namely its domination or only partial visibility of a building (i.e. the "size" of the building). The algorithm is based on a geometric analysis of a digital city model, more precisely the visibility of a pair of points in space. The simulation helps to develop a map showing all locations from which the planned facility is visible, depending on its height. The analysis of the visual impact range was used to determine possible locations of visual exposure of the project in the city. It provides an objective visualisation of the total impact range in Toruń and it delimits appropriate viewing zones. The analysis is also a basic tool for determining the location of strategic viewpoints, which are then documented and subject to further simulations.

2.2. DIGITAL SURFACE MODEL (DSM)

The analyses used the Digital Surface Model (DSM). Such models are available to the public for the entire area of Poland. They are provided by the Head Office of Geodesy and Cartography. Data have been collected under the flood protection project in 2013. DSM model, used for the purpose of the study, has mesh density of 50 cm. This enables creating a precise picture of the city space and creates a very good environment for conducting visual analyses (Rubinowicz 2020). It covers all cityscape components, such as buildings, tall greenery, bridges and overpasses, mapped with equal accuracy (for each component). The proposed method can also be applied using other types of city models, such as CityGML (Arroyo Ohori et al. 2018). However, compared to the DSM, it has one disadvantage, since most often it does not include tall greenery, which significantly affects the visibility of architectural facilities in space. Moreover, it is particularly important for the results of the research described in this article.

3. Case study

3.1. THE PLANNED BUILDING OPPOSITE THE OLD TOWN IN TORUŃ

The research in question was carried out for the city of Toruń, located in central Poland, Europe. The city has a very well-preserved medieval urban fabric. In 1952, the old town area was listed by the UNESCO World Heritage. Since then, the old town complex has been protected, together with its wide buffer zone shown in Figure 1

(Fig. 1). The planned tall building is to be located on the opposite side of the river, just outside the UNESCO buffer zone.

The analyses described in the article aimed to determine the visual impact of the planned building on the historic part of the city, locations from which it would be visible, and the strength of its exposure. Although the building is not located in direct contact with the old town, there is a visual connection between the two areas due to the building height (approx. 85 m). Although the areas are separated by the river and the densely tree clad Kępa Bazarowa, the planned tall building would project above treetops and command an attractive view of the old town. This represents a visual intrusion into the UNESCO buffer zone and a disturbance to the previously unspoilt natural landscape on the other side of the river (Fig. 2).

Figure 1. 3D model of Toruń and the UNESCO protection zone, UNESCO buffer zone, dense forest complex of Kępa Bazarowa, and the location of the planned tall building behind the buffer zone (above); historical panorama of Toruń (below)

3.2. SCOPE AND PROCESS OF DIGITAL ANALYSIS

The study was carried out for an area of 260 km^2 using the DSM (50 cm grid). The model was corrected to eliminate technical errors and updated to include changes in the city's built environment. All visualisations and analyses were developed using original C++ software developed with the involvement of the author.

Firstly, visual impact maps were prepared. The analysis was carried out in two variants: a) for the entire study area (260 km^2) on a 100 x 100 cm mesh, b) for the area around the centre (25 km^2) on a 25 x 25 cm mesh. Nine height ceilings were introduced:

30m, 40, 50m, 60m, 70m, 80m, 90m, 100m, and 120m, each had a different colour on the map assigned reflecting the exposure of the building. The larger part of the building could be seen from a given point, the more dominant it potentially was in that space.

Secondly, photographic documentation was developed for most of important strategic vistas of the planned building. Viewpoints were selected on the basis of the VIS map. Views from the old town area and the quays were particularly important. These were decisive in verifying how the new tall building would fit into the cityscape and whether it would disturb the harmony in its relationship with the old town area. Twenty-one strategic vistas were extracted and analysed. These included exposures listed by the UNESCO World Heritage.

Precise simulations were prepared for the selected strategic vistas while taking into account the presence of the future tall building in the landscape. Each simulation featured a rectangular block of the building with colour divisions every 10 m and a height ruler. This helped to distinguish the so-called Lower Visibility Threshold (LVT), i.e. the ceiling from which the studied facility is visible. The ruler also facilitates verification of the total height and other scenarios for the development of a tall building. An example of a simulation is shown in Figure 3 (Fig. 3). The VIS analyses and strategic view simulations helped to develop conclusions and guidelines for designing.

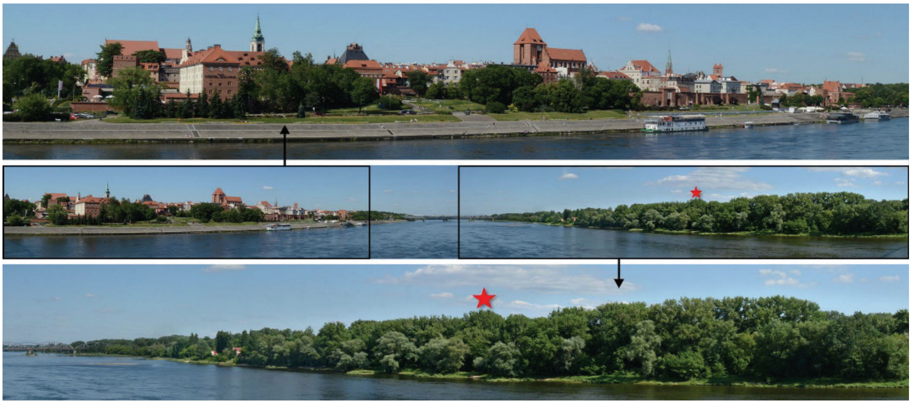


Figure 2. Panorama of Toruń seen from a bridge; grouping of trees at Kępa Bazarowa and the location of the planned tall building

Figure 3. Simulation of the planned tall building with the view from the riverfront. The simulation uses a height ruler and every 10 m colour coding for the facility

4. Results

4.1. RESULTS OF THE VIS ANALYSIS IN TORUŃ

The VIS maps helped to determine the extent and strength of visual impact. In general, it should be noted that the field of visibility is not particularly large. It mainly concentrates in the area of the river, which provides the view foreground. The most significant areas of exposure are therefore located along the right bank, immediately adjacent to the old town within a distance of approximately 1 km, as well as in the immediate vicinity of the site (Fig. 4). In the map, these areas are coloured red, orange and yellow, and they denote the visibility of the new building from 30, 40 and 50 m upwards. In areas of higher building density, the visibility of the proposed building significantly decreases or completely disappears. Therefore, the visual impact on the Old Town area (north of city walls) is limited to small areas and certain points. These have been identified as key exposures for the planned building and they became subject to simulation at a later stage.

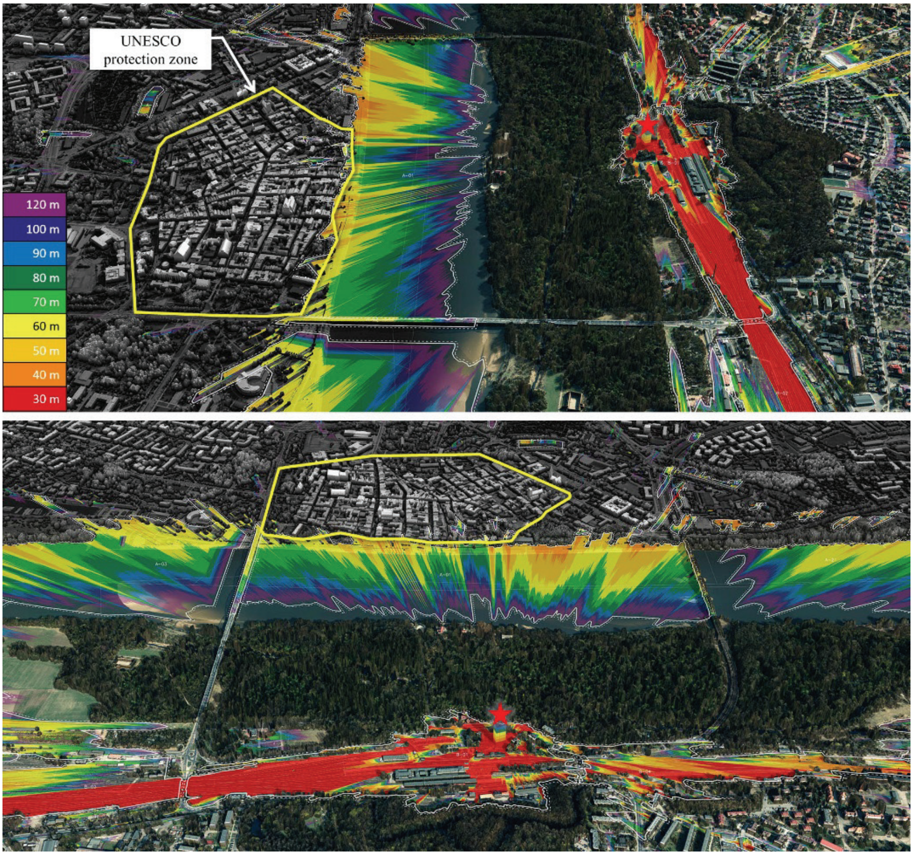


Figure 4. Visualisations of the 3D model and the VIS analysis.
Colours indicate the visibility of the planned facility from a specific height

4.2. SIMULATION RESULTS OF STRATEGIC VIEWS

A common feature of exposures studied are wide panoramic openings. This primarily applies to views from the riverside boulevards. These are exposures from the north, north-east or north-west (Fig. 3). Light access to the site therefore not very favourable. The planned building will usually be visible against the sun and mainly its contour can be discernible.

Given the height of the new tall building, which exceeds the visual barrier of trees on Kępa Bazarowa, the building will interfere with the exposures of the old town, but it will be separated from them by the width of the river. Views of this kind can be seen from bridges (including Piłsudski Bridge) (Fig. 2). The planned building should therefore be well balanced regarding its form, height, and façade finishing.

The views from the axis of streets leading from the Old Town towards the river form a special group of exposures. This sequence of images involves a gradual transition from a densely urbanised area to the open space of the Vistula waterfront and Kępa Bazarowa. This image is further emphasised by the need to cross a gate in the wall of the medieval fortification (Fig. 5a, c). The new tall building located on the other side of the Vistula River has a strong spatial (not necessarily visual) connection with the historic part of the city, not only in a sense that it can be captured in photographs and visibility analyses, but this also encompasses a wider spectrum of perception taking into account the afterimage that we have in mind after passing through the city gate. The analysis of the sequence and views from bridges indicated the necessity to relate the shape and façade material of the new building to the historical landmarks (Fig. 5b, d, e).

4.3. DESIGN GUIDELINES FOR THE PLANNED INVESTMENT

Based on the results of the analyses, guidelines for the design of the tall building were developed. They have to be made by the responsible urban planners (Stamps, 1996). Digital analytical techniques can only deliver a solid basis for decision making. Guidelines were grouped to relate to the following: building's immediate surroundings, its height and form, and architectural solutions for the façade. It was crucial to establish a formal dialogue with historic buildings of the old town regarding the form of the building and façade finishing. Following the example of historic towers, it was suggested that the silhouette of building should become slender towards the top. This solution reduces the domination of the building in the cityscape (Fig. 5d, e). A 100 m slender building appears visually smaller than a 70 m tall rectangular block. Furthermore, it was recommended that similar façade colours and material solutions should be used similar to those in the old town. The material in colour and texture should relate to ceramic architectural coatings used in the main historic buildings. The recommendations are a small part of the guidelines developed under the study (Czyńska, Marzęcki, Rubinowicz 2020).

5. Discussion

5.1. IMPORTANCE OF RESEARCH

UNESCO-listed urban areas are tangible examples of cultural heritage. They are 'non-renewable resources' that must be protected and preserved unchanged for future generations (Ashrafi, Kloos, Neugebauer, 2021). New buildings may pose the largest threat to the World Heritage. This applies especially to high-rising buildings which may cause partial or total degradation of the historic environment quality (Kuçak Toprak, 2020). In many cities around the world, we can find examples of the adverse impact of tall buildings on their historic surroundings. Often, tall buildings appear in the background of important vistas. In Istanbul, for example, they are behind the iconic Hagia Sophia (Akdag, Cagdas, Guney 2010), in London they disrupt the integrity of St. Paul's Cathedral (Czyńska 2018), in Paris such buildings disrupt the symmetry of one of the city's major compositional axes (Hollister 2013), and in Vienna they compete in the axial skyline of the Belvedere (Ashrafi, Kloos, Neugebauer 2021).

Figure 5. Axial exposure from the streets of the old town towards the river and the planned tall building (a, b, c); d) simulation of the tall building and a height ruler (building height 85 m); e) simulation of the building slender towards its top; colour of the tall building is a reference to brick facades of the old town (building height 100 m).

The Heritage Impact Assessment (HIA) is a procedure used to identify and analyse the potential impact of human-induced threats on cultural heritage and therefore the method supports better protection and management of heritage (Ashrafi, Kloos, Neugebauer, 2021; Seyedashrafi, 2017). Appropriate digital analytical techniques, such as those cited in the article, offer greater control over the future impact of new facilities on historic cityscapes. The techniques can support the implementation of a sustainable cityscape strategy.

The VIS method has a tremendous potential and has bene used in the planning practice in several cities in Poland. In studies for Gdańsk it was used to investigate a relatively low-rising but wide spread urban investment (Czyńska, 2019) and also a cluster of high-rise buildings and its visual impact on the cityscape. In study for Szczecin VIS analyses were an element of developing a strategy for tall building location in the city (Czyńska, Marzęcki, Rubinowicz, 2022). In study for Katowice reciprocal relations between the viewshed of planned facilities and historical dominants were examined (Czyńska, 2020). The range of potential applications is wide.

5.2. THE METHOD AND OTHER STUDIES

The analytical methods presented in the article focus on the visual impact of new buildings, including the extent and strength of their visual impact. Similar studies are usually based on viewshed analyses (Petrasova et al. 2018; Caha 2017). The theory associated with the measuring of the viewshed has developed for several decades (Tandy 1967, Felleman 1986, Bishop 2003). With the enhancement of computation capability and the availability of spatial data, it is often used, among other things, to analyse tall buildings. Results of the method are increasingly accurate (Tabik, Zapata & Romero 2013; Czyńska 2019).

Visual analyses can use various software environments. Karimimoshaver and Winkemann (2018) introduced ArcGIS to calculate the visibility on a single tall building by measuring the ratio of the visible area of the building to the visual field. Grasshopper and Rhino 3D are comprehensive tools that allow for geometric creation, simulation, and visualisation within one interface. Puspitasari & Kwon (2018) suggested the use of this software in the planning and designing of tall buildings in a cluster to shape the city skyline.

The research also covers the problem of three-dimensional visualisation of planned architectural facilities in the existing urban space. The method of mapping planned buildings (by means of spatial rulers) used in this research is an attempt to expand possibilities of using the usual visualisation of an architectural concept design (Pullar, Tidey, 2001). It can be used in the planning process to determine the desired height of a new building based on the analysis of its spatial context in a given exposure (Saedi et.al. 2019).

6. Conclusion

The research examines the use of digital techniques to assess the impact of a tall building on the historic protected cityscape. The proposed methodology is based on the combination of visual impact analysis using the VIS method and the simulation of selected cityscape exposures. The application of the method is presented on the case study for Toruń, Poland. The VIS method helped to identify a number of sometimes very distant visual correlations, correlations which are difficult to determine intuitively. Selected exposure locations identified on the VIS maps are examined using height lines. Simulations verified the total height and examined other scenarios regarding the form of the planned building. Comprehensive analyses of the relationship between the tall building and the historic skyline supported the development of guidelines for the height, form and the colour of the new building.

The analyses covered an urban area of 260 km^2 and used the full accuracy of the DSM model (50 cm grid). All visualisations and analyses presented were developed using C++ software. The author of the article was involved in the development of the software. However, the presented methodology is universal and can be used in the planning practice in different cities and in combination with other software.

References

Akdag, S. G., Cagdas, G., Guney, C. (2010). Analyzing the Changes of Bosphorus Silhouette. *28th eCAADe Conference Proceedings*, Zurich, 815-823 https://doi.org/10.52842/conf.ecaade.2010.815

Arroyo Ohori, K., Biljecki, F., Kumar, K., Ledoux, H., Stoter, J. (2018). Modelling Cities and Landscapes in 3D with CityGML. In Borrmann, A., König, M., Koch, C., Beetz, J. (eds) *Building Information Modeling*. Springer, Cham. https://doi.org/10.1007/978-3-319-92862-3_11.

Ashrafi B., Kloos M., Neugebauer C. (2021) Heritage Impact Assessment, beyond an Assessment Tool: A comparative analysis of urban development impact on visual integrity in four UNESCO World Heritage Properties, *Journal of Cultural Heritage*, Vol. 47, 199-207, https://doi.org/10.1016/j.culher.2020.08.002.

Bishop I. D. (2003). Assessment of visual qualities, impacts, and behaviours, in the landscape, by using measures of visibility. *Environment and Planning B: Planning and Design*, 30, 677-688. https://doi.org/10.1068/b12956

Caha, J. (2017). Representing buildings for visibility analyses in urban spaces. In: I. Ivan et al. (eds.), Dynamics in GIscience, Lecture Notes in Geoinformation and Cartography, Springer International Publishing. https://doi.org/10.1007/978-3-319-61297-3_2, 17-29.

Czyńska, K. (2015). Application of Lidar Data and 3D-City Models in Visual Impact Simulations of Tall Buildings. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 2015, XL-7/W3, https://doi.org/10.5194/isprsarchives-XL-7-W3-1359-2015, 1359-1366.

Czyńska, K. (2018). Tall buildings in historical urban context - analysis of selected examples, *Space & Form* 36: 159-176. https://doi.org/10.21005/pif.2018.36.C-04.

Czyńska, K. (2019). Visual Impact Analysis of Large Urban Investments on the Cityscape, Proceedings of the 37th eCAADe / SIGraDi Conference, Porto, Vol. 3, ISBN: 978-94-91207-19-8, pp. 297-304. https://doi.org/10.52842/conf.ecaade.2019.3.297.

Czyńska, K. (2020). Computational methods for examining reciprocal relations between the viewshed of planned facilities and historical dominants - their integration within the cultural landscape, Proceedings of the 25th International Conference of the Association for Computer-Aided Architectural Design Research in Asia, Vol. 1, ISBN: 978-988-78917-3-4, pp: 853-862 https://doi.org/10.52842/conf.caadria.2020.1.853.

Czyńska, K., Marzęcki, W., Rubinowicz, P. (2020). *Study of the landscape impact of the facility planned at Dybowska Street in Toruń*. The study commissioned by Borowskie Kopalnie Granitu i Piaskowca Skalimex Sp. z o.o.

Czyńska, K., Marzęcki, W., Rubinowicz, P. (2022). Protection and development of the cityscape and high-rise buildings based on the 2020–2021 Composition Study of

Szczecin. *Teka Komisji Urbanistyki i Architektury PAN Oddział w Krakowie,* Vol. L, *Cracow* 2022, ISSN 0079-3450.
Felleman, J. P. (1986). Landscape visibility. In: *Foundations for visual project analysis*. New York: John Wiley &Sons, 47-61.
Hollister N. (2013). The History of the European Skyscraper, *CTBUH Journal*, Issue II, 52-55.
Karimimoshaver, M., Winkemann, P. (2018). A Framework for Assessing Tall Buildings' Impact on the City Skyline: Aesthetic, Visibility, and Meaning Dimensions. *Environmental Impact Assessment Review* No. 73: 164-176. https://doi.org/10.1016/j.eiar.2018.08.007.
Kuçak Toprak, G. (2020). A guideline proposal for determining design criteria of new building designs in historical cities: ICOMOS and UNESCO. *Social Science Development Journal* Vol. 5, Issue: 22, Pages: 95-116. https://doi.org/10.31567/ssd.284.
Petrasova, A., Harmon, B., Petras, V., Tabrizian, P., Mitasova, H. (2018). Viewshed Analysis. In *Tangible Modeling with Open Source GIS*. Springer, Cham. https://doi.org/10.1007/978-3-319-89303-7_9.
Pullar, D. V., Tidey, M. E. (2001). Coupling 3D visualisation to qualitative assessment of built environment designs. *Landscape and Urban Planning*, Volume 55, Issue 1, pp. 29-40, https://doi.org/10.1016/S0169-2046(00)00148-1.
Puspitasari, A. W., Kwon, J. (2021). A reliable method for visibility analysis of tall buildings and skyline: a case study of tall buildings cluster in Jakarta. *Journal of Asian Architecture and Building Engineering*, 20(3), 356-367. https://doi.org/10.1080/13467581.2020.1787839
Rubinowicz, P. (2018). Application of Available Digital Resources for City Visualisation and Urban Analysis. *Computing for a better tomorrow: proceedings of the 36th International Conference on Education and Research in Computer Aided Architectural Design in Europe*, Vol. 2, ISBN: 978-94-91207-16-7, 595-602.
Rubinowicz, P. (2020). Sustainable development of a cityscape using the Visual Protection Surface method - optimization of parameters for urban planning. In *25th International Conference of the Association for Computer-Aided Architectural Design Research in Asia CAADRIA 2020*, The Association for Computer-Aided Architectural Design Research in Asia (CAADRIA). https://doi.org/10.52842/conf.caadria.2020.1.863.
Saeidi, S., Mirkarimi, S.H., Mohammadzadeh, M., Salmanmahiny, A., Arrowsmith, C. (2019). Assessing the visual impacts of new urban features: coupling visibility analysis with 3D city modelling. *Geocarto International* 34(12), pp. 1315-1331. https://doi.org/10.1080/10106049.2018.1478891.
Seyedashrafi, B., Ravankhah, M., Weidner, S., Schmidt, M. (2017). Applying Heritage Impact Assessment to urban development: World Heritage property of Masjed-e Jame of Isfahan in Iran. *Sustainable Cities and Society*, Vol. 31, 2017, pages 213-224, https://doi.org/10.1016/j.scs.2017.01.002.
Stamps, A., E. (1996). Significant visual impact: Is it or isn't it? Timmermans, H. (Ed.), *Third Design and Decision Support Systems in Architecture and Urban Planning - Part two: Urban Planning Proceedings, Belgium*.
Tabik, S., Zapata, E., Romero, L., (2013). Simultaneous computation of total viewshed on large high resolution grids. *International Journal of Geographical Information Science*, 1-11.
Tandy, C. R., (1967). The Isovist Method of Landscape Survey. In C.R. Murray (ed.) *Methods of Landscape Analysis*. London: Landscape Research Group, 9-10.
Van der Hoeven, F. D., Nijhuis, S. (2011). Planning and Visibility Assessment of High Building Development in The Hague. *High-rise and the Sustainable City*. https://doi.org/10.7480/rius.2.216

A DATA-DRIVEN APPROACH FOR INTERPRETING HUMAN PREFERENCE IN URBAN PUBLIC SPACES

A Case Study in London

YUXUAN LIU[1]andWENZHAO XU[2]
[1] *University College London*
[2] *University of Pennsylvania*
[1] *1ucbq741@ucl.ac.uk, 0000-0003-3160-5750*
[2]*wenzhao@alumni.upenn.edu, 0000-0002-0055-1393*

Abstract. Sound, an often easily overlooked part of the urban composition, plays a crucial role in urban studies regarding how people perceive urban acoustic environments. Influenced by variable natural and artificial acoustic elements, the multi-level nature of urban soundscapes reflects multidimensional urban characteristics and thus demands thorough investigations. Unravelling the acoustic complexity and understanding its influence on public spaces' popularity will benefit the interpretation of urban soundscapes and help to build a more vibrant public life. In this paper, we utilize DepthMapX and QGIS to select specimen sites for further investigation. The behavioural analysis is performed through principle components analysis (PCA) and K-Means analysis based on data obtained within the city of London. This project seeks to develop digital strategies to investigate the relationship between urban public spaces and their soundscapes: to infer the correlation between people's emotions and the richness of urban environments. Through integrating machine-learning analytic approaches into visualizing how citizens perceive or experience the urban sound environment, our research aims to positively assist and inform urban practitioners in planning and designing more benign acoustic environments.

Keywords. Urban Soundscape, Data-driven Analysis, Machine Learning, Multidimensional Data visualization

1. Introduction

Sound is fundamental to people's perception of the city, and people subconsciously or 'compulsively' absorb all kinds of sounds with diversified acoustical properties around them. While people are sensitive to certain kinds of acoustic changes such as traffic noise in urban areas, a majority of sounds existing in everyday life are buried without notice. These sounds, whether being noticed or ignored constitute the urban soundscape (Porteous and Mastin, 1985). The term "soundscape" was first coined by

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 625-634.

urban planner Michael Southworth and popularised by R. Murray Schafer as the acoustic environment in which people experience, understand or perceive (International Organization for Standardization [ISO], 2014).

Similar to the vast scope of sounds that attract no attention, the urban soundscape is often underappreciated by urban designers and researchers (Hong and Jeon, 2015). Those that caught attention are often pertaining to the emergence of noise, and more specifically, traffic noise (Jiang and Nellthorp, 2020). Urban noise impacts all aspects of human life, including economic and cultural life and physical and mental health (World Health Organization. Regional Office for Europe, 2011). Part of the urban noise generation is determined by the cumulative effect of irrational urban design elements in the city, in which noise can be destructive to stable urban planning (Nieuwenhuijsen and Haneen Khreis, 2018). Therefore, it poses great demands on the urban soundscape for people to savour a better living environment and a more sustainable urbanity.

Due to the COVID-19 pandemic, there has been a decrease in urban transportation and commercial activity as more individuals are working from home. Consequently, urban public space utilization has decreased significantly, along with fewer social, economic, and recreational activities. The pollution caused by urban noise and the number of noise-related complaints has decreased as a result as well (Bergan, 2020). Where urban investigations reveal a significant decline in human-related building sounds in traffic and construction sites, there is a rise in wildlife sounds like bird melodies (Miranda, Halpern, and Duarte, 2020). Individuals prefer these natural sounds over artificial ones, with traffic and construction noise being the most despised soundscape elements (Ma, Mak, and Wong, 2021).

Given the diversity of the soundscape, the perception of the urban acoustic environment requires not only the assessment of the physical features, but also the psychoacoustic measurements from the listeners. Previous research mostly focuses on gathering data through interview-based investigations, which is rather limited and only describes a narrow aspect of people's perception of urban sounds. It is thus vital to embed digital technology to effectively capture and analyse data to examine the relationship between people's interactions with various soundscape combinations. In this study, London is chosen as the study area regarding to the complexity and diversity of its urban soundscapes. The study is conducted in three steps: 1）Gather data of urban sounds and people's emotions; 2) Identify and analyse sites' physical features through DepthMapX and GIS; 3) Process and visualize the analysed data using data processing techniques PCA and K-Means from the machine learning community. This approach allows for a convenient integration of the study of human perceptions into the acoustic considerations of sound identification and classification, revealing the correlation between soundscape diversity and people's preference of urban public spaces. The research intend to serve as a practical reference for the design and adjustment of urban spatial soundscapes, as well as an important stepping stone for the future development of urban sound environments.

2. Methodology

We develop a methodology dedicated to capturing and interpreting the urban sound environment, assessing the correlation between soundscapes, human activities, and spatial openness. The sound and emotional data obtained from social media constitute the dataset for further analysis and research. To analyse the data in the urban dimension, we utilize QGIS and Space syntax tools to establish the base map for the multidimensional data visualization. Then we introduce Principal Component Analysis (PCA) and K-Means, the core machine-learning tools used in our approach, to analyse the selected data. PCA is a dimensional reduction algorithm, which condenses the dataset into major components to transform them into a more manageable form for further analysis (Ringnér, 2008); the K-Means cluster heatmap is an indication of how different datasets are represented within different clusters (Sinaga and Yang, 2020). In this approach, we sort the soundscape and emotional data sequentially to obtain a distribution of the public spaces that has maximum gains in positive human reactions with the high complexity of soundscape composition.

2.1. DATASET RETRIEVAL

To allow for a better understanding of how the city is perceived by various sounds, we first identify urban traffic noise and natural noise in the urban soundscape dataset. To obtain a macro-scale map of sound composition and distribution, we gather data containing specific tags and geolocation from social media, such as Twitter and Flicker. Data from social media usually carries location information, e.g. the exact location where the post has been made public (latitude and longitude). According to the first urban sound dictionary (Aiello et al., 2016), the tags include transport, nature, human, music, indoor, and mechanical aspects. A more detailed tag sets of which are listed in Table 1.

Table 1. Urban Sound Taxonomy

Indoor	shower, toilet, window, church, organ, computer, glitch, paper, slam, bathroom, bottle, cooking, door, elevator, flush, game, glass, knock, office, sanitario
Natural	animal, animals, atmosphere, wind, bark, bird, birds, birdsong, cat, chirp, countryside, crickets, dog, eating, farm, forest, frog, frogs, growl, howl, howling, leaves, nature, park, roar, trees, wood, woods, agua, beach, drip, dripping, fountain, lightning, liquid, ocean, playa, rain, river, sea, shore, splash, storm, thunder, thunderstorm, water, waves
Human	choir, chatter, breath, breathing, yell, yelling, voices, voice, talking, talk, spoken, speech, speaking, speak, shout, scream, screaming, fight, fart, cry, crowd, cough, coughing, conversation, baby, child, children, kid, kids, kitchen, laugh, laughing, laughter, school, footstep, footsteps, human, people, running, street, walk, walking
Music	acoustic, applause, arpeggio, audience, audio, bass, beat, breakbeat, cello, chordophone, clap, clapping, clarinet, club, cymbal, cymbals, dance, dj, drums, drum, drumset, dubstep, flute, group, hiphop, hip-hop, instrument, jazz, keyboard, melodic, melody, microphone, music, orchestra, percussion, percussive, piano, plane, playing, resonance, rhythm, rhythmic, sax, saxophone, sing, singing, song, synth, synthesizer, trumpet, violin, vocal, vocals, whistle, woodwind, resonant, recording, radio
Mechanical	clang, construction, factory, industrial, machine, machinery, mechanical, tools, alarm, bell, bells, carillon, chimes, chime, clock, ring, ringing, crash, drill, explosion, fire, fireworks, hammer, hammering, impact, jackhammer
Transport	autobus, beep, bike, bus, camion, car, cars, coach, coaches, engine, motor, race, road, traffic, truck, vehicle, underground, tube, subway, reverb, railway, rail, metro, aeropuerto, air, aeroplane, airport, helicopter

Using a similar approach, we gather social media contents on emotion as the parameter to evaluate how people react to surrounding soundscapes. All words in social media with emotional and geolocation hashtags are filtered and captured with, e.g. "I have an unsatisfied visit at … plaza". These labels are developed following Paul Ekman's research's six fundamental emotions, including fear, sadness, anger, love, surprise, and joy (Table 2).

Fear	alarm, shock, fear, fright, horror, terror, panic, hysteria, mortification, horror, anxiety, suspense, uneasiness, apprehension, fear, worry, distress, dread, nervousness
Sadness	agony, anguish, hurt, suffering, depression, despair, gloom, glumness, unhappiness, grief, sorrow, woe, misery, melancholy, sadness, dismay, displeasure, disappointment, guilt, regret, remorse, shame, alienation, defeatism, dejection, embarrassment, homesickness, humiliation, insecurity, insult, isolation, loneliness, rejection, neglect, pity, mono no aware, sympathy
Anger	aggravation, agitation, annoyance, grouchy, grumpy, crosspatch, imitability, frustration, exasperation, anger, outrage, fury, wrath, hostility, ferocity, bitterness, hatred, scorn, spite, vengefulness, dislike, resentment, rage, revulsion, contempt, loathing, disgust, jealousy, envy, torment
Love	adoration, fondness, liking, attraction, caring, affection, tenderness, compassion, sentimentality, lust, sexual desire, compassion, sentimentality, longing
Surprise	surprise, amazement, astonishment
Joy	amusement, bliss, gaiety, glee, jolliness, joviality, joy, delight, enjoyment, gladness, happiness, cheerfulness, jubilation, elation, satisfaction, ecstasy, euphoria, zest, enthusiasm, zeal, excitement, thrill, exhilaration, contentment, pleasure, pride, triumph, optimism, eagerness, hope, enthrallment, rapture, relief

Table 2. Emotions Taxonomy

The visualization of multidimensional urban data is non-trivial due to the complex intertwining of different layers of information (e.g. London Borough, London Lower Layer Super Output Areas). In the field of urban design and research, QGIS is widely adopted as an effective method for relaying and mapping urban information, visualizing multidimensional data in a geologic map for site selection, site land suitability analysis, land use and transport modelling, the identification of planned action zones, and impact assessment (Menke et al., 2016). Remapping and resampling data through QGIS allows us to restructure the vast quantity of data linkages into two or three dimensions that are more convenient for us to define, interpret, and analyse. As shown in Figure 1, we compute pixel values by mapping a spatially varying dataset into a 200 by 200 grid with QGIS. This allows all data to be presented within the same pixel system, establishing the base map for all the subsequent steps of our approach.

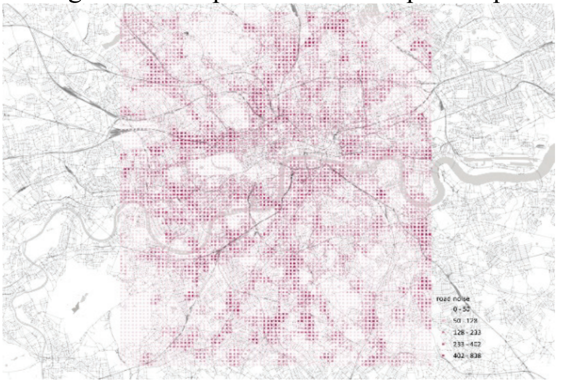


Figure 1. Road noise resample visualization (This mapping is based on the Lden descriptor, which indicates a 24-hour annual average noise level with separate weightings for the evening and night periods.)

2.2. SPATIAL FEATURES AND SITE SELECTION

The characterization of sound environments is commonly influenced by spatial features. For instance, a place with higher traffic and building densities is related to lower sound permeability and more prosaic acoustic environments. In this research, the accessibility of an urban area is defined as the "openness", which affects the road noise in London. We utilize DepthMapX to analyse and visualize the openness in central London to examine the linkage between soundscape and human preference. Space syntax is one of the most important analytical theories and methods for spatial arrangement and urban morphology (Hillier et al., 1976), and we use DepthMapX for spatial syntactic analysis and variables derivation. Approaching the issue regarding the centrality of specific urban areas, Choice and Integration maps are used to evaluate the location of different phenomena. Specifically, Choice quantifies the number of times each spatial unit in a spatial system is crossed by the path with the least angular transition between any two other spatial units, indicating its traverse ability (Van Nes, 2021). Integration is an algorithm that calculates each spatial unit's angular distance from other spatial units in a spatial system, determining its centrality or potential as a trip destination (Van Nes, 2021). Through the spatial network Integration and Choice analysis, the structure of primary and secondary roads is revealed, underlying the overall London-scale circulation hierarchy and the neighbourhood-scale street organization (Figure 2).

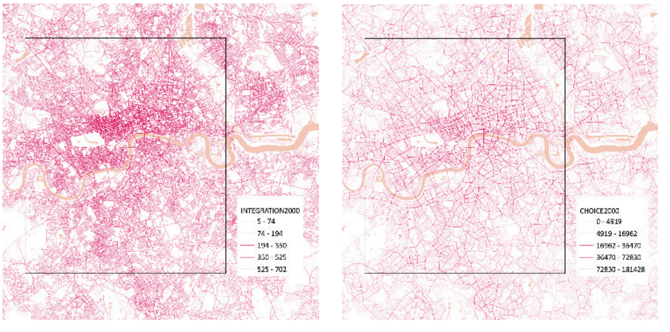


Figure 2. Urban openness in London - Integration and Choice

In order to investigate the variety of sound and the emotional inclinations of individuals toward urban spaces, we choose six metropolitan public spaces with comparable levels of integration and choice. Figure 3 illustrates the openness degree of selected sites.

Figure 3. Site selection

2.3. DATA ANALYSIS

We utilize data processing tools from machine learning, specifically PCA and K-Means, to analyse the urban sound dataset we have collected for further comparative studies at the selected sites.

The principal component analysis is an effective technique for analysing high-dimensional data with complex attributes. This method lends itself to a suitable candidate to extract the main features of a dataset (Ringnér, 2008). Concretely, PCA computes new feature directions that lead to the most changes in the data space, and new attributes could be computed from the linear combination of these principal directions. The advantage of leveraging PCA is two folds. The first is to identify features that are noticeably different from the rest of the data. In other words, extract qualities that aim to most accurately represent the variability in the data (Ringnér, 2008). The second point is to allow for a concise representation of the original data attributes. On a technical level, PCA analysis achieves both points by reducing n-dimensional features into k-dimensional features, where k is a much smaller number than n. These k-dimension vectors are referred to as principal components, which could then be used to partially or fully reconstruct the initial data (Abdi and Williams, 2010). From the results of the PCA analysis, we filter the most representative data group-the PC1 group and extract the first seventeen principal components out of the twenty-eight data categories of acoustic and emotional data (Figure 4).

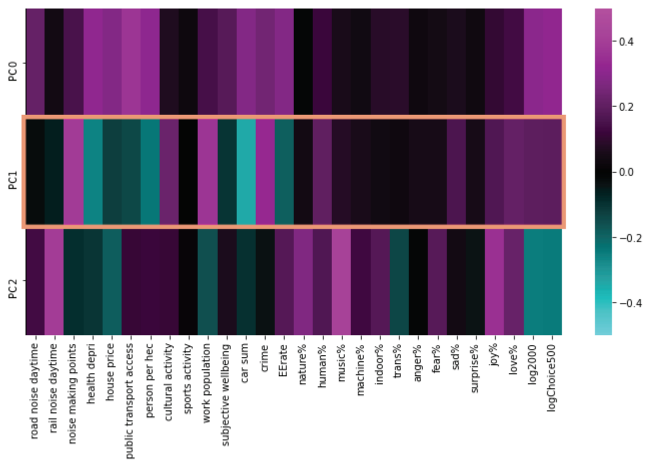


Figure 4. The principal component analysis heatmap (The degree of positive and negative correlation of each component is illustrated by colour saturation.)

After PCA analysis, we import the extracted dataset into the K-Means algorithm to further explore the correlation between soundscapes and emotions. K-Means is a fairly typical example of unsupervised learning (Sinaga and Yang, 2020). A wide range of industries, including banking, cybersecurity, document clustering, and image segmentation, uses K-Means clustering to analyse small-dimensional, continuous, numeric data. Urban planners mostly use PCA and K-Means clustering techniques to determine the urban identity of public spaces (Chang, Bus, and Schmitt, 2017). In this research, the cluster properties of K-Means facilitate us to test and determine how sound and emotion are distributed in clusters within the same public space.

The data distribution is described as five clusters based on the K-Means results (Table 3), which are visualized and remapped by QGIS (Figure 5). Cluster 1 and 4 are considered as problematic clusters due to higher noise levels as well as negative emotions, while 2 and 3 are considered dominant clusters with lower noise levels.

Table 3. K-Means Clusters Results

Cluster	Data characteristics	Value	Reference range
Cluster0	Most data with the lowest value	/	/
Cluster1	Highest railway noise	12.3291	4~13 Lden
	Highest fear emotion	33%	100% (All "Fear" data)
Cluster2	High road noise	13.5205	50~140 Lden
	High human sound	29.95%	100% (All "Human sound" data)
	High music sound	5%	100% (All "Music sound" data)
	High transport sound	17.45%	100% (All "transport sound" data)
	High joy emotion	20.3%	100% (All "Joy emotion" data)
	High love emotion	35.9%	100% (All "Love emotion" data)
Cluster3	Nothing Special	/	/
Cluster4	Highest road noise	136.6679	50~140 Lden
	More sad emotion	5%	100% (All "Sad emotion" data)

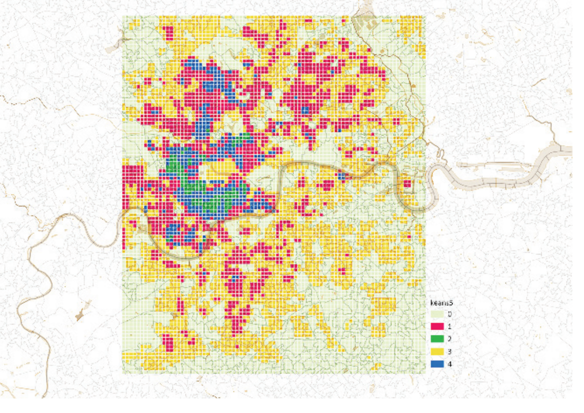
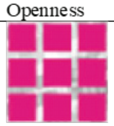
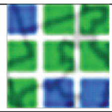
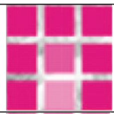
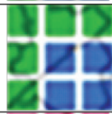
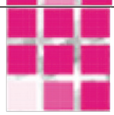
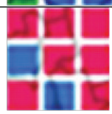
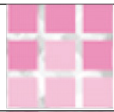
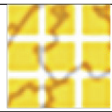
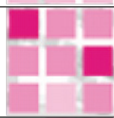
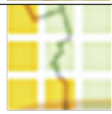
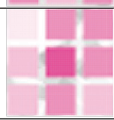
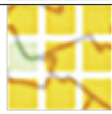


Figure 5. K-Means cluster visualization

The analysis is then carried out accordingly within the six sites we selected (Table 4). Table 4 illustrates that the richness of the urban soundscape is greatest in regions where people have the most emotions, such as sites 1 and 2. In contrast to sites 2, the distribution of emotions at site 3 are mainly apprehensive feelings. Despite similar urban openness, soundscape composition at site 3 is dominated by noise with low acoustic diversity. The statistics for less open regions, such as sites 4, 5, and 6, therefore, exhibit fewer dominant features.

Table 4. K-Means Clusters Results

	Urban Fabric	Openness	K-Means Cluster	Site Data Distribution
Site1				**Cluster2 and Cluster4** Many data have high values at Site 1: High road noise, low railway noise, good health, high human, music and transport sound, high joy and love emotion, high choice, high noise, more sad emotion.
Site2				**Cluster2 and Cluster4** Many data have high values at Site 2: High road noise, low railway noise, good health, high human, music and transport sound, high joy and love emotion, high choice, high noise, more sad emotion.
Site3				**Cluster1 and Cluster4** High noise, more sad emotions.
Site4				**Cluster3** No particularly prominent values.
Site5				**Cluster0 and Cluster3** No particularly prominent values. Most low value in Cluster0.
Site6				**Cluster0 and Cluster3** No particularly prominent values. Most low value in Cluster0.

3. Discussion

The richness and composition of an urban soundscape encode the spatial vibrancy of the city, the cultural diversity of the city, and other multidimensional aspects (Miranda, Halpern, and Duarte, 2020). Therefore, one of the most demanding tasks for urban planners and designers is to incorporate digital technologies into soundscape design. While it has been previously explored by the Tao G. Vrhovec Sambolec's Reality Soundtrack, to able to extract the sound from its context, i.e. to decontextualize it, remains an open challenge. Moreover, it is also unclear how to build a connection between individual listeners and larger groups (Cláudia Martinho and Labelle, 2011). Along the lines of soundscape digitalization, virtual reality and binaural techniques show promising leads when it comes to transferring acoustic contents between different environments, i.e. sound recontextualization. Through the immersive experience of combining various emotions and sensations collectively in a virtual environment, the investigation of the spatial and acoustic linkage becomes more convenient (Tveito, n.d.). On this note, previous experiments (Chua et al., 2016) demonstrate that through utilizing digital platforms, e.g. social media, data extraction for urban sound can be performed in an effective and inexpensive fashion. The effectiveness originated from the fact that noise data can be classified geographically and fits naturally to the standard noise exposure levels for various road types. The method is inexpensive because it has no requirement for creating additional services or infrastructure (Aiello et al., 2016). Thanks to the potential of synthesizing and modelling soundscapes enabled by digital technology, urban practitioners can better comprehend how soundscapes have an impact on us and how to leverage them as an essential toolset for public space design.

4. Conclusion And Future Work

According to the result of this research, regions with diverse soundscapes also entailed with rich emotions with a high proportion of positive feelings (e.g., joy and love), whereas areas with high openness and traffic noise levels indicate more apprehensive feelings (e.g., fear and anger). This aligns with our prior belief that "good sound" should be associated with happy feelings while "poor noise" should be associated with negative emotions. Therefore, when developing with urban soundscape, designers should prioritize lowering traffic noise and boosting the attractiveness of urban public space by enhancing the richness of the soundscape of the place.

By studying soundscapes in a data-driven manner, we intend to examine and consider the potential intersection of data science and urban planning through this research. We demonstrate that data analysis techniques and machine learning tools could be used to examine correlation between urban settings and people's perceptions of the place. Meanwhile, exploring urban soundscapes using crowdsourcing data presents a successful and useful method for integrating digital technologies with conventional soundscape research, which is much less computationally expensive than the conventional approach of soundscape recording.

Our research also bears a few limitations which we leave for future works. The data we can obtain simply represents the average value, thus, the actual soundscape variations of the venue have not yet been taken into consideration, as we discovered in the field trips that followed. We will further utilize machine learning to categorize various sound types based on prior studies. With a more abundant dataset, our next step is to thoroughly investigate how to use the development of numerous acoustic simulations to assess people's emotional shifts and forecast people's decisions.

Acknowledgements

This research was part of the studio “Machine Thinking Urbanism”, instructed by Roberto Bottazzi (Programme Director) and Dr. Tasos Varoudis (Associate Professor), at the University College London. We would like to acknowledge Fan Liu for her contributions to this project, and Yue Li for proofreading.

References

Abdi, H., & Williams, L. J. (2010). Principal component analysis. Wiley Interdisciplinary Reviews: Computational Statistics, 2(4), 433–459. https://doi.org/10.1002/wics.101

Aiello, L. M., Schifanella, R., Quercia, D., & Aletta, F. (2016). Chatty maps: constructing sound maps of urban areas from social media data. Royal Society Open Science, 3(3), 150690. https://doi.org/10.1098/rsos.150690

Bergan, B. (2020). Mass Quarantine From COVID-19 Mutes Urban Noise, but Earth Roars On. Interesting Engineering. https://interestingengineering.com/mass-quarantine-from-covid-19-mutes-urban-noise-but-earth-roars-on

Chang, M.C., Bus, P., & Schmitt, G. (2017). Feature Extraction and K-means Clustering Approach to Explore Important Features of Urban Identity. 16th IEEE International Conference on Machine Learning and Applications (ICMLA), Cancun, Mexico, 2017. https://doi.org/10.1109/ICMLA.2017.00015.

Martinho, C. & Labelle, B. (2011). Site of sound : of architecture and the ear. 2. Errant Bodies Press ; New York. ISBN: 978-0-9827439-0-4

Hillier, B., Leaman, A., Stansall, P., & Bedford, M. (1976). Space syntax. Environment and Planning B: Planning and Design, 3(2), 147–185. https://doi.org/10.1068/b030147

Hong, J. Y., & Jeon, J. Y. (2015). Influence of urban contexts on soundscape perceptions: A structural equation modelling approach. Landscape and Urban Planning, 141, 78–87. https://doi.org/10.1016/j.landurbplan.2015.05.004

International Organization for Standardization. (2014). ISO 12913-1:2014 Acoustics — Soundscape — Part 1: Definition and conceptual framework. (ISO Standard No. 12913-1:2014) https://www.iso.org/standard/52161.html

Jiang, L., & Nellthorp, J. (2020). Valuing transport noise impacts in public urban spaces in the UK: Gaps, opportunities and challenges. Applied Acoustics, 166, 107376. https://doi.org/10.1016/j.apacoust.2020.107376

Ma, K. W., Mak, C. M., & Wong, H. M. (2021). Effects of environmental sound quality on soundscape preference in a public urban space. Applied Acoustics, 171. ISSN: 0003-682X. Volume 171, 107570. https://doi.org/10.1016/j.apacoust.2020.107570

Menke, K., Smith, R., Pirelli, L., & John Van Hoesen. (2016). Mastering QGIS : go beyond the basics and unleash the full power of QGIS with practical, step-by-step examples. Packt Publishing. ISBN 978-1-78439-868-2.

Miranda, A. S., Halpern, D., & Duarte, F. (2020, July 17). Sonic Cities: Listening to Parks During Lockdown. Senseable City Lab. https://medium.com/senseable-city-lab/sonic-cities-listening-to-parks-during-lockdown-f9027b7deb46

Nieuwenhuijsen, M. J., & Haneen Khreis. (2018). Integrating human health into urban and transport planning : a framework. Springer. https://doi.org/10.1007/978-3-319-74983-9

Porteous, J. D., & Mastin, J. F. (1985). SOUNDSCAPE. Journal of Architectural and Planning Research, 2(3), 169–186. http://www.jstor.org/stable/43028767

Ringnér, M. (2008). What is principal component analysis? Nature Biotechnology, 26(3), 303–304. https://doi.org/10.1038/nbt0308-303

Sinaga, K. P., & Yang, M.-S. (2020). Unsupervised K-Means Clustering Algorithm. IEEE Access, 8, 80716–80727. https://doi.org/10.1109/access.2020.2988796

Tveito, H. (2016). The architecture of the ear and the design of soundscapes. | Interactive Architecture Lab. http://www.interactivearchitecture.org/the-architecture-of-the-ear-and-the-design-of-soundscapes.html

Van Nes, A. (2021). Spatial Configurations and Walkability Potentials. Measuring Urban Compactness with Space Syntax. Sustainability, 13(11), 5785. https://doi.org/10.3390/su13115785

World Health Organization. Regional Office for Europe. (2011). Burden of disease from environmental noise: quantification of healthy life years lost in Europe. Who.int. https://doi.org/9789289002295

SPATIAL ANALYTICS OF HOUSING PRICES WITH USER-GENERATED POI DATA, A CASE STUDY IN SHENZHEN

MUXIN JIA[1] and TARO NARAHARA[2]
[1,2]New Jersey Institute of Technology.
[1]mj8@njit.edu, ORCID 0000-0002-2213-6027
[2]narahara@njit.edu, ORCID 0000-0002-2200-1253

Abstract. Housing is among the most pressing issues in China. Researchers are eager to identify housing property's internal and geographic factors influencing residential property prices. However, few studies have examined the relationship between social media users' point of interest (POI) data and house prices using big data. This paper presents a machine learning model for regression analysis to reveal the relationship between housing prices and check-in POI density in Futian District, Shenzhen. The results show that our proposed price prediction model using additional features based on POI data proved to provide higher prediction accuracy. Our results indicate that incorporating POI features based on current feeds from location-based social networks can provide more up-to-date estimates of housing market price trends.

Keywords. Check-in POI, Kernel Density Estimation, Hedonic Pricing Method, SVR Model.

1. Introduction

Housing prices have always been a hot research topic in China. Researchers are eager to learn what factors influence housing prices, particularly in metropolitan areas such as Shenzhen, where property prices are extremely high. Properties near shopping malls, green spaces, and other amenities typically have higher marginal prices, according to location and land use theory based on analysis of urbanization patterns (Alonso, 1964). In many metropolitan cities, unusual concentrations of urban amenities can significantly impact the land value of the surrounding areas. Residential properties near popular shopping centers, for example, tend to have higher property values due to increased demand for rental apartments from densely populated residential areas and the positive externalities of commercial land in China. Previous research has shown that house prices are influenced by a variety of factors, including both internal and external factors. However, few studies used big data to investigate the underlying relationship between the spatial pattern of various POIs and house prices.

Today, many people use social media to share information about places they have visited. Rich resources for urban research are made available by the large amounts of geo-tagged data that can be collected and displayed on location-based social networks (LBSNs) as points of interest (POIs). Information pertinent to user contexts, such as

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 635-644.

current location, time, and check-in, is included in POI-based datasets. There are some existing studies that estimate the spatial correlation between POIs and nearby settlements and show how the locations and numbers of POIs affect commuters' travel distance and time in urban areas (Liu et al., 2014). However, only a few studies have examined how user-generated POI density distribution affects housing costs on a large urban scale. In this paper, we propose a hedonic framework for assessing how the spatial distribution of user-generated POI data affects real estate prices, which can contribute to urban research on housing prices. This paper presents the results of our case study in the Futian district in China, which quantitatively shows the correlation between user-generated POI data and housing prices.

This paper is organized into three sections. In Section 2, we use kernel density estimation to analyze the spatial pattern of check-in data based on various service types of POIs and visualize density with geographic heat maps. In Section 3, the kernel density estimators are obtained in ArcGIS using raster-to-point features, which are treated as new independent variables for regression analysis. In Section 4, support vector regression (SVR) models are employed to test whether the new indicators can explain housing prices significantly.

2. Dataset Creation

Shenzhen is south of Guangdong province, China, bordered by Hongkong. With a total permanent population of 17.56 million as of 2020, it is one of the youngest cities in mainland China but the third-largest metropolis by GDP. With the economy's ongoing development, the rapid rise in housing prices in Shenzhen has become a growing concern. Futian is Shenzhen's central business district and one of the busiest inner cities with the most expensive housing. As a result, the populous area's exceptionally high population density and massive geospatial data provide an ideal setting for studying the impact of POI activity measured by check-in datasets scraped from social media platforms such as Sina Weibo (2022) on housing prices. The study area covers the Futian District's entire boundary and comprises built-up areas, mountains, and beaches.

2.1. GEOGRAPHICAL FEATURE DATASET

The geographical elements are the primary datasets in this study, serving as the foundation for all subsequent analyses. Through the Shenzhen platform for common geospatial information services, we downloaded the most recent GIS shapefiles of geographical elements such as administrative boundaries, land use, water bodies, green space, infrastructure, road networks, bus stops, and metro stations in Futian District (Map World, Gov. 2022). In addition, we collected population distribution data from one of Shenzhen's largest telecommunications operators, which can be displayed as point elements in ArcGIS. Before data processing on ArcGIS, all datasets were categorized and saved based on element attributes (i.e., points, lines, and polygons).

2.2. CHECK-IN POI DATASET

Location check-ins on LBSNs generated by mobile hardware with built-in GPS, such as smartphones or tablets, are one of the most common types of crowdsourced geospatial data used to record information about a user's status update at a given time.

A large number of check-in data points are typically clustered in densely populated areas, which can be used for POI recommendation by analyzing the spatial correlation between contextual information and positions posted by social media users.

In this paper, we first crawled the check-ins in Shenzhen's Futian district from April 1 to July 1, 2022, using the API of the official platform of Sina Weibo (2022), China's most popular blog-based social media application. To protect the privacy of Weibo users, we did not collect information about individuals' identities. Second, to reduce the raw dataset's redundancy, incorrect and duplicate information about users' IDs and check-in locations at the same time was removed using the text discriminant. The processed dataset was then restricted to the Futian administrative boundary and imported into ArcGIS for projection. Finally, we collected 185,162 check-in records, including valid ID fields, username, position, timestamp, and device type. Table 1 shows the format of the check-in dataset.

By analyzing the three-month records of check-ins by users on Sina Weibo, we can calculate the number of public visits to a location, which usually appears on the map as geo-tagged spots (POIs). We identified eight service types for user-generated POIs through data crawling: *retail facilities, restaurants, transportation facilities, healthcare facilities, recreational facilities, hotels, financial and office buildings, and scenic locations* (Figure 1). Table 2 shows the number of user check-ins for each type of POI. The most widely distributed retail facilities, such as "convenience stores" and "supermarkets," have the highest number of check-ins in the Futian district. While *scenic places* have the fewest POI counts, it has a relatively high number of check-ins because it is a freely accessible public location that draws many citizens and tourists.

2.3. HOUSING PRICE DATASET

To obtain reliable and up-to-date housing price information, we crawled the second-hand residential transaction data relating to 726 residential buildings with a total of 4488 properties from April to July 2022 via the API of the LianJia website (2022), a residential real-estate online platform with the largest database of various properties in China. This website contains a wide range of information about properties for sale, such as the type of house, transaction price, size, age, floor, number of bedrooms, etc. It is worth noting that the property types in this study are limited to ordinary commercial properties, with the exception of duplexes and villas, which have a low number of transactions. Furthermore, because only three-month transaction data were used for this study, it is reasonable to disregard the potential impact of time on prices.

3. Features

This section explains our methods to generate features based on the hedonic pricing framework, which were then used as inputs for price prediction models using SVR in Section 4. First, the spatial distribution of each type of check-in spot (POI) was examined using kernel density estimation (KDE). The average value of the kernel density estimator within each cell was also measured as explanatory variables in the subsequent regression models. Then, we categorized features into internal, GIS-data-based external, and POIs-based external factors based on characteristics. Additionally, 3.5 provides a stepwise regression analysis of the statistical significance of variables.

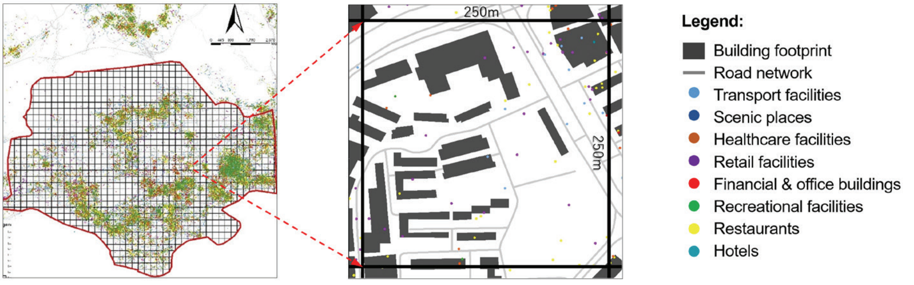


Figure 1. Spatial distribution of eight categories of user-generated POIs.

Id	Name	Time	Date	Wgs-lon	Wgs-lat	Device_type
47498	zz_gigga	8:28:15	2022-05-06	114.052558	22.530727	OPPO Reno 6
47499	big shark	12:45:03	2022-05-06	114.032615	22.614266	iPhone 13
47500	Wang_w	15:12:09	2022-05-06	114.052575	22.530611	Huawei Mate 50E

Table 1. Format of check-in dataset

TYPE	ABBREVIATION	POI_COUNTS	PERCENTAGE	CHECK-IN NUMBERS	PERCENTAGE
Retail facilities	REF	28188	45.70%	64825	46.54%
Restaurants	RES	18828	30.53%	47068	33.79%
Transport facilities	TRA	4820	7.82%	7721	5.54%
Healthcare facilities	HEF	3635	5.89%	852	0.61%
Recreational facilities	RCF	2474	4.01%	5944	4.27%
Hotels	HOT	2112	3.42%	6333	4.55%
Financial & office buildings	FOB	1141	1.85%	2166	1.56%
Scenic places	SEP	477	0.77%	4370	3.14%
TOTAL		61675	100.00%	139279	100.00%

Table 2. Counts of user-generated POIs with check-ins for eight service types

3.1. FEATURES BASED ON A GRID

To simplify the entire analytical process based on different types of vector datasets, we converted the study area into a feature class containing a net of 427 rectangular 250m cells as a base map for spatial analysis over Futian District using the Fishnet tool in ArcGIS. We determined the size of each grid cell to be 250m x 250m after several tests, which is consistent with the average scale of a block in Shenzhen, as grid sizes larger than 250m would mask local differences. On the other hand, a much smaller size would result in an unnecessarily strong bias influenced by local features. Section 3 describes the list of grid features, including internal and external factors.

Figure 2. (a) Spatial position of POI; (b) Heatmap visualization of POI; (c) Convert raster dataset to point features; (d) Zoom-in map of points with kernel density estimator for one grid cell.

3.2. SPATIAL DISTRIBUTION ANALYSIS

In this subsection, KDE is used to analyze the spatial distribution of POIs with check-in numbers, as it can visualize the distribution of discrete measurements and generate a smooth density surface centered on a point within a continuous geographic space. The formula for calculating the predicted density at the new position is determined by:

$$\hat{f}(s) = \frac{1}{(h)^2}\sum_{i=1}^{n}[\frac{3}{\pi}\cdot K_i(1-\left(\frac{d_{is}}{h}\right)^2)^2]$$

Where $\hat{f}(s)$ is the predicted kernel density at location s, $i = 1, \dots, n$ are the input points within a radius distance of location s, K_i is the population field value of point i representing the spatial weight, d_{is} is the distance between point i and location s, and h is the bandwidth of KDE, also known as the search radius in ArcGIS. The predicted kernel density values depend on the search radius, spatial weight, and grid sizes.

Applying the POI density value to a specified grid map involves three steps (Figure 2). We determined the output parameters after several tests prior to performing kernel density analysis. The output raster cell size is set to 50m x 50m to improve estimation precision, and the search radius is set to 250m, taking into account the maximum distance to maintain walking comfort. It is worth noting that the population field's value is the number of check-ins (i.e., spatial weight) for each POI, as a higher number of check-ins indicates that this location has a higher level of attention and a larger underlying influence on the neighborhood. The exported raster heatmap was then converted into an evenly distributed spot matrix with kernel density estimates. Finally, we associated attributes of each point to a specified grid where they belong to and computed the average kernel density estimator for each grid cell.

3.3. FEATURES BASED ON A HEDONIC PRICING FRAMEWORK

We proposed a hedonic pricing framework to investigate the housing attributes and external environmental factors that influence the price of target houses and the contributory value of each factor. Since the home value is not only affected by its inherent characteristics such as floor area, age, and stories but also by external environmental factors that can be perceived by consumers, real estate economists often employ the hedonic pricing method to estimate the selling price of a property because

	Variable	Variable definition	Mean	Std. Dev.	Min	Max
Response	Ln_price	semi-logarithmic form of the total price	15.63	.551	13.93	17.27
INT: Internal factors	H_area	Area of residential living spaces (m^2)	93.654	43.35	26.29	387
	H_age	Age of the unit (years)	24.031	49.36	3	1032
	Bedr_n	Number of bedrooms in the unit	2.64	.90	1	9
	Floor	Floor on which the unit is situated	17.65	10.16	1	35
	B_den	Building density	.244	.104	0	.659
	FAR	Floor area ratio	2.475	1.202	0	7.384
EXT: External factors based on GIS data	R_den	Road density for each grid (m/m^2)	.044	.022	0	.17
	RWater	Ratio of water body area for each grid	.005	.024	0	.261
	RGreen	Ratio of green space area for each grid	.017	.072	0	.751
	P_den	Population density for each grid (p/m^2)	3.71	1.23	0	5
POI: External factors based on POIs	FOB	KDE of financial & office buildings for each grid	3672.6	459.50	0	31561.63
	HEF	KDE of healthcare facilities for each grid	6974.88	408.32	0	19945.13
	HOT	KDE of hotels for each grid	3329.29	300.95	0	16714.35
	RCF	KDE of recreational facilities for each grid	4942.70	347.18	167.56	24435.08
	RES	KDE of restaurants for each grid	3096.11	280.35	0	14601.83
	REF	KDE of retail facilities for each grid	4745.45	754.41	103.26	75495.82
	SEP	KDE of scenic places for each grid	575.54.67	611.12	29.97	5335.45
	TRA	KDE of transport facilities for each grid	10727.69	516.701	959.28	26072.12

Table 3. Selected variables (features) for regression analysis in Section 4.

of the different marginal contribution of environmental factors caused by spatial heterogeneity. POI with high-frequency access records can be considered as an important external environmental factor influencing the housing price of private property. The regression equation is shown as follows:

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + \cdots + b_kx_k + \varepsilon$$

Where $\hat{y}$ is the predicted value of the house price, which is ln_price in this study. $x_1, x_2 \ldots x_k$ are k distinct independent variables. b_0 is the constant, and b_1 through b_p are the regression coefficients. ε is the model error.

We pre-processed the dataset before regression analysis to improve data quality, including removing outliers and replacing missing values. Also, we selected the semi-logarithmic form, i.e., $ln(price)$, as the dependent variable, which has been shown to have better explanatory power than the total price and unit price per square meter in previous studies (Soler and Gemar, 2018). After data cleaning and transformation, we performed descriptive statistics, and the results are shown in Table 3.

3.4. RESULT OF SPATIAL DISTRIBUTION ANALYSIS

The spatial distribution patterns for different service types of check-in POIs are analyzed using KDE. In figure 3, the blue-shaded area indicates higher kernel density

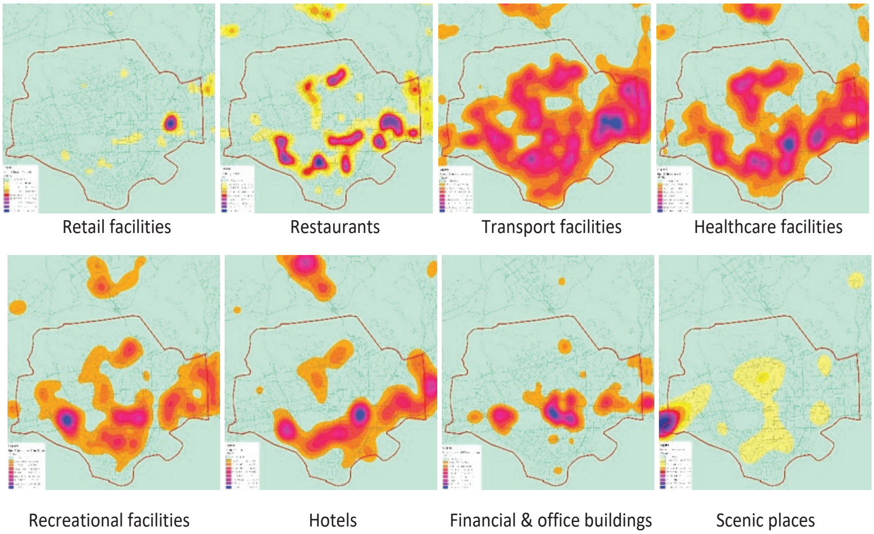


Figure 3. Heatmaps of eight types of check-in POIs

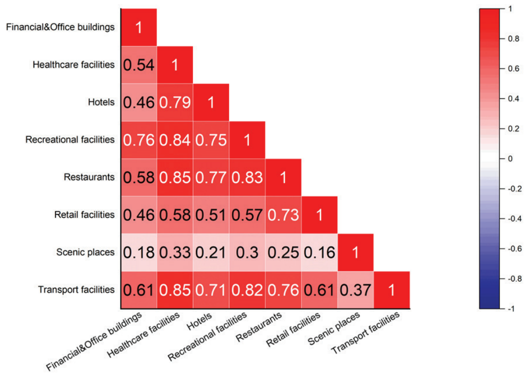


Figure 4. Correlation heatmap of average kernel density estimator for each type of POI

with a higher frequency of visits, whereas the yellow-shaded area indicates lower kernel density with relatively lower visitors. Additionally, from the visual observations of the results, the POIs of transportation, medical, financial, and recreational facilities are more evenly distributed than the others in the Futian District, which implies that such public facilities' layouts have been planned according to the principle of spatial equity in China. Through the correlation analysis of the kernel density estimators for various check-in POIs (figure 4), we can find that except for *scenic spots*, most service facility types have good positive correlations among each other because various service facilities tend to be concentrated in overpopulated areas, while scenic places without a high number of check-ins are often far away from the city center, which is also consistent with our general perceptions and understanding of cities.

Table 4. Analysis results of statistically significant variables in 3.5.

ln_price	Coefficient	Std. Error	t-value	p-value	[95% Confidence Interval]	
H_area	7.66E-03	6.72E-04	11.33	8.25E-04***	6.34E-03	8.98E-03
H_age	3.00E-03	0.02	0.16	1.29E-04***	-1.20E-03	1.53E-03
Bedr_n	0.15	0.03	4.47	7.57E-04***	0.08	0.21
B_den	-1.13	0.19	-5.84	3.27E-04***	-1.51	-0.75
FAR	0.08	0.02	4.9	9.66E-04***	0.05	0.12
RGreen	0.04	0.02	2.31	0.02**	5.94E-03	0.07
FOB	6.39E-03	2.37E-03	2.7	7.00E-03***	1.74E-03	0.01
RCF	-2.50E-03	7.40E-04	-3.38	1.20E-03***	-3.96E-03	-1.05E-03
SEP	2.19E-03	9.07E-04	2.41	0.02**	4.05E-04	3.97E-03

3.5. ANALYSIS OF STATISTICALLY SIGNIFICANT VARIABLES

In this subsection, we analyzed the statistically significant explanatory variables by removing those that are less important and overly correlated using backward stepwise linear regression with a series of iterative T-tests. The results in Table 4 show that the internal characteristics of the house are generally more statistically significant than the external environmental factors, indicating that, in Shenzhen, consumers are likely to invest more in the internal attributes of their houses, such as the number of bedrooms, rather than environmental factors related to their neighborhoods. We can also note that the house's age and the building's density have a negative impact on property prices, implying that consumers prefer newer houses with lower density. In terms of external factors, the model demonstrated that the densities of check-in POIs of financial & office buildings and scenic places have significantly positive impacts on housing prices, whereas the densities of recreational areas and hotels have negative impacts. It can be inferred that the concentration of financial institutions and scenic spots promotes the increase in housing prices in surrounding areas. Consumers are generally willing to pay for the positive spatial externalities generated by such facilities.

4. Price Prediction Models using SVR

This paper aims to investigate the contribution and influence of user-generated POI-based data as indicators (i.e., input variables) for housing price prediction. We present price prediction models using machine learning, which have been widely used in prediction models, for examples, in indoor (Narahara et al., 2022) and urban conditions (Xu et al., 2022). This paper focuses on prediction using support vector regression (SVR), a learning machine that estimates a function given a dataset. The support vector machine (SVM), proposed by Vapnik (1999), was originally used for classification but has since been extended to the task of regression and prediction. We used the dataset from Sections 2 and 3, which consists of 427 grids of 250m cells within the Futian district, with features calculated within each cell. We divided the dataset into randomly selected 90% grids for training and 10% for testing. We prepared the following four models using the Gaussian kernel that input combinations of three groups of features listed in Table 3: internal factors, external factors based on GIS data, and external factors based on POIs.

Table 5. Prediction accuracy comparison.

Model	Metric	INT	INT+EXT	INT+**POI**	INT+EXT+**POI**
SVR	R^2	0.658	0.536	**0.773**	0.751
	RMSE	0.306	0.357	**0.250**	0.261
MLR	R^2	0.541	0.556	0.633	**0.638**
	RMSE	0.355	0.349	0.317	**0.315**

- INT: The model with input features based on internal factors of residential properties in each 250m grid, including average per unit of areas of residential living spaces, number of bedrooms, building age, and building density (4 dimensions).
- INT+EXT: In addition to the features in the INT model, this model includes external factors based on GIS data, a ratio of road area density, water area, green area, population density, and FAR within each grid as features (4+5=9 dimensions).
- INT+POI: In addition to the features in the INT model, this model includes user-generated POIs calculated based on KDEs of eight service types of buildings, such as retail facilities (REF), in each grid as features (4+8=12-d).
- INT+EXT+POI: The model with 17-dimensional all input features (4+5+8=17-d) based on internal factors of residential properties, external factors based on GIS data, and user-generated POIs calculated based on the KDE of eight service types.

All feature vectors were standardized before using them for machine learning such that their distributions had a mean value of 0 and a standard deviation of 1. The models were trained to predict the total price in semi-logarithmic form within each 250m grid area (Ln_price in Table 3).

4.1. RESULTS OF SVR MODELS

We applied the coefficient of determination (R2) to measure the correlation between the predicted values and the housing prices in the test data as ground truths. We also evaluated the performance using the root-mean-squared error (RMSE). Note that a higher PCC and a lower RMSE mean better prediction performance. In Table 5, our proposed models using additional features based on the POI data outperformed models without the POI features, such as INT and INT+EXT,. The INT+POI model has the best prediction accuracy in terms of both R2 (0.773) and RMSE (0.250), and R2 values greater than 0.7 are generally considered to have a high correlation level. As a result, quantitative data based on social media network users' check-in POIs are considered effective features for housing prices in Futian District. The forecasting performance of SVR models was also compared to that of multiple linear regression (MLR) models under the same conditions using the four feature combinations. The experimental results show that the SVR model outperforms the MLR models and that an SVR-based approach is an efficient tool for forecasting real estate prices. According to the results, incorporating POI features from real-time feeds by LBSNs for certain durations has the potential to produce more accurate forecasts of price movements in the housing market.

4.2. LIMITATIONS

Our results show that adding features based on external factors from GIS data, such as road density, did not improve the prediction accuracy of the proposed models as much as expected. The quantitative format of our data does not fully describe the local geometrical conditions of each 250m grid. Two grids, identical in environmental elements, such as roads, buildings, and green areas, may have different living qualities for humans based on how they are arranged spatially. Therefore, pixel-wise features such as images incorporating more geometrical conditions of urban spaces can be used for input variables of models, but this is left as future work.

5. Conclusion

We employed multisource open-access datasets to quantify the spatial pattern of user-generated POI based on eight service facility types and studied its underlying impact on housing prices. Our findings show that internal characteristics based on properties have a stronger correlation with house prices than external factors based on environmental conditions. Experimental results revealed that the spatial patterns of three types of POI data, including *hotels, scenic places,* and *financial facilities,* significantly impact the hedonic price of residential properties in Shenzhen's Futian District. Our proposed methods for adopting features based on POI data generated from social network users' check-ins proved effective, as we obtained housing price prediction models with relatively high accuracy (R2=0.773) for the size of the experiment's grid area. The POI-based datasets can fluctuate over time due to more dynamic feeds from social networks, and the addition of the POI-based features can contribute to the better estimation of more up-to-date housing market trends in prices that cannot be adequately predicted by the use of internal and external features alone, which are normally more static in time and place than the POI-based features.

References

Alonso, W. (1964). Location and Land Use: Toward a General Theory of Land Rent, Cambridge, MA and London, England: Harvard University Press, 1964.

Lianjia, (2022). Housing price dataset, Retrieved from https://sz.lianjia.com/.

Liu, Sui, Z., Kang, C., and Gao, Y. (2014). Uncovering patterns of inter-urban trip and spatial interaction from social media check-in data. PloS One, 9(1), e86026–e86026. https://doi.org/10.1371/journal.pone.0086026

Map World, Gov. (2022). Shenzhen platform for common geospatial information services, from https://guangdong.tianditu.gov.cn/shenzhen/.

Narahara, T., & Yamasaki, T. (2022). Subjective Functionality and Comfort Prediction for Apartment Floor Plans and Its Application to Intuitive Online Property Search. *IEEE Transactions on Multimedia.* https://doi.org/10.1109/TMM.2022.3214072

Sina Weibo Open access platform, (2022). Check-in POIs from http://open.weibo.com.

Soler, I. P., and Gemar, G. (2018). Hedonic price models with geographically weighted regression: An application to hospitality. Journal of Destination Marketing & Management, 9, 126-137. https://doi.org/10.1016/j.jdmm.2017.12.001

Vapnik, V. (1999). The nature of statistical learning theory. Springer.

Xu, X., Qiu, W., Li, W., Liu, X., Zhang, Z., Li, X., & Luo, D. (2022). Associations between street-view perceptions and housing prices. *Remote Sensing*, 14(4), 891. https://doi.org/10.3390/rs14040891

THERMAL PERFORMANCE EVALUATION OF LOW-INCOME HOUSING UNITS USING NUMERICAL SIMULATION

FAIZ AHMED CHUNDELI[1] and TANIA BERGER[2]
[1]School of Planning and Architecture Vijayawada, Andhra Pradesh India, [2]Center for Environmental Sensitivity, Danube University, Krems, Austria
[1]faizahmed.arch@gmail.com, 0000-0003-1556-2375
[2]tania.Berger@donau-uni.ac.at

Abstract. The thermal performance of buildings is measured as heat energy transfer between the buildings and the surrounding environment, and there are several heat exchange possibilities. This paper presents the thermal performance of 12 non-air-conditioned low-income single dwellings in warm-humid climates. The Building and material characteristics of the dwellings, including field measurements of the 12 cases, were meticulously documented through a primary survey. The critical indicator for assessing and evaluating the performance of the dwelling unit was hourly simulated indoor temperature data for an entire year. Further, potential planning and design components, viz. building orientation, roof and wall insulation, window size, property & locations, clerestory window, increased floor-to-ceiling height, site setback, and roof profile, were iterated to improve the thermal performance of low-income dwellings. Indoor temperatures as high as 45.9 C were recorded, the mean indoor temperature for the summer months (March-July) was over 34.64 C, and it was always higher than 30 C for the rest of the month. The findings show that the inhabitants are subjected to temperatures exceeding 34 degrees Celsius for more than half of the year. The paper concludes with some suggested design measures to improve the thermal performance of low-income houses. The study also emphasizes the importance of refined early design phase assessment and decision-making to improve the indoor thermal environment.

Keywords. Thermal Performance, Low-income Housing, Building Simulation, Heatwaves, Natural Ventilation

1. Introduction

Climate change is frequently associated with rising temperatures and poor environmental quality in cities (McCarthy et al., 2010). Many studies have been conducted in cities around the world to understand better the phenomenon of heat islands (Banerjee & Chattopadhyay, 2020; Cao et al., 2019; Kataoka et al., 2009; Keikhosravi, 2019; Mathew et al., 2018; Qaid et al., 2016). Heatwaves and high

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 645-654. © 2023 and published by the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA), Hong Kong.

temperatures are associated with discomfort in the home and heat-related health issues (Fisk et al., 2020; Pantavou et al., 2011). Furthermore, the demand for energy to cool spaces rises (Biardeau et al., 2020; Khalid & Sunikka-Blank, 2018; Latulippe & Klenk, 2020; Osunmuyiwa et al., 2020). Studies show a strong correlation between the inhabitants' income and social characteristics and the built fabric's quality, thermal conditions, and energy consumption (Gupta et al., 2020; Healy & Clinch, 2002). According to studies, higher energy consumption patterns have been overserved in low-income housing settlements with insufficient and poor thermal protection (Berger & Höltl, 2019; Synnefa et al., 2017). People living in low-income housing areas have higher health-related mortality and lower thermal comfort due to prolonged heat exposure and a poor thermal environment (Haines et al., 2006). Through various passive cooling techniques, research has been conducted to improve the indoor thermal conditions and overall thermal performance of low-income dwellings (Berger et al., 2022; Malik & Bardhan, 2020; Synnefa et al., 2017).

Cool materials can reduce outdoor and indoor temperatures by several degrees (Alam et al., 2017; Bhamare et al., 2019; Garshasbi et al., 2020; Hwang, 2006; Kaboré et al., 2018; Kong et al., 2014). Mitigation strategies such as using advanced cooling materials on buildings and their surroundings have yielded positive results (Garshasbi et al., 2020; M Santamouris et al., 2007). Other cooling strategies, such as heat sinks, night flushing, vegetation on buildings, and the use of combined natural and hybrid ventilation systems, have been investigated (Mhuireach et al., 2020; Pantavou et al., 2011; Mattheos Santamouris & Kolokotsa, 2013; Zhai et al., 2011). Potential planning and design components include building orientation, roof and wall insulation, window size, property and location, clerestory window, increased floor-to-ceiling height, site setback, and roof profile to improve the environmental conditions and thermal performance of low-income dwellings are investigated in this chapter. The thermal performance of a typical non-air-conditioned low-income single dwelling in Vijayawada, India, is presented. Furthermore, the study identifies potential strategies for improving thermal comfort and lowering indoor temperature applicable to government-provided low-income housing schemes.

2. Location and dwelling characteristics

2.1. CASE STUDY LOCATION

Vijayawada, Andhra Pradesh, India, is the location of the case study. Vijayawada is one of the most populous cities in Andhra Pradesh, located in southern India. It is situated at latitude 16o31' north and longitude 80o37' east. According to the 2011 census, the city has a population of 10,48,000. According to the India Meteorological Department (IMD), the highest temperature in Vijayawada was 44 degrees Celsius in May 2019 and 46 degrees Celsius in May 2020. According to weather data, the annual average ambient temperature was 27.92 degrees Celsius. During the summer of 2019 (March-July), the average ambient temperature was always higher than 31 ◦C. During the May 2019 heatwave, the maximum ambient temperature reached values close to 46 ◦C for four consecutive days.

2.2. CASE STUDY LOCATION

The climate in Vijayawada is generally hot and humid, with heat waves in the summer. As a result, buildings significantly impact how people cope with high indoor temperatures (Berger et al., 2022). Furthermore, heat islands are affected by heat islands because the settlements are close to Vijayawada's downtown. The physical characteristics of dwellings were documented to accurately map building envelope characteristics, insulation, the thermal mass of the building, shading elements used, and vegetation around the dwelling unit. The research was conducted from June to September 2019.

Primarily, the dwelling is rectangular in plan, with a 3.5m width, 8.5m depth, and 3m height, as shown in Table 1. Walls are made of conventional 200mm bricks and 100mm concrete roofs with a calculated overall R-value of 1.398 °F ft2 hr./BTU. The representative dwelling unit is without any additional wall or roof insulation, Figure 1. As a result, indoor temperatures differ only slightly from those outside. Dwelling ventilation was accomplished naturally through windows and doors. However, windows were intentionally blocked in several instances to prevent smoke from entering the kitchen. As a result, windows remained closed, and doors remained open until late evenings for ventilation. Because the local government planned and developed the layout, all dwelling units had a minimum of 0.6m setback on one side of the plot, per prevailing site planning regulations. However, the roofs of the dwelling units have been extended to the plot boundary, obstructing natural ventilation. Vegetation is scarce throughout the settlement.

Further, the e-QUEST 3.65 building energy simulation and performance assessment tool is extensively used for evaluating the indoor thermal performance of the identified dwellings. Further, during documentation for random validation, the Testo-480 and thermal imager camera were used for measuring temperature, humidity, air velocity, and iso-thermal images. Hourly indoor temperature was simulated after entering the appropriate building and contextual characteristics of the dwellings. The design ventilation rate of 15 CFM/Person, the light power density of 0.30, and the plug load density of 0.60 (for all the other miscellaneous equipment loads) were assumed for calculating the overall energy consumption of the dwellings, as shown in Table 1.

Table 1. Representative dwelling building characteristics

Case No.	Building Dimension (feet)	Building Area (sq. ft)	No. of Floors	No. of Occupants	% of the window opening	Building Orientation	Total Annual Energy Consumption (kWh/m2)
1	12 x 28	672	2	4	3.3	West	24

3. Results of the study

3.1. THERMAL PERFORMANCE ASSESSMENT OF REPRESENTATIVE DWELLING UNIT

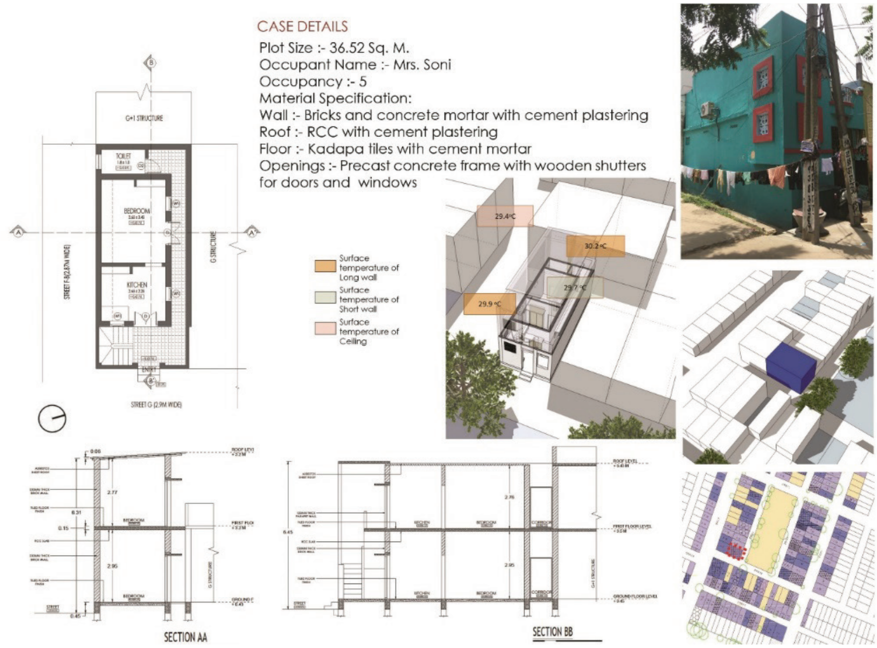


Figure 1 Representative dwelling building characteristics and contexts

Thermal performance assessments were performed on representative dwelling units, and the variation of hourly indoor temperature and dry-bulb temperature is plotted in Figure 2. According to the simulation results, the highest indoor temperatures were recorded in May, followed by April and June. During May 2019, the mean outdoor temperature was consistently above 29.7 degrees Celsius, while the mean indoor temperature was 32.7 degrees Celsius. The highest outdoor and indoor temperatures recorded from weather and simulated data were 44 C and 45.9 C, respectively. Residents are subjected to high temperatures exceeding 33 degrees Celsius for more than 2150 hours, or nearly 90 days, between March and July. High temperatures cause serious health problems and create an unsuitable environment for comfortable living. Understanding the effects of prolonged exposure to high indoor temperatures provides essential insight into the thermal performance of dwelling units. During the summer of 2019, building characteristics were documented, and qualitative interviews were conducted between 10 a.m. and 4 p.m. Furthermore, for over 1600 hours, indoor temperatures of 35 C or higher were recorded. The average indoor temperature was 30.8 C based on temperature recordings made with a Testo-480 and a thermal imager camera during the documentation period.

According to the above findings, the average indoor temperature in the representative low-income dwelling unit was 5 degrees Celsius above the ambient temperature. Furthermore, the average maximum indoor temperature can reach 47.8 C, 12 C higher than the maximum ambient temperature of 36.3 C. Aside from a higher number of people living in each of the dwellings, several reasons for the increase in indoor temperature can be identified, including poor building envelope design, a lack of insulation, an unsuitable thermal mass of the building, and a lack of shading elements and vegetation around.

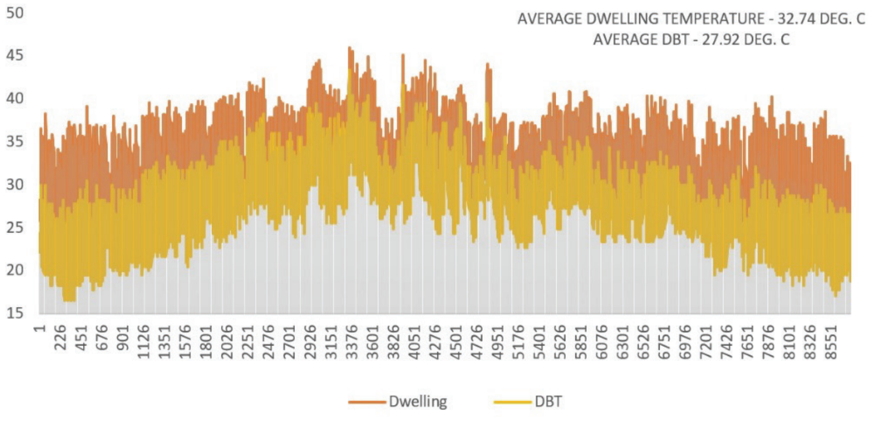


Figure 2 Hourly simulated indoor room temperature of the case dwelling and dry-bulb temperature

3.2. RECOMMENDATIONS FOR IMPROVEMENTS IN DWELLING UNITS

Most residents are protected from heat stress by buildings, and buildings are the most important and effective means of protection. The effectiveness of this protection is limited by design constraints and the use of low-cost building materials, particularly during the evening and night. Unfortunately, changing or improving building design and materials after initial construction is complicated (Berger et al., 2022). As a result, the most prudent measures for reducing heat stress in these households can indirectly be addressed. Several iterations of possible building improvement strategies were simulated to obtain a feasible solution and identify effective building control parameters to improve the thermal performance of the dwellings. The strategies were based on the orientation of the dwellings, wall and roof insulation, the use of appropriate low-E glass material for windows, changing the position of windows based on cardinal direction, using clerestory windows, increased floor-ceiling height for increased ventilation, the use of a pitched roof instead of a flat roof, a combination of a pitched roof and clerestory window, and increased site area (increase in setback). Figure 3 depicts the summary results of the iterations.

Based on the iterations, the dwelling facing south and west performed poorly. A similar indoor average temperature of 33.5 ◦C and above were observed in both cases. At the same time, dwellings facing north and east showed an average indoor temperature of 32.5 ◦C and 32 ◦C, respectively. On average, about 1-1.5 ◦C differences are observed based on the building orientation, and the highest difference of 2 ◦C was observed for buildings facing east. It is not always possible to develop east-facing plots only. However, consideration may be given in the initial phase of neighbourhood planning for developing a climate response layout

One primary function of any building material is to (at least somewhat) disconnect indoor from outdoor temperature, be it to keep the interior warm in cold outside conditions or cool in hot weather. The efficacy in doing so obviously differs considerably for different materials based on their insulation capacity applied to brick and concrete structures. Undoubtedly, insulation would be one of the most effective ways of protecting residents from excess heat (Doctor-Pingel et al., 2019). However, no insulation material was found in any of the documented dwellings. Consequently, indoor temperatures only slightly differ from that outside, and, for a significant part of the months, average indoor temperatures were way above ambient temperature. This is evident from the simulated results obtained, as shown in Figure 2. Dwellings with white clay tile as roof insulation brought average indoor temperature by 1.5 ◦C, while combined wall and roof insulation brought about a 2-2.5 ◦C difference. Therefore, adequate wall and roof insulation or high-performing material for roof and wall should be carefully decided while developing mass housing schemes, mainly while designing low-income dwelling units, where there is less scope for modifications in design by inhabitants.

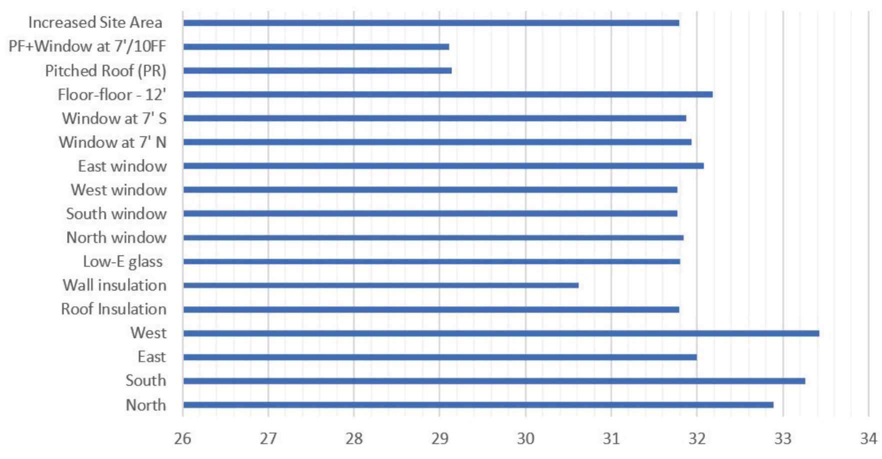


Figure 3 Strategies for improving the thermal performance of the dwellings

A significant reduction in average indoor temperature was observed when the flat roof of the dwelling units was changed to a pitched roof. The average indoor temperature was observed at 29.14 ◦C, which is 4 ◦C, lesser than the mean indoor temperature observed for all the cases. Further, a pitched roof coupled with a clerestory window (window opening at 7'0" level) showed a slightly improved result, i.e. average indoor temperature was observed at 29.11 ◦C. Retrofitting the roof of the dwelling units is not possible now. However, such a strategy may be adopted in mass housing schemes to achieve better indoor thermal conditions since most social housing schemes do not facilitate the beneficiaries to go for significant alternations or additions to their allotted houses. Therefore, early planning and design phase interventions are the only viable option.

4. Conclusions

Low-income housing schemes are characterized by densely populated dwellings and mass reproduction of houses with the same building design template and material characteristics without addressing the climate-related issues at the individual unit level. Therefore, achieving desired thermal environment and comfort conditions are often problematic. Further, continuous exposure to high temperatures affects the health and well-being of the inhabitants. This is one of the significant causes of concern for marginalized sections of society, as their dwellings experience high indoor temperatures for an extended period. This chapter investigates the thermal performance of a representative non-air-conditioned low-income single dwelling. Indoor temperatures as high as 50 ◦C were observed, and the mean indoor temperature for the summer months (March-July) was over 34 ◦C and consistently above 30 ◦C for the rest of the months. The inhabitants are exposed to high temperatures exceeding 33 ◦C for almost half a year. On average, for about 1600 hours or 70 days, the indoor temperature of 35 ◦C and above were recorded in many dwellings. From the study, East-facing dwellings and dwellings with roof insulation lowered the average indoor temperature by 1.5-2 ◦C and for a combination of wall and roof insulation by 2.5 ◦C. The highest reduction in indoor temperature, i.e. as far as 4 ◦C, was observed for dwellings with pitched roofs. Thermal performance analysis of existing dwelling units helps to understand the dynamics of indoor thermal environments. It is evident from the analysis that indoor thermal environments are greatly influenced by many building and context characteristics. Given the prolonged exposure to high temperatures in cities like Vijayawada and alike, knowledge and understanding of the thermal dynamics of the building are essential to developing appropriate and feasible retrofit measures and new designs. Notably, this understanding is vital while planning and designing mass social housing projects, where one informed decision in the early design phase could make a massive difference to the overall performance of the buildings in the long run and the well-being of inhabitants. The findings also emphasize the need for refined early design phase assessment and decision to improve the indoor thermal environment. The study findings could further help architects, planners, and decision-makers make informed design decisions while planning and designing low-income housing settlements.

References

Alam, M., Sanjayan, J., Zou, P. X. W., Ramakrishnan, S., & Wilson, J. (2017). Evaluating the passive and free cooling application methods of phase change materials in residential buildings: A comparative study. Energy and Buildings, 148, 238–256. https://doi.org/https://doi.org/10.1016/j.enbuild.2017.05.018

Ali, U., Shamsi, M. H., Hoare, C., Mangina, E., & O'Donnell, J. (2019). A data-driven approach for multi-scale building archetypes development. Energy and Buildings, 202, 109364. https://doi.org/https://doi.org/10.1016/j.enbuild.2019.109364

Banerjee, S., & Chattopadhyay, S. (2020). A meta-analytical review of outdoor thermal comfort research: Applications, gaps and a framework to assess low-income settlements in Indian megacities. Urban Climate, 33, 100641. https://doi.org/https://doi.org/10.1016/j.uclim.2020.100641

Berger, T., & Höltl, A. (2019). Thermal insulation of rental residential housing: Do energy poor households benefit? A case study in Krems, Austria. Energy Policy, 127, 341–349. https://doi.org/https://doi.org/10.1016/j.enpol.2018.12.018

Berger, T., Chundeli, F.A., Pandey, R., Jain, M., Tarafdar, A. and Ramamurthy, A. (2022). Low-income residents' strategies to cope with urban heat. Land Use Policy, 119, p.106192.

Bhamare, D. K., Rathod, M. K., & Banerjee, J. (2019). Passive cooling techniques for building and their applicability in different climatic zones—The state of art. Energy and Buildings, 198, 467–490. https://doi.org/https://doi.org/10.1016/j.enbuild.2019.06.023

Biardeau, L. T., Davis, L. W., Gertler, P., & Wolfram, C. (2020). Heat exposure and global air conditioning. Nature Sustainability, 3(1), 25–28. https://doi.org/10.1038/s41893-019-0441-9

Cao, Z., Wu, Z., Liu, L., Chen, Y., & Zou, Y. (2019). Assessing the relationship between anthropogenic heat release warming and building characteristics in Guangzhou: A sustainable development perspective. Science of The Total Environment, 695, 133759. https://doi.org/https://doi.org/10.1016/j.scitotenv.2019.133759

Census. (2011). District Census Handbook - Krishna, Andhra Pradesh. Directorate of Census Operations, Andhra Pradesh. Retrieved from http://censusindia.gov.in/2011census/dchb/2816_PART_A_DCHB_KRISHNA.pdf

Cerezo, C., Sokol, J., AlKhaled, S., Reinhart, C., Al-Mumin, A., & Hajiah, A. (2017). Comparison of four building archetype characterization methods in urban building energy modeling (UBEM): A residential case study in Kuwait City. Energy and Buildings, 154, 321–334. https://doi.org/https://doi.org/10.1016/j.enbuild.2017.08.029

Doctor-Pingel, M., Vardhan, V., Manu, S., Brager, G., & Rawal, R. (2019). A study of indoor thermal parameters for naturally ventilated occupied buildings in the warm-humid climate of southern India. Building and Environment, 151, 1–14. https://doi.org/https://doi.org/10.1016/j.buildenv.2019.01.026

Fisk, W. J., Singer, B. C., & Chan, W. R. (2020). Association of residential energy efficiency retrofits with indoor environmental quality, comfort, and health: A review of empirical data. Building and Environment, 107067. https://doi.org/https://doi.org/10.1016/j.buildenv.2020.107067

Garshasbi, S., Haddad, S., Paolini, R., Santamouris, M., Papangelis, G., Dandou, A., Methymaki, G., Portalakis, P., & Tombrou, M. (2020). Urban mitigation and building adaptation to minimize the future cooling energy needs. Solar Energy, 204, 708–719. https://doi.org/https://doi.org/10.1016/j.solener.2020.04.089

Gupta, S., Gupta, E., & Sarangi, G. K. (2020). Household Energy Poverty Index for India: An analysis of inter-state differences. Energy Policy, 144, 111592. https://doi.org/10.1016/j.enpol.2020.111592

Haines, A., Kovats, R. S., Campbell-Lendrum, D., & Corvalan, C. (2006). Climate change and human health: impacts, vulnerability, and mitigation. In Lancet (London, England) (Vol. 367, Issue 9528, pp. 2101–2109). https://doi.org/10.1016/S0140-6736(06)68933-2

Healy, J. D., & Clinch, J. P. (2002). Fuel poverty, thermal comfort and occupancy: results of a national household-survey in Ireland. Applied Energy, 73(3), 329–343. https://doi.org/https://doi.org/10.1016/S0306-2619(02)00115-0

Hwang, I. S. Y. (2006). When does stacking become vertical sprawl? WIT Transactions on Ecology and the Environment, 93, 283–292. https://doi.org/10.2495/SC060271

Kaboré, M., Bozonnet, E., Salagnac, P., & Abadie, M. (2018). Indexes for passive building design in urban context – indoor and outdoor cooling potentials. Energy and Buildings, 173, 315–325. https://doi.org/https://doi.org/10.1016/j.enbuild.2018.05.043

Kataoka, K., Matsumoto, F., Ichinose, T., & Taniguchi, M. (2009). Urban warming trends in several large Asian cities over the last 100 years. Science of The Total Environment, 407(9), 3112–3119. https://doi.org/https://doi.org/10.1016/j.scitotenv.2008.09.015

Keikhosravi, Q. (2019). The effect of heat waves on the intensification of the heat island of Iran's metropolises (Tehran, Mashhad, Tabriz, Ahvaz). Urban Climate, 28, 100453. https://doi.org/https://doi.org/10.1016/j.uclim.2019.100453

Khalid, R., & Sunikka-Blank, M. (2018). Evolving houses, demanding practices: A case of rising electricity consumption of the middle class in Pakistan. Building and Environment, 143, 293–305. https://doi.org/https://doi.org/10.1016/j.buildenv.2018.07.010

Kong, X., Lu, S., Li, Y., Huang, J., & Liu, S. (2014). Numerical study on the thermal performance of building wall and roof incorporating phase change material panel for passive cooling application. Energy and Buildings, 81, 404–415. https://doi.org/https://doi.org/10.1016/j.enbuild.2014.06.044

Latulippe, N., & Klenk, N. (2020). Making room and moving over: knowledge co-production, Indigenous knowledge sovereignty and the politics of global environmental change decision-making. Current Opinion in Environmental Sustainability, 42, 7–14. https://doi.org/https://doi.org/10.1016/j.cosust.2019.10.010

Malik, J., & Bardhan, R. (2020). Energy target pinch analysis for optimizing thermal comfort in low-income dwellings. Journal of Building Engineering, 28, 101045. https://doi.org/https://doi.org/10.1016/j.jobe.2019.101045

Mathew, A., Khandelwal, S., & Kaul, N. (2018). Analysis of diurnal surface temperature variations for the assessment of surface urban heat island effect over Indian cities. Energy and Buildings, 159, 271–295. https://doi.org/https://doi.org/10.1016/j.enbuild.2017.10.062

McCarthy, M. P., Best, M. J., & Betts, R. A. (2010). Climate change in cities due to global warming and urban effects. Geophysical Research Letters, 37(9). https://doi.org/10.1029/2010GL042845

Mhuireach, G. Á., Brown, G. Z., Kline, J., Manandhar, D., Moriyama, M., Northcutt, D., Rivera, I., & Van Den Wymelenberg, K. (2020). Lessons learned from implementing night ventilation of mass in a next-generation smart building. Energy and Buildings, 207, 109547. https://doi.org/https://doi.org/10.1016/j.enbuild.2019.109547

Osunmuyiwa, O. O., Payne, S. R., Vigneswara Ilavarasan, P., Peacock, A. D., & Jenkins, D. P. (2020). I cannot live without air conditioning! The role of identity, values and situational factors on cooling consumption patterns in India. Energy Research & Social Science, 69, 101634. https://doi.org/https://doi.org/10.1016/j.erss.2020.101634

Pantavou, K., Theoharatos, G., Mavrakis, A., & Santamouris, M. (2011). Evaluating thermal comfort conditions and health responses during an extremely hot summer in Athens. Building and Environment, 46(2), 339–344. https://doi.org/https://doi.org/10.1016/j.buildenv.2010.07.026

Qaid, A., Bin Lamit, H., Ossen, D. R., & Raja Shahminan, R. N. (2016). Urban heat island and thermal comfort conditions at micro-climate scale in a tropical planned city. Energy and Buildings, 133, 577–595. https://doi.org/https://doi.org/10.1016/j.enbuild.2016.10.006

Santamouris, M., Pavlou, K., Synnefa, A., Niachou, K., & Kolokotsa, D. (2007). Recent progress on passive cooling techniques: Advanced technological developments to improve survivability levels in low-income households. Energy and Buildings, 39(7), 859–866. https://doi.org/https://doi.org/10.1016/j.enbuild.2007.02.008

Santamouris, Mattheos., & Kolokotsa, D. (2013). Passive cooling dissipation techniques for buildings and other structures: The state of the art. Energy and Buildings, 57, 74–94. https://doi.org/https://doi.org/10.1016/j.enbuild.2012.11.002

Synnefa, A., Vasilakopoulou, K., Kyriakodis, G.-E., Lontorfos, V., De Masi, R. F., Mastrapostoli, E., Karlessi, T., & Santamouris, M. (2017). Minimizing the energy consumption of low income multiple housing using a holistic approach. Energy and Buildings, 154, 55–71. https://doi.org/https://doi.org/10.1016/j.enbuild.2017.07.034

Zhai, Z., Johnson, M.-H., & Krarti, M. (2011). Assessment of natural and hybrid ventilation models in whole-building energy simulations. Energy and Buildings, 43(9), 2251–2261. https://doi.org/https://doi.org/10.1016/j.enbuild.2011.06.026

EXTENDING VISUOSPATIAL ANALYSIS IN DESIGN COMPUTING

An Exploration with a Novel GPU-Based Algorithm and Form-Based Codes

GEOFF KIMM[1], MARCUS WHITE[2] and MARK BURRY[3]
[1,2,3]*Swinburne University of Technology.*
[1]*gkimm@swin.edu.au, ORCID 0000-0002-9643-8968*
[2]*marcuswhite@swin.edu.au, ORCID 0000-0002-2238-9251*
[3]*mburry@swin.edu.au, ORCID 0000-0001-7743-7719*

Abstract. This paper responds to a gap observed between the contemporary capacity for calculation and analysis of visibility of built environment features, such as buildings, in digital urban and architectural computational research models and the functionality of off-the-shelf software tools available to professionals. The research investigates the potential of visibility analysis to be embedded and extended within computational-based workflows of software tools to better meet urban design and planning industry needs. We introduce a novel method for visibility calculation that exposes output data for further analysis within a computational workflow and implement it in a game development engine used by software tool providers. Based in our engagement with a local government authority, we then use that method to demonstrate a workflow in the context of form-based building codes in which the visual impact of a building is considered rather than prescriptive limits on dimensions and use. Our results indicate the novel method has substantial performance improvements compared to an alternative mode of visibility calculation and that software providers could more thoroughly integrate and extend visibility analysis to meet industry needs.

Keywords. Design Computing, Viewsheds, Isovists, GPU Shader, Unity 3D, Genetic Algorithm, Generative Design, Form-based Building Codes

1. Introduction

Built environment design computing workflows that support visibility analysis, if well implemented, can be an invaluable tool in urban densification design and planning. Growing urban populations place pressures on the resources and infrastructure of cities and the capacities of those responsible for their planning, design, and maintenance (United Nations 2019). The consequent urban densification, if managed skilfully, can contribute positively to social well-being through efficiencies of scale, improved sustainability, and health outcomes (Buxton 2014;

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 655-664. © 2023 and published by the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA), Hong Kong.

Trubka et al. 2010). Aspects of good urban densification design and planning rely on the understanding of visuospatial influences of urban features, such as form-based building codes, in which the visual impact of a building is considered rather than stringent limits on use and dimensions, or urban greening strategies (Talen 2013; Yang et al. 2009). This paper questions how design computing workflows may efficiently and accessibly incorporate and extend urban visibility analysis as an integral component of the urban design process to address real-world problems.

2. Background and Context

The fundamental element of urban visibility analysis is the relationship between a feature of interest – an attractive park or a selected building, perhaps - and a location from which its visibility is to be tested – possibly a point on a building façade or a vantage point on a nearby street, correspondingly. Notably, that relationship between the feature and location is 1D and is, in effect, a hit test of a notional ray extending from the viewpoint location toward where the feature of interest may (or may not) exist. That test, although 1D, may be oriented in 2D, 3D, or, as 2.5D, pseudo-3D space; it, in its sweep through these higher dimensions, becomes a viewshed (Figure 1).

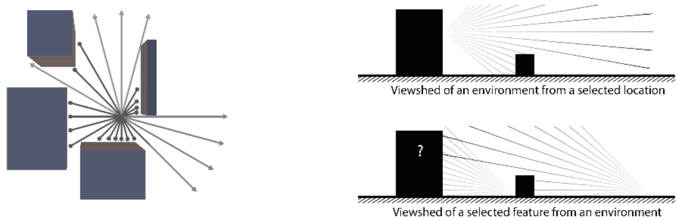


Figure 1. Conceptual diagrams of a 2D viewshed (left) and viewsheds from a location and of a feature (right).

It is that latter pseudo-3D viewshed analysis that has been prevalent in visuospatial computing workflows. Early work focused on calculating viewsheds in digital elevation models (DEM) in which each cell or raster pixel represents a height value, beginning with methods for which “preparing terrain data in computer-readable form” would “usually be punched cards” (Travis 1975). This 2.5D approach has continued with a succession of improvements on algorithms and data structures to lead to more efficient calculation, including with methods that exploit the parallel processing of the graphics processing unit (GPU) (Zhao et al. 2013). The DEM method is suited to terrain analysis that is not impeded by the 2.5D modelling limitation that no surface may exist above any other.

Other methods apply in 2D space, typically for the architectural plan, in which the viewshed is calculated as a bounding polygon, akin to what the protagonist of Abbott’s 1884 Flatland: A Romance of Many Dimensions would experience of his 2D universe. The essential concept of this polygon, or isovist, has been extended to statistical measures of perceptual qualities – at the most elementary, area and perimeter length – in 2D and 3D, and isovist fields in which measures taken throughout the space of an environment can be mapped as a scalar field or heatmap-style visualisation (Benedict 1979, Batty 2001).

Contemporary advances in computing power and the take up of 3D modelling in BIM, PIM, and other workflows are strengthening the case for effective 3D approaches. Extending the quantised modelling of DEMs, Morello and Ratti (2009) efficiently computed 2.5D isovists within a 3D voxel grid and evaluated their application to quantitative indicators of visual elements. For vertex-based 3D modelling, the parallel computing of GPUs is used to efficiently calculate viewsheds by Chao et al. (2011) and others, and 3D isovists have been explored (Derix et al. 2008). Adapting tools from outside the AEC industry to urban issues, White and Langenheim (2014) explored the appropriation of light sources within 3D animation software to assess view quality from buildings.

Visuospatial modelling as described above typically stops at an immediate stage of presentation of visualisation of viewsheds or their meta-analysis, at which the design computing workflow halts and the human is brought into the design loop. Users can benefit from this visual analysis and apply their understanding of urban environments as complex systems and of the nuanced considerations of real-world projects – that may not easily be accommodated in design computing algorithms – to further iterations of the design loop. Might not, though, the human-in-the-loop benefit also from extending further the computation on visuospatial analysis outputs in that same design loop?

Some approaches do propagate analysis results further within the design computing workflow. Viewshed analysis in 2D and 3D has been linked with agent-based modelling (ABM) to simulate pedestrian movement with respect to experiential qualities of line of sight, spatial openness, or visual occlusion (Aschwanden et al. 2008; depthmapX development team 2017). Akin to ABM, Lee and Stucky (1998) employed viewsheds within a DEM to calculate least-cost paths that maximise or minimise their visibility from or of non-path DEM cells. Via the DecodingSpaces Toolbox (toolbox.decodingspaces.net), viewshed analysis has been embedded within the Grasshopper visual scripting environment to allow visible entities to be explicitly identified and for visibility data to be passed further along the computational design workflow of a Grasshopper script. In 2.5D, viewshed analysis has been adapted to optimise precinct building envelopes for solar amenity as experienced by ABM pedestrians (Kimm 2020). In other generative work, Schneider and König (2012) applied computational optimisation to diverse isovist field statistical measures to generate building footprints.

However, a review we conducted of leading commercial precinct modelling software platforms targeted to design and planning professionals indicates that this data extensibility does not typically carry over to off-the-shelf tools available to industry. Spacemaker (autodesk.com/spacemaker), an Autodesk product for designing site development proposals, permits users to undertake view analysis as either view distance – a measure of how distant the view is at a point on a façade – or view to area – a measure of whether an area of interest is visible from a given façade point. As with the view to area of Spacemaker, CityEngine (esri.com/cityengine) supports the calculation of viewsheds for which the visibility information from a point, and within a limited field of view, is projected onto the urban model. That visibility information – default red for hidden and green for visible – may then be viewed by the user from other perspectives within the model. The planning and design urban digital twin application VU.CITY (www.vu.city) provides viewshed analysis as *zone of theoretical visibility* in which ground-level analysis of the visibility of a selected building is undertaken. The heatmap-style visualisation provides a colour scale corresponding to the visible vertical proportion of the building. Unsigned Studio (unsignedstudio.com), a division of the international engineering, design, and advisory firm Aurecon, produces siteLab, a built environment visualisation tool that operates from precinct level to the scale of architectural detail, that provides what-if view frustum testing of CCTV camera installation for public safety. Many geographic information system (GIS) packages can perform viewshed analysis. For example, GRASS GIS (grass.osgeo.org) can calculate an output image of which modelling cells are visible from a selected point on a raster terrain.

While these representative industry examples are varied in their application of visibility analysis, their use has characteristics in common. They are highly graphically based analyses that are primarily oriented to visual display of data within the 3D scene. Although some like CityEngine, which includes a breakdown of the share of layer visibility and like interpretation, go further, information typically is focused toward human- rather than machine-oriented additional processing.

Our research therefore considers how visuospatial analysis may be embedded within a design computing workflow in which its outputs may propagate further, and how this may be implemented in a way accessible and relevant to industry software solution providers. It first presents a novel GPU-based algorithm demonstrated in a game development engine. Using that algorithm, it then examines a context developed in collaboration with a local government authority of outer Melbourne, Australia, to test extending visuospatial analysis in a design computing workflow.

3. Method

A novel urban visibility algorithm was developed in the Unity 3D game engine ("Unity") to efficiently expose viewshed analysis for further use in a design computing workflow. The algorithm was then tested in a real-world context.

The GPU-based urban visibility analysis algorithm extends two precedents. Schneider et al. (2014) used a GPU-based method for real-time solar analysis that can assess performance criteria at the early design stage. In their method, a cube map is rendered for a given point within a digital 3D urban model against a sky image that embodies radiation values for a single instant in time or a specified period. The illuminance and daylight factor at the given point can be calculated from the rendered image by summating the intensity values of the sky that is unoccluded by the massed buildings of the 3D model. Chirkin et al. (2018) extended the approach of Schneider et al. to evaluation of massed objects of a 3D urban scene directly, and demonstrated the generalised applicability of their approach to analyses including distance to the closest object and visible ground area. In the methods of Schneider et al. and of Chirkin et al., evaluations may be taken across a multitude of points to display a heatmap-style precinct-level visualisation.

The adaptation of the 3D GPU pipeline of the precedents was extended in an algorithm for viewshed analysis of individual programmatically specified urban features and their occlusion by their context; through this flexibility, it is extensible to experiential perspectives within the urban scene rather than simply global or static perspectives. Visibility analysis data of a specified feature is exposed to subsequent stages of a design computing workflow.

The algorithm is a workflow of six key steps implemented in this research in the Unity 3D game engine ("Unity") (Figure 2).

- Two cameras, A and B, are instantiated in the scene and positioned to view the subject building from a desired location. Camera A is set to clear a shared target render texture to a predefined background colour and to render only the subject building. Camera B is set to render any obscuring context.
- A shader, being a small program that instructs a GPU how to draw parts of the scene, uses camera A to render the building with a gradient according to its height.
- A second shader on camera B renders all context with a predefined context colour.
- A compute shader, being a program that allows user-defined parallel data processing on the GPU, accepts the shared render texture and calculates a histogram and other statistical measures of the visibility of the rendered building.
- The output of the compute shader is waited on asynchronously and is added to a queue data structure when available.
- Results read from the queue are processed per application requirements.

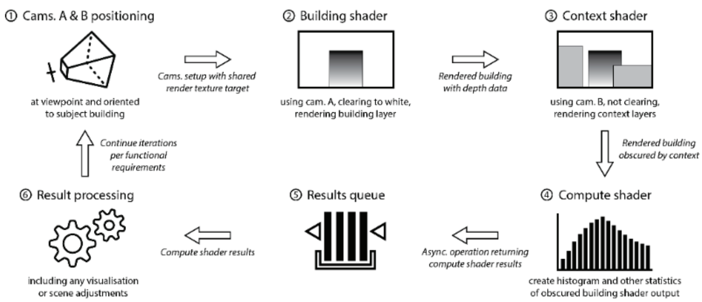


Figure 2. The algorithm of the viewshed calculation as implemented in Unity.

Unity was selected as an implementation platform for its industry acceptance, accessibility, and broad use in built environment analysis and simulation contexts (Huang et al. 2021). It is utilised in the AEC and urban design and planning industries by platforms including VU.CITY and siteLAB, each discussed above, and offers the Unity Reflect product suite of BIM AR and VR with connection to Autodesk Revit (unity.com/products/unity-reflect).

The application of the algorithm for exploiting viewshed analysis outputs in stages of the design workflow beyond immediate visualisation was tested. For this, a test context was identified in consultation with an outer Melbourne local government authority ("the LGA"). The results were then analysed for performance.

4. Results

The test context of a "serious gap" was identified in discussion with the LGA. "One of the challenges we have in [planning and design for] an urban context" is that "often everything is perceived [by the community] from a height perspective [of number of floors alone]". There is a need for form-based building code digital tools that can assist visualising buildings as "a full package", rather than in simple terms of "height and setbacks", and that could look at the "form of the built environment and how that interface is with the surroundings" and help answer "how does it [a building] actually visually relate to the space of the surround?"

To test the algorithm for form-based codes, an elementary precinct modelling environment was developed in Unity (Figure 3). The environment supported three key precinct entities: building massing and its creation and arrangement; a subject building envelope of parameterised footprint and height; and one or more viewsheds, each oriented towards the subject building envelope and anchored by a user-moveable point. For each viewshed, an in-scene screen shows to the user a view of the subject building. A pattern was applied to the subject building envelope in 3 m bands in the X, Y, and Z axes to highlight its form. A simple user interface was provided to set for the building envelope minimum and maximum allowable heights and a preferred average height (and, by extension, a preferred floor area using an assumed floor-to-floor height of 3 m). The user may also assign a [0,1] importance weighting to each viewpoint.

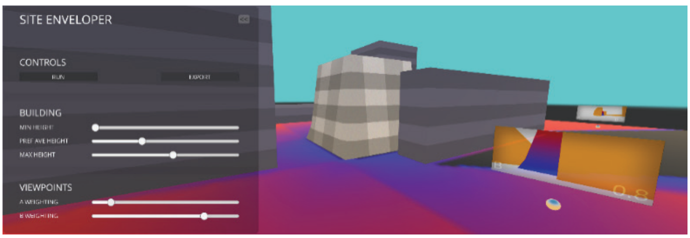


Figure 3. The elementary precinct modelling environment in Unity showing building massing (left rear and mid-right); subject building envelope (centre) and visualisation of its visible height heatmap (ground plane); viewpoints and their screens (right); and a simple user interface (right fore).

A genetic algorithm we developed in C# was applied to generate candidate form-based code solutions for the subject building envelope with respect to the weighted viewsheds (Figure 4). The genetic representation of the subject building envelope for each individual was encoded as a scalar field of heights. The fitness function was formulated to A) minimise the difference of the solution building envelope average height and the user-set preferred average height and B) minimise the weighted visibility of the building envelope from the viewpoints. In lay terms, the envelope optimisation seeks to maintain a desired building volume while hiding that volume from each viewpoint according to its assigned importance insofar as is possible.

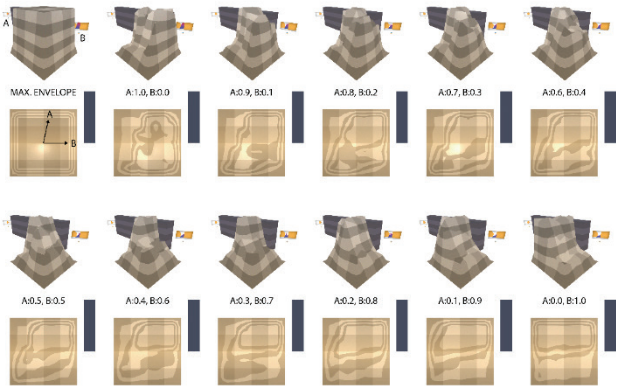


Figure 4. Perspective and orthographic views of 11 candidate solutions for different weightings of two points, A and B, from A: 1.0, B: 0.0 to A: 0.0, B: 1.0 in 0.1 increments. The configuration of building massing is the same as seen in Figure 3.

Performance was tested for solution convergence of the genetic algorithm and computational speed. The tool was tested with a genetic algorithm of population size of 150 individuals each of chromosome length 81 representing a site envelope of an 8 x 8 cell grid. Results indicate the fitness plateaus towards 200 generations (Figure 5). Each generation had a mean calculation time of 901.83 ms (n = 200) on a Windows 10 laptop (i7-8750H, 32GB RAM, NVIDIA GeForce RTX 2070, 1080p). A complex environment was tested by replacing the buildings in the standard scene with massed tree models, thereby increasing the number of triangles the GPU must process for viewpoint A from 2.8k to 673.1k and for viewpoint B from 1.4k to 325.5k. Each generation in this scenario had a mean calculation time of 1,130.4 ms (n = 200).

A traditional hit test method was appraised in the standard scene for comparison. Given the target image used for rendering in the novel approach was 512 x 256 pixels, a minimum of 131,072 hit tests would be needed to achieve the same level of detail. To account for the population size of 150 individuals, each being evaluated for two viewpoints, sets of 39,321,600 building hit tests were undertaken (n = 5) giving a mean time of 58.54 s, and not including any overheads of site envelope geometry manipulation and the genetic algorithm framework.

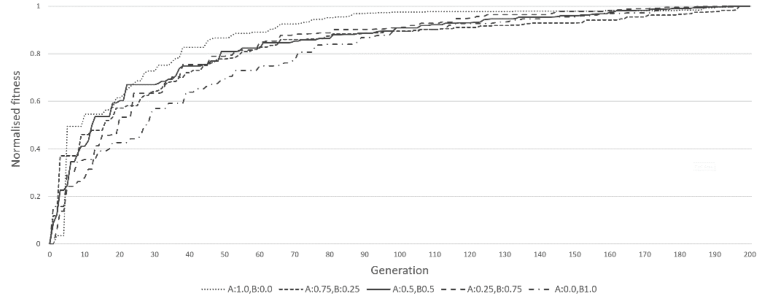


Figure 5. Normalised fitness versus generations for the urban configuration used in Figure 4.

5. Discussion

The experimentation presented in this paper has tested how visuospatial analysis may be embedded and propagated in a design computing workflow. The exploration was conducted with a test case of a real-world industry problem with application of a novel GPU-based visuospatial analysis algorithm.

The GPU shader urban visibility algorithm introduced in the method of this paper exposes visuospatial data for further processing in a design computing workflow. It demonstrated performance approximately 64 times faster than a comparable hit test implementation and has an efficiency that permits interactive application for even complex urban scenes. The level of detail of the analysis can be altered for use case requirements by adjusting the resolution of the render target image. The method does not rely on physics-based intersection testing of 3D geometry and is hence applicable to analysis of 3D models - vegetation, for example - that could otherwise be computationally expensive to test accurately.

Although the algorithm was implemented in this research to analyse building visibility, modifying the processing on the GPU can adapt it to other urban features and metrics, such as vegetation and green view. This study exploited specific features of the Unity 3D game engine graphical system and is extensible to other development platforms like Unreal Engine, and to direct implementation with technologies including OpenGL and DirectX.

The applied test considered the embedding of visuospatial analysis within a design computing workflow and used a platform used in industry software solutions. The building envelopes generated reflected a spatial variance in accord with viewpoint weightings. In the context of form-based building codes developed in collaboration with a local government authority, and community perceptions of the number of floors as a foremost planning and design criteria, the outcomes indicate a possibility for visuospatial digital tools to change the way in which the public is engaged. While the detailed investigation of this possibility is beyond the scope of this research, early indications suggest these tools have the potential to provide the community a more nuanced understanding beyond simple floor count metrics.

Two extensions were considered but not implemented for this research phase. First, linking viewpoint vectors and weightings to ABM pedestrian circulation would be straightforward and would provide a ready basis for each of building and precinct level modelling. Second, envelope generation could accommodate more detailed exploration of design options; for example, mass modelling could permit cantilevers and spatial voids, visibility assessment could consider facade treatments, and additional analysis such as shadowing of public space could be incorporated. In implementing these factors, envelope optimisation would need to be meticulously balanced against rivalrous considerations and practical limitations lest it converges on unworkable outcomes. In this research the optimisation results in recognisable building envelopes due to the assessment of building volume against its visibility and to the inherent constraints of scalar field modelling of envelope height. In further development, a naive implementation could spawn unworkable forms; modelling spatial voids, for instance, may create forms that merely converge on the spatial volume hidden from viewpoints or that poorly allow for structural elements.

Although moving closer to the "full package" of visual analysis desired by the LGA, these two extensions were not pursued in this early exploration as they risk the output solutions becoming difficult to intuitively comprehend or to match to the input conditions and hence less persuasive if used in community engagement. We see this bind as raising the broader question, that appears elsewhere in AI research, of balancing the use of AI outputs that may be black box and not explainable versus sticking to potentially less powerful or less nuanced but more intuitive solutions.

The digital data of view analysis is carried further in the design computing workflow in this research than immediate visualisation, yet ultimately the human is brought back into the design loop. The resulting envelopes of the tool are indicative only and there is no readily definable stopping condition for the algorithm that is optimal in objective terms. Each envelope is not prescriptive and must be interpreted by the designer or planner as spatial bounds to observe or break according to their nuanced understanding of urban environments. Further research could study the utility and interpretation of generated envelopes and how far such data can or should flow in a design computing workflow - for the human must return to the design loop if we are not considering a fully automated design machine with all its innate social and technical challenges.

In our research, we have introduced a novel analysis algorithm to investigate a gap between the potential to further propagate visuospatial data within a design computing workflow and the functionality offered by typical software tools available to urban built environment professionals. In the context of form-based codes, the results demonstrate the human within the design computing loop can benefit from extension of computation on visuospatial analysis outputs, and that this strategy can be feasibly employed by developers of industry software solutions.

References

Aschwanden, G., Halatsch, J., & Schmitt, G. (2008). Crowd simulation for urban planning. In *Education and Research in Computer Aided Architectural Design in Europe (eCAADe) 2008*, https://doi.org/10.52842/conf.ecaade.2008.493.

Batty, M. (2001), Exploring isovist fields: space and shape in architectural and urban morphology. *Environment and Planning B*, 28, 123–150, https://doi.org/10.1068/b2725.

Benedict, M.L., (1979). To take hold of space: isovist and isovist fields. *Environment and Planning B*, 6, 47-65, https://doi.org/10.1068/b060047.

Buxton, M., (2014). The expanding urban fringe: impacts on peri-urban areas. In B. Maheshwari, R. Purohit, H. Malano, V. P. Singh, & P. Amerasinghe (Eds.), *The Security of Water, Food, Energy and Liveability of Cities* (pp. 55-70). Springer, Dordrecht.

Chao, F., Chongjun, Y., Zhuo, C., Xiaojing, Y., & Hantao, G. (2011). Parallel algorithm for viewshed analysis on a modern GPU. *International Journal of Digital Earth*, 4(6), 471-486, https://doi.org/10.1080/17538947.2011.555565.

Chirkin, A., Pishniy, M., & Sender, A. (2018). Generilized visibility-based design evaluation using GPU. In *Learning, Adapting and Prototyping - Proceedings of the 23rd CAADRIA Conference - Volume 2* (pp. 483-492) Tsinghua University, Beijing, China, https://doi.org/10.52842/conf.caadria.2018.2.483.

depthmapX development team. (2017). depthmapX (Version 0.6.0) [Computer software]. Retrieved from https://github.com/SpaceGroupUCL/depthmapX/

Derix, C., Gamlesæter, Å., & Carranza, P.M. (2008). 3D isovists and spatial sensations: two methods and a case study. *Movement and Orientation in Built Environments: Evaluating Design Rationale and User Cognition*, p.67.

Huang, X., Kimm, G., & Burry, M. (2021). Exploiting game development environments for responsive urban design by non-programmers: melding real-time ABM pedestrian simulation and form modelling in Unity 3D, https://doi.org/10.52842/conf.caadria.2021.2.689.

Kimm, G. (2020). Actual and experiential shadow origin tagging: A 2.5 D algorithm for efficient precinct-scale modelling. *International Journal of Architectural Computing*, 18(1), pp.41-52, https://doi.org/10.1177/1478077119895218.

Morello, E. & Ratti, C., (2009). A digital image of the city: 3D isovists in Lynch's urban analysis. *Environment and Planning B*, 36(5), 837-853, https://doi.org/10.1068/b34144t.

Schneider, S. & König, R. (2012). Exploring the generative potential of isovist fields: the evolutionary generation of urban layouts based on isovist field properties. *30th Education and Research in Computer Aided Architectural Design in Europe (eCAADe)* (p. 355), https://doi.org/10.13140/RG.2.1.2569.9042.

Schneider, S., Tonn, C., Bielik, M., Donath, D., and Ruth, J. (2014). Real-time solar analysis - introducing a GPU-based method for calculating solar related performance criteria in the early design stages. In *Building Simulation and Optimization (BSO14)*, London.

Stucky, J.L.D. (1998). On applying viewshed analysis for determining least-cost paths on digital elevation models. *International Journal of Geographical Information Science*, 12(8), pp.891-905, https://doi.org/10.1080/136588198241554.

Talen, E. (2013). Zoning for and against sprawl: the case for form-based codes. *Journal of Urban Design*, 18(2), pp.175-200, https://doi.org/10.1080/13574809.2013.772883.

Travis, M.R., Elsner, G, Iverson, W, & Johnson C. (1975). *VIEWIT: computation of seen areas, slope, and aspect for land-use planning (Vol. 11)*. Department of Agriculture, Forest Service, Pacific Southwest Forest and Range Experiment Station.

Trubka, R., Newman, P., & Bilsborough, D. (2010). The costs of urban sprawl–Infrastructure and transportation. Environment design guide, pp.1-6.

United Nations, Department of Economic and Social Affairs, Population Division (2019). World Urbanization Prospects: The 2018 Revision (ST/ESA/SER.A/420). New York: United Nations.

White, M. & Langenheim, N. (2014) Quantitative assessment of view from within high density development using a perverted light modelling approach, In *7th Making Cities Liveable Conference*

Yang, J., Zhao, L., Mcbride, J., & Gong, P. (2009). Can you see green? Assessing the visibility of urban forests in cities. *Landscape and Urban Planning*, 91(2), pp.97-104. https://doi.org/10.1016/j.landurbplan.2008.12.004

Zhao, Y., Padmanabhan, A., & Wang, S. (2013). A parallel computing approach to viewshed analysis of large terrain data using graphics processing units. *International Journal of Geographical Information Science*, 27(2), 363-384, https://doi.org/10.1080/13658816.2012.692372.

AN URBAN BUILDING ENERGY SIMULATION METHOD INTEGRATING PARAMETRIC BIM AND MACHINE LEARNING

Forecast The Impact Of Surrounding Conditions On Solar Accessibility

JONG BUM KIM[1], SEONGCHAN KIM[2] and JAYEDI AMAN[3]
[1,3]*University of Missouri Columbia.*
[2] *Western Illinois University*
[1]*kimjongb@missouri.edu, 0000-0001-8762-9554*
[2]*s-kim@wiu.edu, 0000-0002-0780-383X*
[3]*jayediaman@mail.missouri.edu, 0000-0002-6128-8293*

Abstract. This research investigates a method of urban building energy simulation (UBES) by integrating Building Information Modeling (BIM), building simulation, and algorithm-based prediction to forecast the impact of surrounding conditions. In the urban context, building energy performances are determined not only by the individual building design but also by the building's surrounding context. Many energy performances are sensitive to outdoor and surrounding building conditions, such as neighbouring building volumes, heights, and spaces between buildings. However, such surrounding conditions were overlooked because they can exponentially increase the complexity of urban modeling and simulation. In that regard, the research sought to investigate a novel framework to take advantage of accurate performance simulations and algorithm-based fast predictions. This paper presents our UBES method implemented from three research phases: (i) building a parametric urban model in BIM to provide simulation inputs, (ii) creating a parametric simulation interface to produce training and validation data, and (iii) creating a prediction interface using a Support Vector Machine (SVR) algorithm. Lastly, the paper elaborates on the findings from the prediction results.

Keywords. Urban Energy Simulation, Solar Accessibility, Surrounding Conditions, Parametric BIM, Machine Learning, Support Vector Machine, Sustainable Cities and Communities

1. Introduction

This research investigates a method of urban building energy simulation (UBES) to forecast the impact of surrounding conditions on solar accessibility in urban development scenarios. The proposed method integrates parametric Building Information Modeling (BIM), building simulation, and Machine Learning (ML)-based prediction.

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 665-674.

Substantial UBES approaches have explored building-physics-based simulation for accuracy and historical data-driven estimation for urban-scale expeditious prediction. They are distinctive in simulation cost, accuracy, scalability, level of detail, and predictability for unbuilt development scenarios. Previous studies presented that the inaccuracy or ambiguity in UBES originates from excessive abstraction and neglect of the surrounding condition effects in individual building simulation (Nutkiewicz et al., 2021).

The research focuses on the surrounding conditions in the urban context. Building energy performances are determined not only by the building design but also by the building's surrounding context. Many energy performances are sensitive to outdoor and surrounding conditions, such as neighboring building volumes, heights, and spaces between buildings. However, such surrounding conditions were simplified or not fully addressed in many UBES methods because they can exponentially increase the complexity of energy modeling. Consequently, this simplification could make the simulation results less applicable to the real world.

In that regard, the research seeks a novel framework to take advantage of accurate building simulations and algorithm-based predictions. This paper presents our UBES method implemented in three research phases: (i) building a parametric urban model in BIM to provide physical attributes of surrounding conditions, (ii) creating a parametric simulation interface to produce training and validation data, and (iii) creating a prediction interface using ML algorithms. First, the research collected the modeling information from the recently adopted urban design regulations, such as the urban design regulation variables and the geography information of the surrounding blocks. Second, the simulation module analyzed the solar accessibility of individual parcels, iterated a large number of simulations to test the surrounding conditions, and produced training and validation data for ML prediction. For ML, the research explored Linear Support Vector Machine (SVM), a widely applied algorithm for energy simulation and sensitivity analysis.

2. Research Context

2.1. SURROUNDING CONDITION SCENARIOS OF THE PLANNED BUILT ENVIRONMENT

Surrounding conditions in this research refer to heterogeneous physical conditions of individual blocks, parcels, and buildings that can impact the environmental context of individual sites. Surrounding conditions variate the environmental context, such as neighboring building height, setbacks, streetscapes, traffic lanes, landscape, etc. The energy performance in the urban context is determined by how individual buildings perform, whereas individual performances are influenced by their surrounding conditions and urban morphology.

For the most part, information on the existing conditions is collectible from a geospatial dataset or on-site data collection. When considering growing demands for urban retrofitting, urban planning and design guidelines are critical datasets, such as neighborhood master plans, comprehensive plans, and land-use and zoning regulations. This research collected the existing condition information from GIS and the planning information from Smart Growth regulations. Smart Growth regulations are fast-growing urban planning and design methods developed by U.S. Environmental Protection Agencies (EPA). Smart Growth regulations control building form and design with several tools. The Transect is a classification tool in Smart Growth to assign the type of design requirements to individual parcels and blocks. The idea is to provide diverse housing and workspace options with gradually changing building use, density, and design characteristics.

Figure 1 shows two development plans of the current Smart Growth regulations. Two examples show how Transect is used in the recent planning regulations and how it diversifies individual parcels' surrounding conditions. The block colors indicate the Transect types, and other annotation symbols designate the entrance locations, setbacks, front façade locations, etc. P1 and P2 in the left map have similar parcel dimensions and orientations, but they have different surrounding conditions, including different transects of south-facing parcels, the main traffic lane location, and the block orientations. P3 and 4 in the right-side map have opposite directions of the front façade. Even though two parcels belong to the same Transect and need to follow the same building design standards, their different surrounding conditions may change their environmental context.

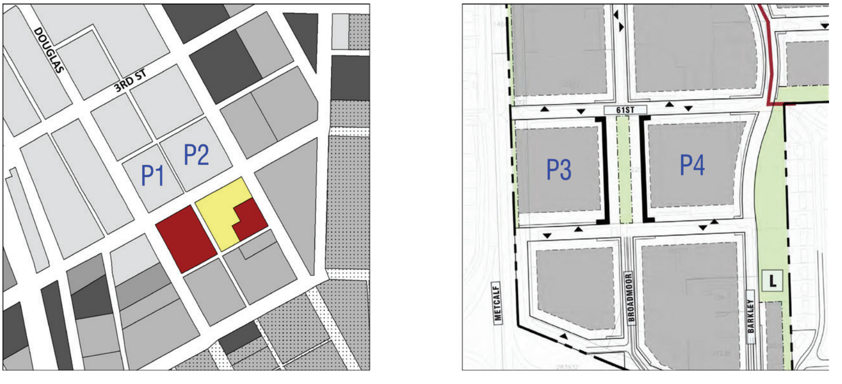


Figure 1 Various surrounding conditions in Smart Growth regulations. Notes: (Left) West Gateway Form-based Code, Kansas (Right) Lee's Summit Comprehensive Plan, Missouri

Transect-based planning promotes consistent styles and homogeneous functions within the same transect zones. However, where multiple transects are assigned, the pattern of surrounding conditions becomes more complex. In other words, the individual blocks within the same Transect have varying surrounding conditions. In UBES research, this random surrounding condition can quickly increase the range of what-if scenarios, making the modeling and simulation drastically time and effort-intensive or less viable. It increases complexity in UBES research because they yield bi-directional and dynamic energy flow between buildings, parcels, and blocks.

3. A UBES METHOD FOR SURROUNDING CONDITION SIMULATION

This research proposed a UBES method to address the surrounding conditions in energy modeling and simulation. The proposed method fused parametric BIM, physics-based simulations, and algorithm-based prediction to mediate the above problems and challenges. This research was built on our previous works on parametric BIM modeling, multi-criteria simulations, and immersive visualization (Kim et al., 2021).

- Parametric BIM: The research collected the modeling and geography information from the recently adopted urban design regulations and GIS data. Then, parametric urban models were built using spatial operations in BIM and a set of custom parameters that can change the BIM objects' geometry.
- Simulation iteration: Second, the simulation module analyzed solar accessibility of individual parcels, iterated a large number of parameters to test the surrounding conditions, and produced training and validation data for the ML prediction.
- SVM model: Lastly, for ML development, the research explored Linear SVM to perform prediction of the impact of surrounding conditions on the solar radiation.

The research implemented the method with parametric modeling in BIM and software prototype development. For parametric modeling, we used Autodesk Revit, its custom object modeling interface, and custom parameters to manipulate the object geometry. For energy simulation, we built simulation interfaces for solar radiation analysis. The research used crossover simulation tools such as Ladybug tools and TT Toolbox written in Grasshopper to take advantage of rich simulation interfaces and validated simulation engines. To connect Revit and Grasshopper, Rhino.inside was used. For algorithm-based prediction, the research explored SVM with an open-source library such as scikit-learn (Pedregosa et al., 2011.) and Python scripting.

This research explored SVM for regression analysis using the training data obtained from physics-based energy simulations. The parametric BIM modeling, simulation module creation, and parametric BIM/ simulation integration were built on our prior works. The following sections describe the methods and processes of the SVM model creation.

3.1. SVM FOR ENERGY SIMULATION

For years, machine-learning approaches have been tested in the energy simulation sector (Chen et al., 2020). Support Vector Machine (SVM), developed by Vapnik with colleagues in 1995, is a supervised learning technique for data classification and regression analysis. Two major applications of SVMs are data classification and regression analysis. In classification analysis, SVMs assign a given dataset to the binary linear classifier, distributes data points in high-dimensional space, and maximize the distance (margin) between the classified data groups. In regression, SVMs map the input dataset into kernel-induced feature space to place the data points inside a tube space of a given radius. The kernel is an essential concept of SVMs. When the dataset in low-dimensional space does not present linear separable patterns, SVMs can transform this low-dimensional space into high-dimensional space. These mapping rules between two spaces are the kernel function. Widely used functions include linear, nonlinear, polynomial, radial basis, and sigmoid. The selection of kernel functions considerably affects the SVM results.

Because of powerful data classification capabilities, SVMs have been widely used in text categorization, handwriting character recognition, bioinformatics, etc. In addition, SVMs are increasingly used in building energy simulations. Example areas include building energy consumption prediction (Dong et al., 2005; Jung et al., 2015), annual electric load forecast (Azadeh et al., 2008), estimating heating energy (Paudel et al., 2014), PV system performance (Ogliari et al., 2013), and wind power system (Chen et al., 2013). Studies show that SVMs can solve nonlinear problems with a small number of training data (Zhao et al., 2012). Other studies presented that SVM is more capable than different algorithms such as Artificial Neural Networks (Li et al., 2010). Several studies described SVR's benefits over decision tree (DTR) and random forest (RFR). Ma et al. (2019) claimed that the SVR could correctly predict building energy use with a significant nonlinear capacity, improving accuracy and stability in energy performance forecasting. Furthermore, Others explained that the SVM outperforms other techniques when limited samples are available (Chen et al., 2020).

SVR's regression function is summarized as follows:

$$f(x) = w \cdot \phi(x) + b$$

Where f(x) is the forecasting value, x represents the input variable, w represents the weight coefficient, b represents the deviation value, and (x) represents the high dimensional feature space. In addition, we used the R-squared value and mean absolute error (MAE) to evaluate the SVR model as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} |y_i - \widehat{y_i}|^2}{\sum_{i=1}^{n} |y_i - \overline{y_i}|^2}$$

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |y_i - \widehat{y_i}|$$

Where y_i is the simulation results, (y_i)ˆ indicates the prediction results, y ̄_i is the average of the simulation results, and n represents the sample numbers of the testing dataset.

3.2. DATA FLOW

Figure 2 explains the workflow of surrounding condition simulations using the SVR model. First, we produced inputs, including regulation variables as IVs and energy simulation results as DVs. This research simulated the solar radiation levels from a building-physics-based simulation (S1). The input dataset was then split into training

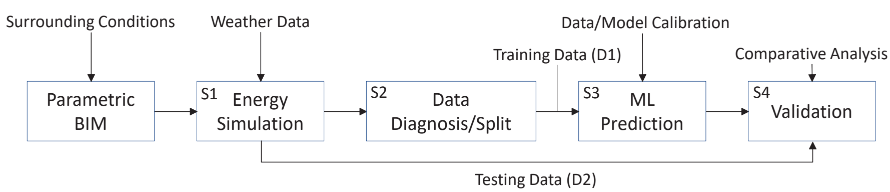


Figure 2. Workflow of the surrounding condition simulation

(D1) and testing datasets (D2). In this phase, the data process was performed to normalize the variables before the SVR model building. Then we fed the training IVs and DVs into the training module (S3). The SVR then projected the testing IVs dataset into kernel-induced feature space. Lastly, the SVR results were compared with the testing dataset selected from simulation results (S4) for validation.

4. Case Study

The research performed case studies of two study areas with different surrounding conditions. Two sites were selected from the Downtown Form District of Overland Park Unified Development Plan. The City of Overland Park is a southern neighborhood of the Kansas City metropolitan area. The City adopted the master plan and Smart Growth regulations to revitalize the downtown area and the Metcalf corridor, a historical corridor connecting southern cities with the Kansas City downtown area. First, we analyzed the regulating map, transect types, and building design standards. Then we built two block models using Revit and its custom modeling interface. GIS data provided the district subdivision information. The building design standards in the regulation provided modeling information of the building sections (Table 1).

Parameters	Transect 1	Transect 2	Transect 3
Building height	7.3m~18.2m	7.3m~18.2m	7.3m~14.6m
Building depth	18.2m~21.3m	18.2m~21.3m	15.2m~18.2m
Setback	1.5m~3m	1.5m~3m	1.5m~3m
Roof height	2.7m~3.6m	2.7m~3.6m	2.7m~3.6m

Table 1. Regulation variables of three transects for parametric BIM.

Figure 3 shows two block models showing different subdivision shapes and surrounding conditions. The regulating map in Smart Growth indicated the position of the building front façade, street types, setback requirements, etc. The building design standards in Smart Growth regulations showed building use, height, roof shape, and other sectional configurations. The street color indicates the streetscape types and the building types. The maps indicate the allowed building front façade location, making

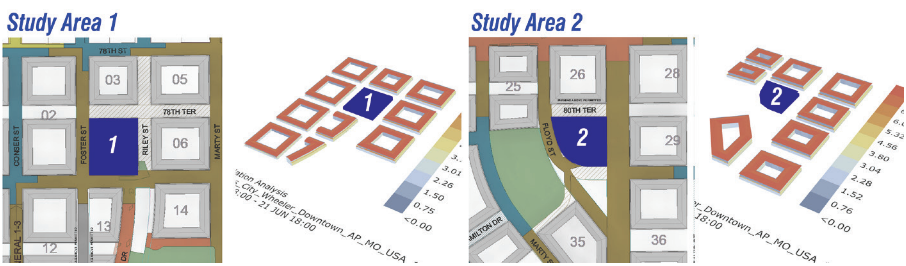


Figure 3. Regulating plans and solar radiation maps of two study areas

the building masses similar to the parcel shapes.

Next, the simulation interface iterated parameter sets and ran 500 simulations for each case. We simulated solar radiation values using the Revit district models and Ladybug tools. The resulting dataset consisted of input variables as IVs and solar radiation values per parcel as DV. The inputs include building height, depth, setback, and roof pitch. The system interface written in Python redefined the result datasets and crosschecked the missing values for data pre-processing.

In the final phase, we split the result datasets randomly into 80% for training and 20% for testing before the SVR model training. For each case, 500 independent data points were divided into 400 for training and 100 for testing. Lastly, we fed the training and testing data into the SVR kernel and produced ML-based prediction models.

5. Results and Findings

The research implemented several calibration methods to improve the reliability and accuracy of input and output of prediction models, including random state, learning curves, and K-fold cross-validation. Then, a comparative study was performed between the simulation results and the ML-prediction results.

Input data diagnosis: To diagnose the simulation dataset, we analyzed multicollinearity, normality, homoskedasticity, and independence. Standardized Regression Coefficients showed that the building depth and setback have the most substantial effect on solar accessibility.

Random State: We used the random permutation function to select the reliable training and data sets from the simulation results. The function generated the pseudo-random number of the selected datasets, calculated the regression coefficient, and selected the most reliable dataset. The pseudo-code made the selected datasets not only trackable but also reusable. The selected dataset presented significance (Case1: R2adj=0.979, Case2: R2adj=0.977).

Learning Curves: Learning curves are used to plot the error of the training set and validation set as a function of training size to measure the model's generalization performance. Overfitting occurs when a model performs well on training data but poorly on validation data. In contrast, if it does poorly on both, it is underfitting. The learning curve for Case 1 is presented on the left, and the learning curve for Case 2 is presented on the right in the figure 4. As seen in the learning curve plots, the line of the training set begins at zero. As additional instances are added to the training set, the model's ability to fit the training data worsens. As a result, the error on the training data increases until it approaches a plateau indicating additional instance has no effect on the average error. For the validation set, when the model is trained on a few instances, it cannot generalize appropriately, so the validation error is initially greater than the training set. The model then learns as more training instances are provided, and validation error gradually decreases until it reaches a plateau close to the training set. The two learning curve plots demonstrate that the SVM models used in this study generalize adequately, without overfitting or underfitting issues.

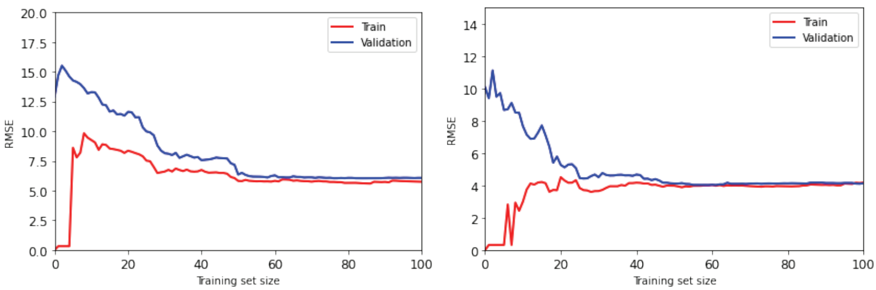


Figure 4. Learning cures (Left: Study area1, Right: Study area2)

K-fold cross-validation: In addition, we conducted K-fold cross-validation to compare and select a regression model with a lower bias. In K-fold, the given dataset is split into a K number of folds. In the first iteration, the first fold is used for testing, and the rest are used for training. In the second iteration, training is done with the second fold data. Then, the cross-validation process is repeated k times, with each of the k subsamples used exactly once as the validation data. The k results can then be averaged to produce a single estimation. The dataset was randomly shuffled using the "Gamma" and "Epsilon" hypermeters of the SVR model. Seven gamma (¥) and six epsilon (€) values together with 42 possibilities resulted in 420 fits of the sample sets. For accuracy testing, we employed R-square and root mean square deviation (RMSE). In case 1, the most fitted ¥ value was 0.1, whereas, in case 2, the most fitted value was 0.5. The € value that fit both instances the best was 1. The advantage of this method over repeated random sub-sampling is that all observations are used for both training and validation, and each observation is used exactly once. 10-fold cross-validation is commonly used [16], but in general k remains an unfixed parameter.

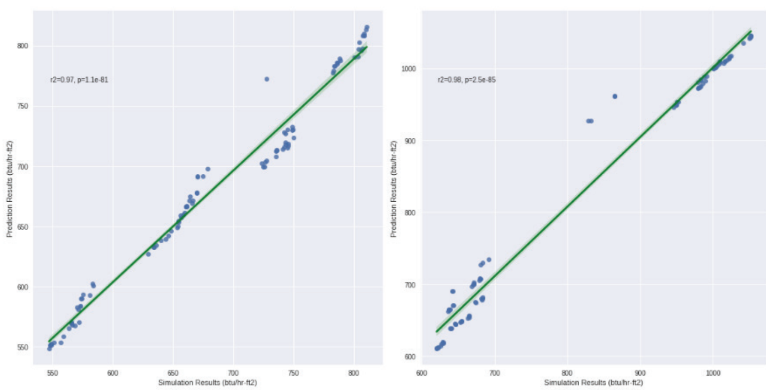


Figure 5. Regression analysis (Left: Study area 1, Right: Study area 2)

Comparative Analysis: Finally, we carried out regression analysis of simulation and prediction results in Figure 5. Strong correlation is observed (r2 is 0.97 and 0.98

respectively), while several datasets presented weak correlations. Table 2 shows a subset of samples to illustrate the percentage variance in the analysis results. Each sample contains 12 IVs as well as solar radiation results generated by the physics-based simulation and ML-based prediction model. For both cases, very few samples show deviations greater than 1%. The results demonstrate the high accuracy and reliability of the proposed method.

Study area 1				Study area 2			
Case No.	Simulation (btu/hr-ft2)	ML-prediction (btu/hr-ft2)	Deviation (%)	Case No.	Simulation (btu/hr-ft2)	ML-prediction (btu/hr-ft2)	Deviation (%)
103	810.19	814.44	0.52	178	682.29	678.77	-0.52
219	640.69	634.51	-0.97	360	1024.45	1017.45	-0.69
422	546.84	549.68	0.52	128	673.83	675.25	0.21
111	668.55	675.29	1.00	144	1041.23	1035.64	-0.54
373	574.18	590.43	2.75	203	1000.86	1000.61	-0.02

Table 2. A comparison of selected samples from simulation and prediction

6. Conclusion

In this paper, we investigated the UBES method to handle many simulation scenarios associated with the surrounding conditions. The proposed approach attempted to strike a balance between the accuracy and the reliability of simulations, simulation size and latency, and the level of detail and scale. To do so, we examined the fusion of parametric BIM, simulations, and ML with SVR.

Algorithm-based prediction can handle a large number of district layout scenarios with less effort and computing time, particularly when multi-criteria performances in urban development are considered. As a result, the substitution of simulations with ML predictions will yield significant benefits in saving cost and time. The software prototype outperformed physics-based parametric simulations in terms of reliability, capacity, and speed. The heterogeneous connection between modeling and simulation platforms required high computing power, but it eliminated redundant and error-prone input processes by using BIM information. The case study tested small-scale urban blocks. In reality, there are infinite potential layout patterns because of the curse of dimensionality and numerous provisions in Smart Growth regulations. As a result, the proposed method still has limitations in generating optimal development scenarios to minimize the environmental impacts.

The simulation results demonstrated that the surrounding conditions significantly impact each block's solar accessibility, which needs to be considered in urban energy modeling and simulations. From case studies, we inferred that SVR predicts different solar radiation patterns according to each parcel's surrounding conditions. However, it is still required to address "black box effect" of algorithm-based prediction. This research used the linear kernel, but it is suggested to investigate different kernel types. In addition, we noticed that SVR is sensitive to the sample size and the sampling data sets. Potential overfitting issues and sensitivity in sampling need to be addressed in future research to gain the validity and credibility of SVR-based surrounding condition predictions.

References

Azadeh, A. & Ghaderi, S.F. (2008). A simulated-based neural network algorithm for forecasting electrical energy consumption in Iran. Energy Policy. 36. 2637-2644. https://doi.org/10.1016/j.enpol.2008.02.035

Chen, J., Zhu, J., Wang, Z., Zheng, X., & Zhang, B. (2013). Scalable Inference for Logistic-Normal Topic Models. NIPS.

Wu, M., Liao, J., Yu, L., Lv, R., Li, P., Sun, W., Tan, R., Duan, X., Zhang, L., Li, F., Kim, J., Shin, K. H., Seok Park, H., Zhang, W., Guo, Z., Wang, H., Tang, Y., Gorgolis, G., Galiotis, C., & Ma, J. (2020). 2020 Roadmap on Carbon Materials for Energy Storage and Conversion. Chemistry, an Asian journal, 15(7), 995–1013. https://doi.org/10.1002/asia.201901802

City of Overland Park, Kansas (2016). Downtown Form-Based Code. Retrieved from http://online.encodeplus.com/regs/overlandpark-ks/doc-viewer.aspx#secid-7075

Dong, B., Cao, C., & Lee, S.E. (2005). Applying support vector machines to predict building energy consumption in tropical region. Energy and Buildings, 37(5), 545-553. https://doi.org/10.1016/j.enbuild.2004.09.009.

Jung, H.C., Kim, J.S., & Heo, H. (2015). Prediction of building energy consumption using an improved real coded genetic algorithm based least squares support vector machine approach. Energy and buildings, 90, 76–84. Elsevier B.V. https://doi.org/10.1016/j.enbuild.2014.12.029.

Kim, J. B., Aman, J., & Balakrishnan, B. (2021). Forecasting performance of Smart Growth development with parametric BIM-based microclimate simulation. Projection, Proceedings of the 26th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2021, Hong Kong. https://doi.org/10.52842/conf.caadria.2021.1.411

Li, R., Chen, X., Balezentis, T., Streimikiene, D., & Niu, Z. (2021). Multi-step least squares support vector machine modeling approach for forecasting short-term electricity demand with application. Neural Computing and Applications, 33(1), 301–320. https://doi.org/10.1007/s00521-020-04996-3

Ma, Z., Ye, C., Li, H., & Ma, W. (2018). Applying support vector machines to predict building energy consumption in China. Energy Procedia, 152, 780–786. https://doi.org/10.1016/j.egypro.2018.09.245

Nutkiewicz, A., Choi, B., & Jain, R. K. (2021). Exploring the influence of urban context on building energy retrofit performance: A hybrid simulation and data-driven approach. Advances in Applied Energy, 3, 100038. https://doi.org/10.1016/j.adapen.2021.100038

Ogliari, E.; Grimaccia, F.; Leva, S.; Mussetta, M. Hybrid predictive models for accurate forecasting in PV systems. Energies 2013, 6, 1918–1929. https://doi.org/10.3390/en6041918

Paudel, S., Elmtiri, M., Kling, W.L., Corre, O.L., & Lacarrière, B. (2014). Pseudo dynamic transitional modeling of building heating energy demand using artificial neural network. Energy and Buildings, 70, 81–93. https://doi.org/10.1016/j.enbuild.2013.11.051.

Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, & M. Blondel. (2011). "Scikit-Learn: Machine Learning in Python." Journal of Machine Learning Research 12: 2825–2830.

Zhao, H., & Magoulès, F. (2012). A review on the prediction of building energy consumption. Renewable and Sustainable Energy Reviews, 16(6), 3586–3592. https://doi.org/10.1016/j.rser.2012.02.049

Building Information Modelling

METAVERSE EXPANDS THE COMMON DATA ENVIRONMENT

TSUKASA ISHIZAWA[1]
[1]*Project Associate Professor, The Institute of Industrial Science, The University of Tokyo*
[1]*ishizawa@iis.u-tokyo.ac.jp, 0000-0001-5089-2877*

Abstract. Building information modeling (BIM) is critical to the industry's digital transition. Its diffusion is also expected to improve project economics. However, maintaining adequate information in the common data environment (CDE) typically becomes the bottleneck. Conversations still take place outside of the model workflow, easily becoming untraceable. This article proposes the fusion of the BIM and metaverse, which can serve as a CDE (MaCDE). The metaverse workspace hosts the project meetings, presenting the model as the virtual venue. Avatars' utterances can be automatically recorded, indexed, and tracked. The enhanced traceability of information reduces the burden of dealing with contractual issues and motivates professionals to join, stay, and utilize the model environment. The concept demonstrates that the model can be continuously leveraged regardless of its Level of Development. The increased visibility helps align stakeholder expectations with the model deliverable. The improved coverage and searchability of the archive help recapture the lost revenue, improving the project profitability and the ongoing effort to enhance productivity.

Keywords. Metaverse, Common Data Environment, Building Information Modeling, Digital Twin, Philosophy of Engineering

1. Introduction

1.1. THE PENETRATION OF BIM USE

Building information modeling (BIM) is increasingly fundamental for the industry's digital transition (Dx). Dodge's recent report showed that the adoption rates of technologies such as reality capture, the Internet of Things, or wearable devices in BIM projects are significantly higher than the others (A. Jones and Laquidara-Carr, 2021). In order to improve the industry's productivity stagnating for decades, continuous effort to spread BIM use is a paramount issue.

Although BIM products have been in the market for more than a quarter-century, only a limited number of organizations and projects have fully implemented them to date (Malleson, 2018; A. Jones and Laquidara-Carr, 2021). There is a global shortage of talented professionals. The struggle of academia to produce qualified practitioners

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 677-686.

is still ongoing (Elias, Issa, and Wu, 2022). For various reasons, the model often ceases to be updated in the middle of the project, and the BIM process itself tends to be abandoned. We are at the stage to encourage the BIM projects to complete the model life cycle throughout the project period.

1.2. COMMON DATA ENVIRONMENT

Data coverage and accuracy management typically become the bottleneck for broader use of digital building information. ISO19650 defines Common Data Environment (CDE) as an "agreed source of information for any given project or asset, for collecting, managing and disseminating each information container through a managed process." Therefore, "information exchange involves the sharing and coordination of information through a CDE, using open standards whenever possible and clearly defined operating procedures to enable a consistent approach by all organizations involved" (British Standards Institution, 2018, p.5). However, human-based synchronization of the CDE can easily lead to lapses and incompleteness. Non-BIM-based communication is even more difficult to capture, track and manage appropriately. Little confidence in BIM and CDE's information quality demotivates stakeholders from utilizing the information they contain.

2. Research Objective

BIM still has room to demonstrate its greater value. Projects expect BIM to increase productivity and improve project economics. It should be viewed, verified, and trusted by a wider range of stakeholders; then it could serve as a Single Source of Truth (SSoT). Provided that any activity can be recorded at all times, such an archive will become compelling, persuasive evidence for the project closure process. How can we provoke such changes without altering the local treaty cultures?

The metaverse, which has rapidly received attention in recent years, conceptually encompasses not only digital native spaces that exist only virtually but also digital twins (DT) that represent the physical world. Its development could compensate for the current shortcomings and contribute to fulfilling the original expectation.

This paper proposes a novel concept for implementing the metaverse to overcome the problems confronted by today's BIM projects. The proposed framework, Metaverse as a CDE, aims to embrace more participants, increase the trustworthiness of the information it contains, and harness more benefits of the digital model.

3. Conceptual Framework

3.1. BIM AND METAVERSE

The metaverse, which originated in 1992, appears to have been dramatically popularized by Facebook's rebranding in 2021. Its technical intersection with BIM is an emerging field at the time of this writing. So far, published papers have analyzed its potential use incorporated with blockchain technology (Huang et al., 2022), proposed a novel design environment (Lee and Moon, 2022), and explored the new possibilities of architectures there (Tang and Hou, 2022). This domain has been previously researched as the topics of extended reality (XR) or telexistence. The visual advantage of immersive experiences has expanded creativity and improved consensus-building in projects (Alizadehsalehi, Hadavi and Huang, 2020). It can also be viewed as a virtual urban space from the smart city perspective (Moneta, 2020).

Like a well-organized building model, the metaverse-ready digital environment seems ready to invite the building users. Avatars represent the occupants, visitors, or end-users. These digital beings could be the professionals involved in the design and construction process. This shift brings about the changes explained in the difference between the "metaverse as a target" and "metaverse as a tool" (Dwivedi et al., 2022). Duan et al. listed the three-layer architecture of the metaverse, namely, ecosystem, interaction, and infrastructure, and the ecosystem includes user-generated content, economics, and artificial intelligence (Duan et al., 2021). When the project progress and model fall on the interaction and infrastructure, the ecosystem can be automatically understood as the generated documents, conversations, professional insights, and any other valuable information transacted there. When stakeholders seek such a virtual environment, the goal of improving overall productivity and project economics will be spontaneously strengthened.

3.2. LEVEL OF DEVELOPMENT AND DIGITAL TWIN PROTOTYPE

DT can be regarded as a form of the metaverse. Lee et al. proposed the DT-native continuum, in which the digital places explicitly intended to correspond are the DT among the infinite imaginable digital native spaces (Lee et al., 2021). The project model ultimately aims at the DT of the completed physical building. Level of Development (Level of Detail, LoD) is a measure that refers to the degree of geometric or attributional information maturity. For example, LOD100 model defined by the BIM Forum (Ikerd et al., 2021) is typically recognized as a simple massing representation containing few properties inside, which seems distanced from DT. However, it can also be considered a coarse and transient DT at the relevant moment as one of the Digital Twin Prototypes (DTP). DTP is the virtual description of a prototype product, which consists of the totality of the possible forms termed the Digital Twin Aggregate (DTA) (Jones et al., 2020; Singh et al., 2021). While immature digital representations may be of little use from the perspective of "metaverse as a target," the model that concurrently reflects the up-to-date state well should demonstrate its value to the stakeholders as a DTP in "metaverse as a tool."

4. Methodology

This study adopts the methodology of Staples' Philosophy of Engineering (Staples, 2014, 2015). Based on Popper's three-world ontological framework, BIM and the possibility of the metaverse as a tool can be understood along the elements of engineering and the roles played by the engineering theories (Figure 1). Among the 1–5 roles represented as arrows, this study focuses specifically on role #1: "characterizations of changes required to be brought about in usage situations represented as requirements specifications," where the observation in World 1 (world of physical entities and phenomena) transforms into World 3 (world of objective content: knowledge and products of thought that can be explicitly recorded or spoken.) First, necessary changes for further BIM utilization are identified based on the existing use cases (Section 5.1). Second, we attempt to link them to the specification represented as requirements (Section 5.2). Finally, we will focus on the main function of the artifact (Section 6.1) and consider supplemental technologies to enhance the functionality (Section 6.2).

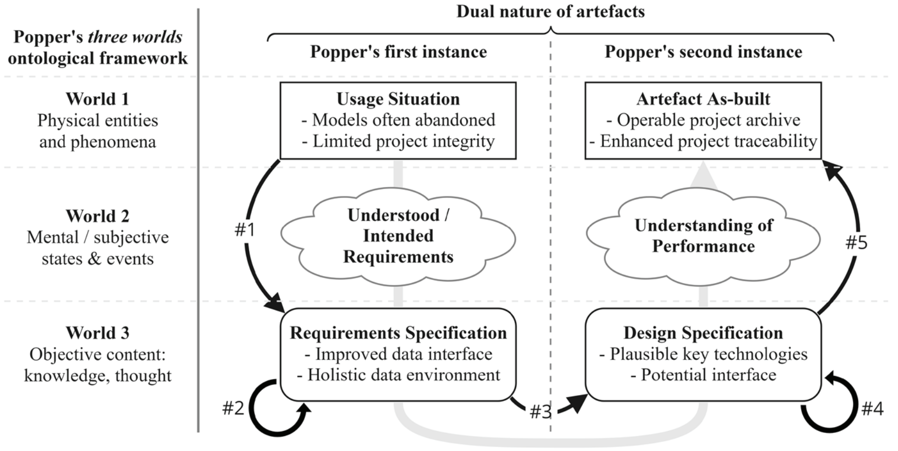


Figure 1. Proposal Framework based on Staple's Philosophy of Engineering.

5. Analysis

5.1. EXPECTED CHANGES IN USAGE SITUATION

Since its use typically declines over the phases, the incompleteness of the model lifecycle is a severe problem. The current approach to LOD generally assumes that it steadily grows throughout the project. While the BIM team aims to deliver a complete DT at the end, "offshore" communication outside the model-centric protocol can jeopardize the quality of the as-built model. Worse, the contract often requires the contractor to submit overwhelmingly detailed attribute lists without articulating post-occupancy operation. Eventually, BIM is abandoned mid-project without fulfilling the expected return on investment (ROI).

Limiting the as-built model requirements to those truly needed by stakeholders saves BIM by preventing such inflated expectations. Alignment of understanding can be fostered through consistent exposure to the model environment. Even though the CDE ensures accessibility, individuals rarely take their time to open the model file stored in the repository. It is unrealistic to expect all members to have adequate BIM skills; we always work with non-BIM professionals. Instead of this open-access style, mandatory meetings should be held in the virtual environment of the project-model metaverse, not in a physical meeting room.

BIM has played only a partial role in periodic project meetings, whether physical or virtual, typically like interior visualization or clash reports. Naturally, the model has only retroactively reflected the consensus recorded in the documents as the primary source of information. While BIM could become the primary database, it remains the adjunct to the classic set of documents such as meeting minutes, requests for information, and drawing sheets, as illustrated in Figure 2.

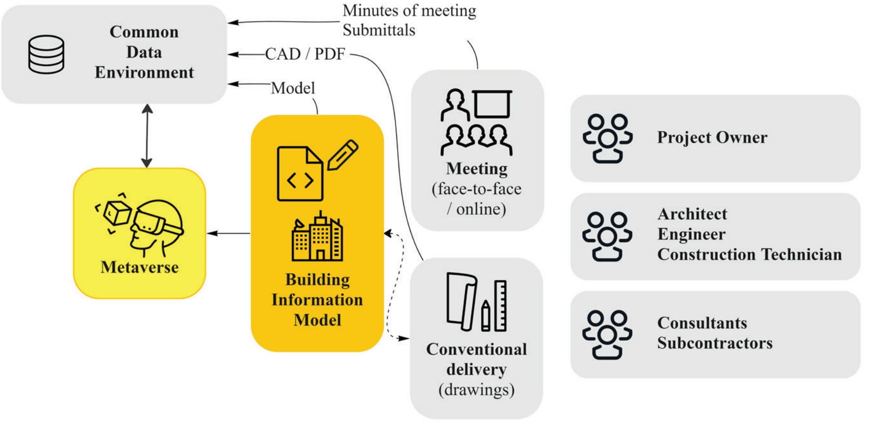


Figure 2. The current typical use case of BIM, CDE, and metaverse

The more we expect from BIM, the more we are disappointed when deficiencies are found. In reality, the compromised accuracy of geometry and attributes have been relatively painless for projects because the documents serve as enough evidence for project closeout, including the cases like litigations or disputes. Data correctness can easily fall into a downward spiral.

Suppose that BIM could successfully retain most of the information about the project. The low-LoD model can then serve as such functionality of a virtual project workspace. It is even more than just building digital representation. The avatars' voices can be recorded as audio, transcribed into texts, and related to the specific model place where the conversation took place. Accumulated conversations, tracked in the continuous digital workspace, form the basis for claiming design changes or construction defects. It incentivizes stakeholders to centralize every piece of data and knowledge, thus significantly improving the reliability of CDE.

The ROI of BIM is yet to be conclusively positive (Sompolgrunk, Banihashemi and Reza, 2021; Sompolgrunk et al., 2022). BIM enables organizations to optimize the design, reduce downside risk, and minimize contingency costs. However, improving the profits and streamlining the project management itself will be required primarily to recoup the subscription fees, talented human resources, and cloud computing costs incurred by the continued use of BIM. On top of those, it is crucial to enhance the sustainability of this industry.

5.2. REQUIREMENTS SPECIFICATION IN NEXT BIM

First, BIM should be positioned as the project's foundation instead of a mere add-on or parallel to the existing workflow. Encouraging the participation of non-BIM stakeholders by hosting meetings within the model metaverse suppresses decision-making unlikely to be recaptured. Increasing the immersive experience enhances exposure and familiarity with the model design, reducing the over-expectation of the model deliverable. It can encourage facility manager involvement prior to building completion, thus providing more opportunities to optimize the building information deliverable. The project should provide a shared metaverse environment to enjoy the above benefits, which can transform the model into a virtual world to which all stakeholders are invited.

Second, the model must become the SSoT of the project. The rigor of geometry and attributes can be enhanced when the model is exposed to the participants more frequently. It is further reinforced when the traditional meeting toolsets, such as detail drawings, hand-drawn sketches, and referenced images, are tagged to the appropriate model locations. A digital workspace that archives metadata like author, timestamp, and in-place location can multi-dimensionally record the project activities.

Third, the verbal communication of the avatars can be treated similarly to the pieces of project information in the virtual space. Transcriptions allow retrospective retrieval of conversations, improving the traceability of important events, especially variation orders. It automates the burden of creating documentary records and provides better resolution in chronological events, thereby contributing to a fairer fee system. The BIM should be linked to the metaverse meetings' voice (or chat) communication database.

6. Proposal

6.1. METAVERSE AS A COMMON DATA ENVIRONMENT

Suppose all the necessary project communication is migrated to the BIM-metaverse. The team creates a virtual space at the kick-off, where the members gather inside (or around, above) the up-to-date model placed as DTP. All communication is encouraged within the virtual workspace, regardless of the formality. The environment records each avatar's activities, including the location, field of view, speech, and text-based chat. Metadata indexes the captured data, including author, timestamp, and corresponding location. Additional documents or whiteboards used in the virtual space are similarly tracked, tagged, and stored.

The metaverse should chronologically manage the model updates while maintaining coordination points. It ensures the continuity of information by four-dimensionally aligning the DTPs until the completion. The digital world may initially contain only a mass model with a few digital people. As the project progresses, the virtual space gradually evolves into building elements that accommodate more stakeholders inside. Tracked communication as a database allows analysis for better management. For example, the communication heat map will verify the thoroughness of the design and construction discussion by element.

XR technology has been used for rendering and visualization, where luxurious materials and details are more appreciated though this is not the case here. We can count on the model space as the project script supervisor when considering BIM as an ever-changing continuum of cyberspaces (Figure 3). This paper calls it "Metaverse as a Common Data Environment (MaCDE)."

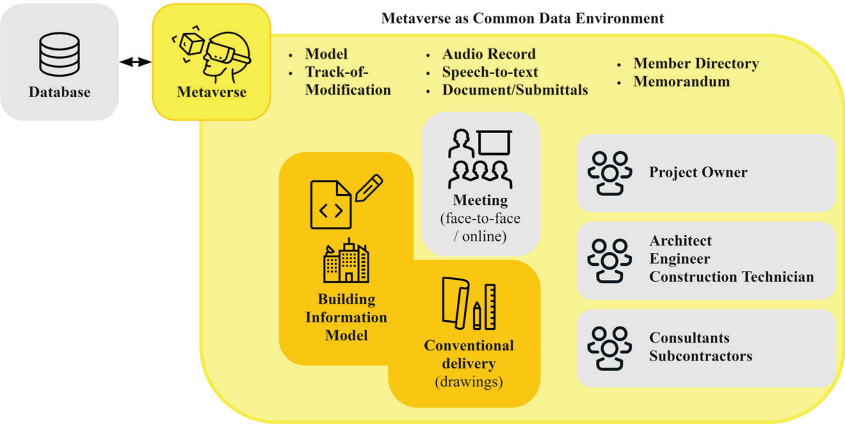


Figure 3. Metaverse as a CDE; based on "metaverse as a tool" concept

6.2. ADDITIONAL TECHNOLOGIES TO BE INCORPORATED

BIM Communication Format (BCF) associates textual comments with location information (Lindhard and Steinmann, 2014). It is helpful to record and manage the clashes detected during model coordination or non-conformities related to the objects' requirements. This structure can be used to store avatar conversations in the MaCDE.

Natural Language Analysis (NLA) rapidly achieves high-level accuracy in many languages. Today, the technique is developed enough to accurately and instantly convert intelligible speech into text data (Hu and Zhang, 2013; Jung and Lee, 2019). Pre-learning of voiceprints can further automatically identify the speakers. Natural Language Generation (NLG) can describe the changes observed in the statistical data, which can facilitate the model version control. While we can imagine extracting variation orders from the spoken context and automatically translating them to the model, this interpretation is challenging because design dialog is often metaphorical (Kahlon and Fujii, 2022). However, the recent rapid development of generative AI may partially overcome this problem.

In an open-access environment, the legitimacy of the model and its modification are paramount. In addition to having an administrative authority that constantly monitors and controls changes, distributed ledger technology (DLT), including blockchain, is plausible to prevent unwanted manipulations (Celik, Petri and Rezgui, 2021; Das, Tao and Cheng, 2021).

7. Discussion

Pandemics and recessions have given momentum to the spread of digitalization in the past few decades (Taylor and Soneji, 2022). To take advantage of learning and further promote the necessary digital transformation, we should use virtual spaces to improve what we create in the future.

Several themes of foreknowledge for the metaverse with BIM were extracted using the concept of falsification in the Philosophy of Engineering to identify the gaps in current use cases. BIM as a trustworthy project database becomes possible when the activities are systematically tracked within the virtual meeting space.

MaCDE could lead to improved model quality, utilization of the model after construction completion, and contribution to the broader urban model. Thus, this support improves the likelihood of closing the project model life cycle. On the other hand, BIM discussions to date often overemphasize the replacement of every aspect of the workflow with three-dimensional representations; model expression can be rather confusing and impractical for experienced yet drawing-oriented professionals. If contractually required drawings and documents are sufficient to solve problems, the BIM-based workflow should prioritize avoiding duplication of effort. The model represents the project from a metaverse perspective, but it need not be everything. BIM serves as a communication space to envision an environment, providing benefits to professionals of all generations and skill levels.

Possible sunk cost reductions through MaCDE include information consolidation, uploading and synchronization, and design variation orders. In lump-sum projects, the requirements are transient and vary significantly from start to finish. However, it is unlikely that the additional work required by the variation orders will be adequately recognized and compensated. Automatic and centralized communication tracking and indexing communications will facilitate the verification of change histories, eliminate legal issues, and fairly compensate for changes; this is expected to work well in business practices and contract cultures that do not emphasize written agreements (commonly found in East Asian countries). Under such business conventions, it is plausible to manage this content as an integral part of the central database since the scene-by-scene approvals constitute parts of the substantive content of the contract.

8. Limitations and Future Prospects

Some analytical processes may be performed outside the metaverse, which requires further investigation. For example, the CDE concept involves the process of reviews, approvals, and authorizations to differentiate the information for sharing and publishing. The document control platform typically handles this complex workflow, which is not necessarily integrated into the MaCDE.

The laggards, the last to adopt the new technology, are always present in every team. Because the concept is relatively new to the industry, most team members may be skeptical about adopting a metaverse platform. An inclusive approach to the laggards and the late majority is essential to encourage wider adoption.

Security for the CDE will become even more critical than it is today. The centralization of intelligence also implies an increased risk of cyber-terrorism. Meanwhile, societal demands will make failure increasingly unacceptable. There needs to be more experts in project information management, including data status, versions, and ownership. Training people and developing professionals to cope with such an environment remains a challenge for the future.

9. Conclusion

This paper concludes that the metaverse should be used as a tool for architectural projects to play the role of CDE. The activities automatically interpreted and tracked therein will strengthen the usage of BIM because the enhanced centrality and trustworthiness will increase the value of building information throughout the project.

The originality of this study is threefold. First, the promising application of the metaverse was deduced from the necessary changes in the current BIM use by the framework in the Philosophy of Engineering. Second, the concept for the project professionals was developed by applying the frameworks of DTP and metaverse as a tool. Third, the proposal's advantages are expected to recover the lost profits that construction processes inevitably experience. Such a measure applies to any project, in addition to the continuous efforts to improve productivity. Overall, this paper proposes a new recommendation for BIM use with the metaverse, which contributes to the industry's sustainability through an approach to enhance project economics.

Reference

A. Jones, S., & Laquidara-Carr, D. (eds.). 2021. *Accelerating Digital Transformation Through BIM: SmartMarket Report*. Dodge Data & Analytics. Retrieved August 24, 2020, from https://proddrupalcontent.construction.com/s3fs-public/DigitalTransformationBIMSMR_111021.pdf

Alizadehsalehi, S., Hadavi. A., & Huang, J.C. (2020). From BIM to extended reality in AEC industry. *Automation in Construction* 116. 103254. https://doi.org/10.1016/j.autcon.2020.103254

British Standards Institution. (2018). Organization and digitization of information about buildings and civil engineering works, including building information modelling (BIM) - Information management using building information modelling Part 1: Concepts and principles (No. BS EN ISO 19650-2:2018). BSI Standards Limited.

Celik, Y., Petri, I., & Rezgui,Y. (2021). Leveraging BIM and Blockchain for Digital Twins. In *2021 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC)* (pp.1–10). https://doi.org/10.1109/ICE/ITMC52061.2021.9570246

Das, M., Tao, X. & Cheng, J.C.P. (2021). BIM security: A critical review and recommendations using encryption strategy and blockchain. *Automation in Construction* 126. 103682. https://doi.org/10.1016/j.autcon.2021.103682

Duan, H., Li, J., Fan, S., Lin, Z., Wu, X., & Cai, W. (2021). Metaverse for Social Good: A University Campus Prototype. In *29th ACM International Conference on Multimedia.* (pp. 153–161). Association for Computing Machinery. https://doi.org/10.1145/3474085.3479238

Dwivedi, Y. K., Hughes, L., Baabdullah, A. M., Ribeiro-Navarrete, S., Giannakis, M., Al-Debei, M. M., … Wamba, S. F. (2022). Metaverse beyond the hype: Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. International Journal of Information Management, 66, 102542. https://doi.org/10.1016/j.ijinfomgt.2022.102542

Elias, R., Issa, R.R.A., & Wu, W. (2022). Progress on Building Information Modeling Education and Talent Acquisition. *International Journal of Construction Education and Research* 1–20. https://doi.org/10.1080/15578771.2022.2115174

Huang, H., Zeng, X., Zhao, L., Qiu, C., Wu, H., & Fan, L. (2022). Fusion of Building Information Modeling and Blockchain for Metaverse: A Survey. *IEEE Open Journal of the Computer Society 3*. 195–207. https://doi.org/10.1109/OJCS.2022.3206494

Hu, Z., & Zhang, J. (2013). BIM oriented intelligent data mining and representation. In *Proceedings of the 30th CIB W78 International Conference*. (pp. 280-289).

Ikerd, W. & Merrifield, D. & Vandezande, J. & Cichonski, W. (2021). *Level of development specification.* BIM Forum.

Jones, D., Snider, C., Nassehi, A., Yon, J., & Hicks, B. (2020). Characterising the Digital Twin: A systematic literature review. CIRP Journal of Manufacturing Science and Technology, 29, 36–52. https://doi.org/10.1016/j.cirpj.2020.02.002

Jung, N., & Lee, G. (2019). Automated classification of building information modeling (BIM) case studies by BIM use based on natural language processing (NLP) and unsupervised learning. Advanced Engineering Informatics, 41, 100917. https://doi.org/10.1016/j.aei.2019.04.007

Kahlon, Y. and Fujii, H. (2022) 'Visualization and inquiry into mental content in design activity: a case study of design interpretation', Artificial intelligence for engineering design, analysis and manufacturing: AI EDAM, 36, p. e2. https://doi.org/10.1017/S0890060421000172.

Lee, J., & Moon, I-Y. (2022). Method Of Design Drawing Conversion Automation for 3D BIM. International Conference on Future *Information & Communication Engineering* 13(1). 123–126.

Lee, L.-H., Braud, T., Zhou, P., Wang, L., Xu, D., Lin, Z., … Hui, P. (2021). All One Needs to Know about Metaverse: A Complete Survey on Technological Singularity, Virtual Ecosystem, and Research Agenda. CoRR, abs/2110.05352. Retrieved from https://doi.org/10.48550/arXiv.2110.05352

Lindhard, K. & Steinmann, R. (2014). BIM-collaboration processes-from fuzziness to practical implementation. In *European Conferences on Product and Process Modeling in The Building Industry (ECPPM).* (pp. 27–32).

Malleson, A. (2018). *National BIM Survey: summary of findings.* In National BIM Report 2018, 16–37. London: RIBA Enterprises Ltd. Retrieved March 24, 2019, from http://www.bimacademy.global/wp-content/uploads/2018/05/NBS0850-BIM-Report-2018-LR-.pdf

Moneta, A. (2020). Architecture, heritage and metaverse: new approaches and methods for the digital built environment. *Traditional Dwellings and Settlements Review 32(2)*. Retrieved November 17, 2022, from http://irep.ntu.ac.uk/id/eprint/41626/

Singh, M., Fuenmayor, E., Hinchy, E., Qiao, Y., Murray, N., & Devine, D. (2021). Digital Twin: Origin to Future. Applied System Innovation, 4(2), 36. MDPI AG. https://doi.org/10.3390/asi4020036

Sompolgrunk, A., Banihashemi, S., Hosseini, M.R., Golzad, H., & Hajirasouli, A. (2022). An integrated model of BIM return on investment for Australian small- and medium-sized enterprises (SMEs). *Engineering Construction and Architectural Management*. https://doi.org/10.1108/ECAM-09-2021-0839

Sompolgrunk, A., Banihashemi, S., & Reza, M.S. (2021). Building information modelling (BIM) and the return on investment: a systematic analysis. *Construction Innovation*. https://doi.org/10.1108/CI-06-2021-0119

Staples, M. (2014). Critical rationalism and engineering: ontology. *Synthese* 191(10). 2255–2279. https://doi.org/10.1007/s11229-014-0396-3

Staples, M. (2015). Critical rationalism and engineering: methodology. *Synthese* 192(1). 337–362. https://doi.org/10.1007/s11229-014-0571-6

Tang, S.K., & Hou, J-H. (2022). Designing a Framework for Metaverse Architecture. In *27th International Conference on Computer-Aided Architectural Design Research in Asia: Post Carbon, CAADRIA 2022* (pp.445–454). https://doi.org/10.52842/conf.caadria.2022.2.445

Taylor, S., & Soneji, S. (2022). Bioinformatics and the Metaverse: Are We Ready? *Frontiers in Bioinformatics* 2. 863676. https://doi.org/10.3389/fbinf.2022.863676

SYSTEMATIC REVIEW OF ORGANIZING THE DATA FLOW OF SYSTEMS USING BUILDING INFORMATION MODELING AND DISTRIBUTED LEDGER TECHNOLOGY

TOMOYA IZUMI[1] and KEISUKE TOYODA[2]
[1,2]*Department of Architecture, Faculty of Engineering, The University of Tokyo.*
[2]*Project Professor, The Institute of Industrial Science, The University of Tokyo.*
[1]*izumito@iis.u-tokyo.ac.jp, 0000-0002-9601-2342*
[2]*ktoyoda@iis.u-tokyo.ac.jp*

Abstract. As the digitalization in the architecture, engineering, and construction (AEC) industry has been rapidly progressing, building end users will require more coverage of quality data for their built environment. Despite its obvious need, it is seldom for the building information modeling (BIM) to be distributed for such broad use cases thus far. A potential reason for the situation is the lack of the public repository that can provide the adequate access control for various users. Distributed Ledger Technology (DLT), including blockchain, is one of the promising resolutions for the problem. Numbers of preceding proposals for DLT-based BIM repository systems (BIM-DLT system) have proved the utility to date. Understanding the statistical facts of their configurations will become a significant assistance to determine the plausible data distribution system, and thus, to strengthen the project information life cycle. This article aims to build a representative framework through the integrated literature review (ILR). The review results showed that the system's data flow was classified into two patterns. The judgment criteria for selecting the pattern of the system are expected to depend on the purpose of system implementation, and the purpose of automation through smart contracts is more important than the oracle problem.

Keywords. Building Information Modeling (BIM), Blockchain, Distributed Ledger Technology (DLT), Integrated Literature Review (ILR).

1. Introduction

Building information modeling (BIM) contributes to the digitalization in the architecture, engineering, and construction (AEC) industry for structuring building information throughout the lifecycle of a building and making it machine-readable (Jeong and Kim, 2022). It can be estimated that more and more innovations will occur in the future as the building information can be easily recorded, queried and transmitted (Ye et al., 2018). At that time, the value of BIM data will be recognized more than at present and used in various places, and thus, a mechanism that stimulates the data supply chain to ensure the supply and quality of data will become necessary (Nugroho et al., 2015). Despite its obvious need, it is seldom for BIM data to be distributed for such broad use cases thus far. A potential reason for the situation is the lack of the public repository that can provide the adequate access control for various users.

One solution to this problem is using distributed ledger technology (DLT), including blockchain. DLT manages data in an autonomous decentralized manner, offering features such as confidentiality, disintermediation, provenance tracking, non-repudiation, inter-organizational recordkeeping, change tracing, and data ownership (Turk and Klinc, 2017). In recent years, many studies have been conducted by taking advantage of these characteristics using systems that apply DLT to data management in BIM and even in the AEC industry (BIM-DLT system). However, this technology remains immature and has been primarily developed at the theoretical level, lacking practical applications (Figueiredo et al., 2022). It is plausible to imagine that the DLT implementation in AEC industry is further challenging than the others because of the overwhelming connections to the physical world.

2. Research question

The research question for this paper is how the data flow of the BIM-DLT system can be summarized. The ultimate goal of this study is to publish construction data with appropriate access controls and create an ecosystem that provides incentives for data producers. However, as mentioned earlier, the adoption of DLT in the AEC industry has been limited. Additionally, frameworks for evaluating system configurations are generally insufficient; thus, comparing systems by considering certain policies is difficult. For this reason, understanding the statistical facts of their configurations will significantly assist in determining the plausible data distribution system and strengthening the project information life cycle. Therefore, this paper aims to classify the data flow of the BIM-DLT system developed in previous studies. This research focuses on the data flow structure and how the BIM data is imported into or referenced from the DLT. If more than one classification is observed, its different uses must be investigated.

3. Method

The research methodology follows the concept of Integrative Literature Review (ILR) as following (Lubbe et al., 2020). In the review, evaluation items were selected with particular attention to the structure of the data flow and the form of BIM data in systems. Figure 1 shows the review procedure.

First, Scopus was used to search for published articles, and Table 1 presents the search queries. The returned list of studies was set as Group 1.

Second, each study in Group 1 was screened, and only the studies in which the BIM-DLT system was implemented were extracted (Group 2).

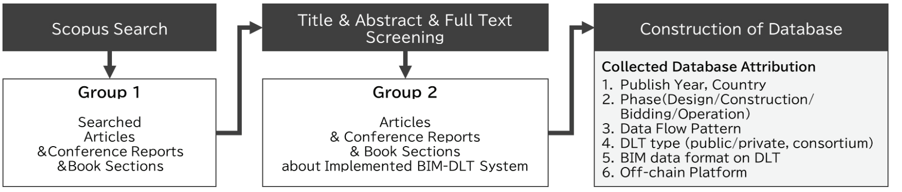


Figure 1. Research method

(TITLE-ABS-KEY (blockchain OR "smart contract*" OR "distributed ledger" OR "block chain" OR "block-chain" OR dlt) AND TITLE-ABS-KEY (bim OR "building information model*" OR "common data environment*")) AND PUBYEAR < 2023 AND (LIMIT-TO (DOCTYPE , "ar") OR LIMIT-TO (DOCTYPE , "cp") OR LIMIT-TO (DOCTYPE , "re") OR LIMIT-TO (DOCTYPE , "ch"))

Table 1. Queries used for Scopus search

Third, for each system in the Group 2 studies, the data was summarized and analysed according to the following four criteria: (1) data flow pattern (details in next section); (2) target phase (design/bidding/construction/operation); (3) unit of construction data to be imported into DLT (BIM file version, building element, property, task, etc.); and (4) how to reference BIM data from DLT (parts, data hash, etc.). The systems were respectively evaluated along the process when a paper contains multiple ones. We counted the system as individual ones when a single proposal applies to different project situations (i.e. to design and operational phases).

4. Analysis

The data retrieval conducted on 10 October 2022, resulted in 157 counts of systems, termed as Group 1. The subsequent screening process narrowed down the list to 39 cases as Group 2. When these were organized by year, a significant increase in the number of studies was observed after 2018 (Figure 2). Only in three studies in Group 2, multiple systems per study (2, 4, and 7 systems in one study) were implemented. A total of 49 systems extracted were analyzed.

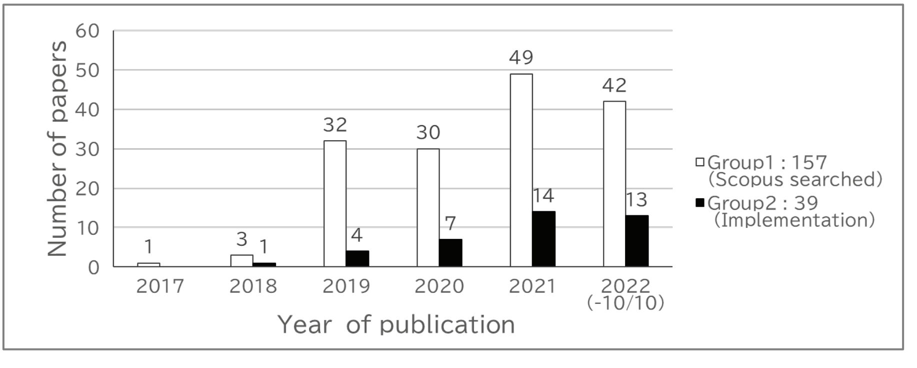


Figure 2. Number of papers by year

4.1. ANALYSIS A: DATA FLOW PATTERN

The observed data flow can be largely classified into two representative patterns termed as "Uni-directional Pattern" and "Bi-directional Pattern", elaborated in the following part. Except for 4 systems that could not be explicitly determined into the specific pattern, 41 out of 49 turned out to be typologized as Uni-directional Pattern and the rest 3 is as the other. The Uni-directional Pattern indicates that data move in only one direction from the off-chain repository to the DLT environment. Data from BIM, applications, and Internet of Things (IoT) devices are input to the off-chain repository and DLT environment and referenced by viewing applications or transaction viewers as needed. The data flow presented in (a) and (c) in the Figure 3 (function to input data into DLT environment) is an oracle implemented by decentralized applications (dApps), various Web3 interfaces, and dedicated oracle services. An off-chain repository is a storage or database that accumulates input data.

The Bi-directional Pattern differs from the Uni-directional Pattern in that it moves data from the DLT environment to an off-chain repository. Thus, there is a function or operation that applies the change defined on the DLT environment to the off-chain repository (data flow (h) in Figure 4). For example, Berawi et al. (2021) advocated the operation of using the finalized model, which was finalized and stored in the blockchain in the post-process. Wang et al. (2022) constructed a system that reflects BIM editing differences saved on the blockchain using the semantic differential transaction method to the user's local model.

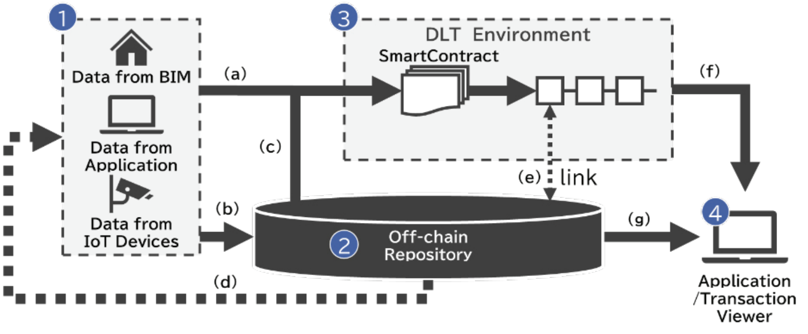


Figure 3. Data flow of DLT systems in the construction sector (Uni-directional Pattern)

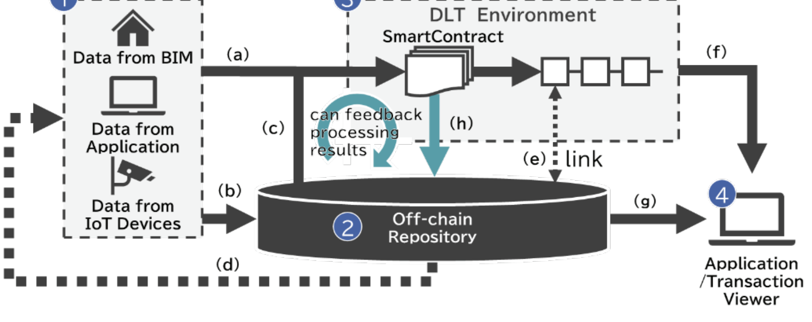


Figure 4. Data flow of DLT systems in the construction sector (Bi-directional Pattern)

4.2. ANALYSIS B: DLT TYPE

The DLT types used in the DLT environment in each system were organized for each target phase (Table 2). It was observed that the majority of the proposed systems were based on private blockchain (30 systems), followed by public blockchain (7 systems) and hybrid blockchain (combination of public and private blockchain, 3 systems).

Target Phase / DLT Type	Number of Systems						
	Design	Construction	Operation	Design & Construction	Design & Operation	Bidding & Construction & Operation	Total
Public Blockchain	3	4					7
Private/Consortium Blockchain	8	14	4	3	1	1	31
Hybrid Blockchain (Combination of Public and Private/Consortium Blockchain)			2			1	3
No explicit type selection available	3	4		1			8
Total	14	22	6	4	1	2	49

Table 2. Summary of DLT type by phase

4.3. ANALYSIS C: DATA UNIT

How finely the construction data handled by the DLT were captured based on the BIM data division unit was investigated (Table 3). Some measures of data fineness depended on not only building data structures, such as file versions, building elements, and properties, but also tasks, such as payments.

Target Phase / Data Unit	Number of Systems (Allow Multiple Counts)						
	Design	Construction	Operation	Design & Construction	Design & Operation	Bidding & Construction & Operation	Total
File Version	7	5	2	3		2	19
Area			1				1
Building Element	2	9					11
Property	4	3	2	1	1		11
Task	2	5	1				8
No explicit type selection available	1	2					3

Table 3. Summary of data units to be incorporated into DLT for each phase

4.4. ANALYSIS D: MODEL REFERENCE

The model reference methods were classified into following three patterns (Table 4). The first is a file-specific ID or hash that refers to the data in the off-chain repository (35 systems). The referencing methods include the content ID of files created by the InterPlanetary File System, hashes of files calculated by using hash functions such as SHA256, IDs assigned to each building element, and URLs of an off-chain repository. The second pattern involves recording part of files or differences between file versions (9 systems). The third pattern is storing no BIM data in DLT (11 systems). Examples include storing data obtained from non-BIM sources such as financial information or secondary data computed from BIM data and other documents.

Model Reference \ Target Phase	Number of Systems (Allow Multiple Counts)						
	Design	Construction	Operation	Design & Construction	Design & Operation	Bidding & Construction & Operation	Total
Hash Value (File)*	4	3	1	3			11
Hash Value (Part of File)*	3	3		1			7
ID (File)	1	1	1			1	4
ID (Part of File)		2			1		3
Link (Part of File)	1	2	1			1	5
NFT (Part of File)	2	1	1			1	5
Part of File		3		1			4
Difference between File Versions	3	2					5
No BIM Data on DLT	2	5	3				10
Others		1					1
No explicit type selection available	1	2					3

* contains CID used in IPFS

Table 4. Summary of model reference

5. Discovery and Discussion

Analysis A: Despite we can imagine another category, such as the combination of Uni-directional Pattern and Bi-directional Pattern, we can regard these two observed categories as the representative models of systems at the moment of writing.

Most of the systems were built with Uni-directional Patterns. A major concern with blockchains is how to guarantee that the external data input to the blockchain is high quality, which is called the oracle problem (Chainlink, 2020). Given this problem, when a system is constructed to ensure data reliability and immutability by implementing the autonomous decentralized and trustless structure of DLT, avoiding the return of the data that have been imported into the DLT environment back to the off-chain repository seems natural.

On the other hand, few systems have been built with the Bi-directional Pattern. However in the AEC industry, which concern real building, it is worth noting that the processing result in the DLT environment can be fed back to an off-chain repository. Because building design and construction management always requires decision-making by experienced professionals, the oracle problem cannot be completely omitted from the construction projects. Then, Bi-directional Pattern should have an advantage to reinforce trustability of information. For example, Blockchain-enabled IoT-BIM platform (BIBP)" advocated by Li et al. (2022) was motivated by making the BIM data agree using the blockchain as a single point of truth.

In summary, how problematic the blockchain oracle problem is, affect the criteria used to determine whether to adopt the Uni- or Bi-directional Pattern. Among these, the Bi-directional Pattern adopted when obtaining data processed by smart contracts is assumed to be more important than the oracle problem.

Analysis B: The DLT-type analysis revealed that many systems use private/consortium blockchains throughout all phases. There are two possible reasons for the adoption of private/consortium blockchains: one is to restrict data access to certain parties, and the other is due to limitations during the proof of concept. However, it is difficult to determine which of these was the case. Methods for selecting DLT type based on system purpose and other factors are organized by Turk and Klinc (2017), Li et al. (2019). By following those methods, it is expected that more systems will use public blockchain when the technology matures.

Analysis C: Various patterns were confirmed in the unit of construction data imported into DLT, possibly to optimize the operation of the system and the information to be managed; however, no clear trends were observed in this survey.

Analysis D: Model referencing methods also varied. Many of these methods were considered to reduce the operational costs of the DLT environment. Among them, the method of referencing off-chain repositories using IDs and hashes is often used to link assets such as digital art to the blockchain, and is also effective for referencing BIM, which is a large database. The technique of importing only portions or differences of the BIM into the DLT can minimize data redundancy (Xue and Lu, 2020) and also contribute to distribute the information needed by users as transactions in the DLT. The review also detected systems that do not directly reference the BIM. This may be because the data handled by the systems were independent of the structure of the BIM, which is a static database of the building.

Additionally, the number of systems was organized for each target phase in this analysis, but the number of samples handled in this study did not show any difference in trends for each phase.

6. Limitations of This Study and Future Work

In this study, only the minimum number of words was included in the search query in Scopus to investigate the method of linking BIM data and DLT. Therefore, the survey did not include many systems that manage AEC industry data with DLT without using BIM. In the future, to investigate the DLT widely used in the AEC industry, a similar survey that includes words related to specific purposes, such as "facility management," and "smart city," should be conducted.

Additionally, this study was part of the author's doctoral course research and was conducted to examine a specific method of importing DLT into BIM data utilization. The findings will be used to develop specific BIM-DLT systems in future work.

7. Conclusion

In this study, data flow of BIM-DLT system were classified into 2 patterns: Uni-directional Pattern and Bi-directional Pattern. These patterns were observed through ILR process focusing on how data flows were deployed in 49 systems implemented in previous studies. "Uni-directional Pattern" that only moves from the off-chain repository to the DLT environment and "Bi-directional Pattern" that involves movement back from the DLT environment to the off-chain repository. The pattern to use when constructing the system depends on the purpose of system implementation and whether the purpose of automation through smart contracts is more important than the oracle problem. Additionally, in many studies, private/consortium blockchains were selected as the DLT type, either intentionally or to meet implementation limitations. Although model reference methods varied depending on the data handled in the system, many were aimed at reducing operational costs and data redundancy in the DLT environment.

This study is expected to contribute to the investigation of implementation policies for blockchain systems using BIM data.

Acknowledgments
We would like to thank Editage (www.editage.com) for English language editing.

References

Berawi, M. A., Zagloel, T. Y., Ariq, M. N. & Sari, M. (2021). A Blockchain-Based BIM Model for the Smart Building Planning Process. Proceedings of the 3rd International Scientific Conference on Innovations in Digital Economy, 253–261. https://doi.org/10.1145/3527049.3527140

Chainlink. (2020). What Is the Blockchain Oracle Problem? Why Can't Blockchains Solve It? Chainlink Blog Accessed: 11 December 2022. https://blog.chain.link/what-is-the-blockchain-oracle-problem/

Figueiredo, K., Hammad, A. W. A., Haddad, A., & Tam, V. W. Y. (2022). Assessing the usability of blockchain for sustainability: Extending key themes to the construction industry. Journal of Cleaner Production, 343, 131047. https://doi.org/10.1016/j.jclepro.2022.131047

Jeong, D. Y. & Kim, H. J. (2022). The Advent of the New Architectural Profession through BIM. Buildings, 12(8), Article 8. https://doi.org/10.3390/buildings12081134

Li, J., Greenwood, D., & Kassem, M. (2019). Blockchain in the built environment and construction industry: A systematic review, conceptual models and practical use cases. Automation in Construction, 102, 288–307. https://doi.org/10.1016/j.autcon.2019.02.005

Li, X., Lu, W., Xue, F., Wu, L., Zhao, R., Lou, J., & Xu, J. (2022). Blockchain-Enabled IoT-BIM Platform for Supply Chain Management in Modular Construction. Journal of Construction Engineering and Management, 148(2), 04021195. https://doi.org/10.1061/(ASCE)CO.1943-7862.0002229

Lubbe, W., Ham-Baloyi, W. ten, & Smit, K. (2020). The integrative literature review as a research method: A demonstration review of research on neurodevelopmental supportive care in preterm infants. Journal of Neonatal Nursing, 26(6), 308–315. https://doi.org/10.1016/j.jnn.2020.04.006

Nugroho, R. P., Zuiderwijk, A., Janssen, M., and de, J. M. (2015). A comparison of national open data policies: Lessons learned. Transforming Government: People, Process and Policy, 9(3), 286–308. https://doi.org/10.1108/TG-03-2014-0008

Turk, Ž., and Klinc, R. (2017). Potentials of Blockchain Technology for Construction Management. Procedia Engineering, 196, 638–645. https://doi.org/10.1016/j.proeng.2017.08.052

Wang, J., Shen, Y., Xiong, X., Wang, X., and Fang, X. (2022). Research on multi-person collaborative design of BIM drawing based on blockchain. Scientific Reports, 12(1), Article 1. https://doi.org/10.1038/s41598-022-20321-5

Xue, F., and Lu, W. (2020). A semantic differential transaction approach to minimizing information redundancy for BIM and blockchain integration. Automation in Construction, 118, 103270. https://doi.org/10.1016/j.autcon.2020.103270

Ye, Z., Yin, Â. M., Tang, L., and Jiang, H. (2018). Cup-of-Water Theory: A Review on the Interaction of BIM, IoT and Blockchain During the Whole Building Lifecycle. ISARC Proceedings, 478–486.

STAIRCASE THROUGHPUT STUDY WITH THE USE OF AGENT-BASED MODELING

MICHAŁ JARZYNA[1]
[1]*Lodz University of Technology, Faculty of Civil Engineering, Architecture and Environmental Engineering, Department of Digital Technologies in Architecture and Urban Planning,*
Aleja Politechniki 6a, 90-924 Lodz, Poland
[1]*michal.jarzyna@dokt.p.lodz.pl, ORCID 0000-0001-8599-5665*

Abstract. Agent-based modelling method was used in the study to simulate the behaviour of the crowd pushing against the "bottleneck" areas in the building. The described research involved a simulation testing of the throughput of the staircase used for the evacuation of a building located on a university campus. The study proposes a proprietary method for investigating throughput in a two-dimensional model using the horizontal compound of the average speeds achieved on the stairs by users, considering their age. The method used makes it relatively simple to study the maximum throughput of a stairwell separately from other factors affecting evacuation performance. Based on the results obtained in simulations, the capacity value in the studied situation was determined and several behaviours related to crowd density and stairway speed were observed. This method can be used to simulate the distribution and number of emergency exits from a building using only the maximum number of evacuated occupants, without modelling all floors of the building.

Keywords. Agent-based Modeling, ABM, Building Evacuation, Staircase Throughput

1. Introduction

Scientists are constantly attempting to conduct more and more precise tests and simulations that are able to reflect reality as accurately as possible. Of course, any computer simulation or test performed in a laboratory is subject to some degree of approximation, and nothing will give as precise results as a test based on data collected in the real world.

The trend towards more and more accurate simulations can also be seen in architecture. Simulations regarding the construction or energy efficiency of buildings are constantly being improved to reflect reality more accurately. However, in addition to physical characteristics, users of buildings are also the subject of interest in terms of precise simulations.

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 695-704.

A building should be intuitive and comfortable to use. It was already Vitruvius who put usability on a par with durability and beauty of a building. Among the features that influence the usability of a building is the safety of the people inside. Fire protection issues determining the dimensions of escape routes are regulated by law, and all distances are determined geometrically, with an indication of the shortest distances to the emergency exits. According to the legal regulations in the country of the article's author, the size of staircases is regulated only in terms of the minimum usable width of the flight and landing and the width determined in proportion to the number of people that can be on the floor where the largest number of people is expected to be present at the same time, assuming a minimum width of 0.6 m per 100 people. No more progressive value for throughput appears. Basing regulations on width can be misleading because it does not take into account many factors related to crowd behaviour so throughput should be considered as another value suitable for legal regulation. Agent-based modeling (ABM) is a research direction that is becoming increasingly important in this regard (Stieler et al., 2022). Evacuation simulation is a useful tool to assess and verify the safety of users. It can be used to predict and analyse emergency situations to evaluate how well a particular procedure works in a building.

In the study described in this paper, an existing stairwell capacity test was conducted for evacuation from a multi-story building by proposing an alternative way to analyze the capacity of the staircase compared to the analysis of three-dimensional models.

Agent-based modeling

Agent-based modeling involves the computational study of the behaviour of independent entities (agents), which can be either individual objects or specific collectives with the ability to make decisions within imposed rules and with awareness of their environment.

In the case of ABM used in architecture, agents can be identified with people who use the building (Lee et al., 2014). Individuals may be attributed with the specific actions that mimic human behaviour, such as: moving through a building in a certain direction and at a certain speed, standing in line, or getting things done at a counter. Because each pedestrian can have their own set of behaviours, modeling different situations is much easier than with other crowd modeling approaches such as: fluid dynamics or game theories (Zheng et al., 2009). Derksen (Derksen et al., 2012) listed the following advantages of ABM over other modeling techniques: flexibility, capture of emergent phenomena, and ease of finding analogues in natural environments. At the social level of behaviour, evacuation can show behaviours such as competition, queuing, or herding (Pan et al., 2007). Thus, by interacting both with each other and with the environment, agents can show certain phenomena that can otherwise only be observed through empirical investigation on real life situations. For example, a study of the length of a queue in a store depending on the number of open cash desks can be carried out already at the design stage providing a portion of valuable information. However, if one considers simulations of dangerous situations, e.g.: those associated with fire, then agent-based modeling has a major advantage over experimental testing in terms of safety for potential users of the study and ease of execution. Unsafe crowd behaviours in emergency situations include pushing, hitting, and trampling others. Studying such behaviours in emergency scenarios is difficult because it usually requires exposing real people to potentially dangerous conditions. Especially in situations where users may have mobility restrictions, such as nursing homes or hospitals (Haghpanah et al., 2021; Yuyang et al., 2018). An intelligent computational tool that accounts for human and social crowd behaviour becomes a viable alternative here (Feng et al., 2020; Pan et al., 2007).

Agent-based modeling can be used to confirm already observed trends or to explore opportunities to improve functioning processes giving quick answers regarding possible room configurations and equipment arrangement, e.g. influence of the right desk placement or the presence of two evacuation doors in a classroom (Delcea et al., 2020). In one study, ABM was used to simulate a subway system in the event of a tunnel fire and to examine the consequences of individual actions and their interactions on overall system performance. The simulation proved useful in the search for effective rescue plans (Zarboutis et al., 2004). In high-rise buildings, stairwells are often the only escape route and, being a bottleneck, have a significant impact on the ability to quickly exit the building (Bao et al., 2021; Ma et al., 2012). To check staircases with agents, 3D models were created (Galea et al., 2008; Jiang et al., 2012; Qu et al., 2014). The study described below proposes a different method based on the use of horizontally projected geometry and velocity vectors onto a plane.

2. Research

The study was conducted using Anylogic 8.7.9 software for creating simulations, including agent-based modeling.

The study was designed to test the throughput of a staircase, i.e., what is the average number of people who can descend the height of one floor using that staircase per minute.

In the study described in this paper, an existing stairwell in a 9-story building located on a college campus was analysed. The staircase has a two-run, return flight of stairs with a landing. The flights of the staircase are 120 cm wide, and the landing is 150 cm deep, which is the minimum specified by national law. There is a 20 cm wide clearance between the flights. Stair treads are 17.5 cm high and 27 cm deep, which gives a stair inclination angle of 33 degrees. There are 10 steps in each flight. The average distance to be covered by a pedestrian walking in the middle of the staircase from the moment of reaching the first step of the upper flight to the moment of descending from the last step of the lower flight is about 716 cm in the horizontal projection.

The drawing entered into the Anylogic program is two-dimensional and is shown on the left side of Figure 1. It was not decided to perform a three-dimensional simulation because in the program, climbing a staircase is equated with climbing a ramp, and the speed of climbing the ramp is determined as the product of the agent's initial speed and a manually entered factor that corrects that speed because of the ramp. For this reason, after reviewing the studies that determine horizontal speed on stairs (Fujiyama et al., 2004) (Ma et al., 2012), it was decided to use a two-dimensional simulation.

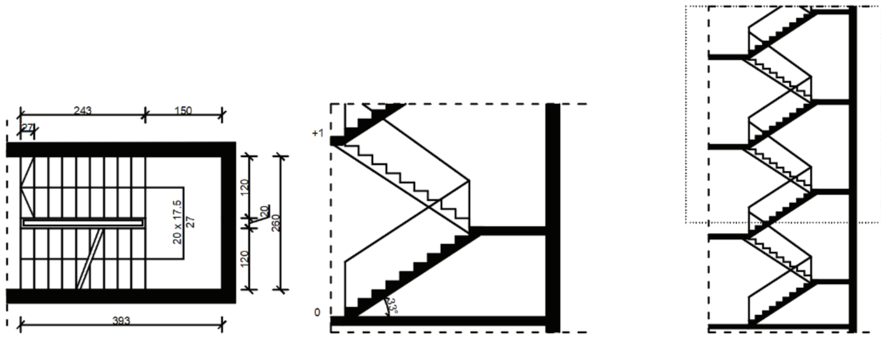


Figure 1. Technical drawing of the staircase. From left: plan at level +1, cross-section through the flight between floors 0 and +1, cross-section through the staircase showing the floors above, which are omitted in the simulation. Dimensions given in cm.

Although the building under study is an existing 9-story building, the two lowest stories are analysed. The analysed higher floors marked in the right part of Figure 1 are not relevant for determining the maximum capacity of the staircase, because it is only a matter of taking the pressure from the largest possible crowd on the section of the staircase between floors +1 and 0. The most critical place of evacuation, in this case, is the staircase, the width of which limits the evacuation speed of a large group of people, because in this place there is competition to increase their own chance to leave the room. Therefore, the simulation assumes the pressure of a large group of people gathered from all floors above level 0, and it does not matter whether they arrive on the staircase from the upper floors or the corridor of the second floor, because they all meet at the entrance to the flight of stairs between level +1 and 0 and are heading toward the 0 floor. Because the study is about evacuation, the entrance to the stairway in the simulation is from a large room filled up with a group of 630 agents within about 3 minutes of the start of the simulation. This is done to create a crowd of people wanting to get to the stairs. Agents have free access to the passage connecting room (A) to the staircase (B), but the number of people who can get to the stairs is limited by the width of the run. All these assumptions are intended to simulate as accurately as possible the potential behaviour of a crowd wishing to overcome the lowest flights of the staircase in order to leave the building as quickly as possible. A study (Fujiyama et al., 2004) was used to determine the speed at which the user travels this distance, in which the speed of descending a stair was examined as a function of direction, speed, stair incline, and age of the user. In the cited study, speed was tested for single trials, so without considering crowding. In another study testing the effective speed in an evacuated building, it was determined to be approximately 0.62 m/s (Ma et al., 2012). In the simulation carried out, it was assumed that the users of the staircase under study are young people about 30 years old, and because the evacuation scenario is being studied, it was assumed that they move down the stairs quickly. Based on an example of a table from the study (Fujiyama et al., 2004) an average horizontal velocity of 0.98 m/s was assumed.

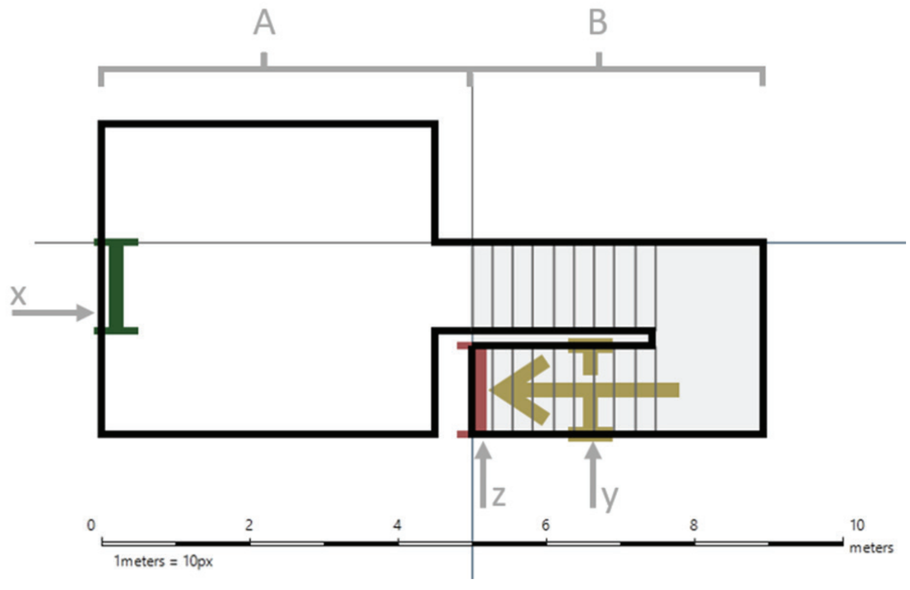


Figure 2. Drawing of the experimental model from Anylogic showing, respectively, the starting room (A), the place where the agents arrive (x), the staircase (B), the place where the number of passing agents is measured (y), the place where the agents move to (z).

In the simulation assuming a single agent, it took 8.1 s to traverse the path, resulting in a horizontal velocity of 0.88 m/s. This coincides with the study of the actual time of one person going down the stairs, when 2 people were tested, obtaining after 3 measurements average times to climb the floor of 7.97 and 8.3 seconds. Any reduction in speed due to crowding or a change in direction of movement will be due to the simulation itself and adjustments are not assumed a priori. As the method of conducting the throughput study is being investigated here, the aforementioned starting data was taken as sufficient.

After entering the staircase area marked B in Figure 2, the agents move with a maximum horizontal velocity of 0.98 m/s and have to cover the first part of the distance in one direction, then have to turn around and go down the second flight of stairs. Within the second flight, there is a gate (marked with a „y" in the Figure 2) that counts the number of people who have passed that way. The results are counted every minute. At the end of the bottom flight of stairs, the agents complete the passage (the line marked with „z"). Random events related to accidents, e.g.: tripping on the stairs, and related disruptions of the traffic flow are not taken into account in this study. In the situation of a building with more than 2 floors, by traversing one floor, users join the crowd of people wanting to leave the floor below and the scenario is repeated, with the throughput still limited by the width of the staircase.

The simulation was run for approximately 14 minutes and this was the time it took for all agents to pass through the stairwell. As a result of the simulation, data was taken to determine the number of people that crossed the control line (y). The value was recorded every minute, and for each record the quantity was counted from the beginning of the simulation. The obtained values are shown in Table 1.

Explanation of the information collected in the table:

- Measurement time - the time counted from the beginning of the trial;
- Number of agents - the number of agents that have passed the measurement point located in the middle of the second flight of stairs - „y";
- Average - the number of agents per second based on the measurement so far;

Measurement time [sec.]	60	120	180	240	300	360	420	480	540	600	660	720	780	840
Number of agents [unit]	41	93	145	198	251	303	359	408	453	493	531	567	607	630
Average [unit. /sec.]	0.68	0.78	0.81	0.83	0.84	0.84	0.85	0.85	0.84	0.82	0.80	0.79	0.78	0.75

Table 1. Graphical representation of the test results showing when the throughput values increase, stabilize and decrease

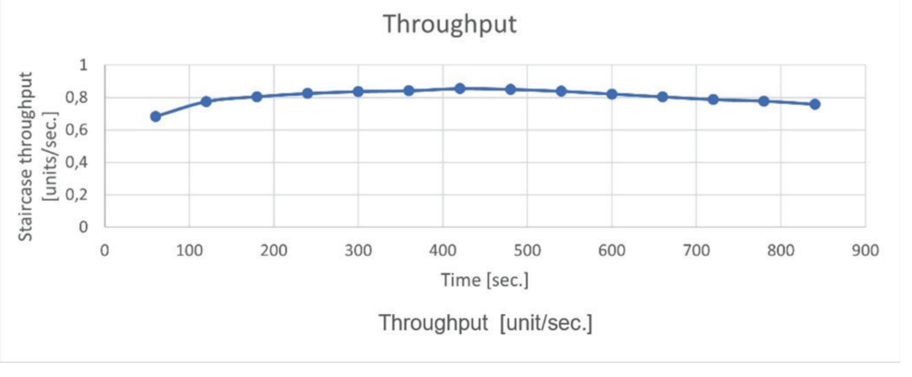


Figure 3. Graphical representation of the test results showing when the throughput values increase, stabilize and decrease

In the graph above (Figure 3) the three stages of throughput formation can be seen. In the first stage, covering the time from the beginning to about 180 seconds, the value gradually increases and this is due to the small initial value when there were few people in the room. Initially, all created agents had free access to the staircase, so they only needed time to traverse room A from the starting point (line x) to the entrance. As time progressed, the room filled up. From the 180th second, the value stabilizes at a result between 0.8 and 0.85 units/sec., and this coincides with the maximum filling of the room. The number of people entering the staircase already depended only on the capacity of the staircase, since a constant supply of agents was provided from the starting room. This value is maintained until about 660 seconds, when the room begins to become empty, and the crowd pressure is no longer so high. A graphical representation of the test for several trials is shown in Figure 4 and the meanings of the data in specific fields of the figure are explained in Table 2.

Measurement time [s]	Number of agents [unit]
Floorplan	

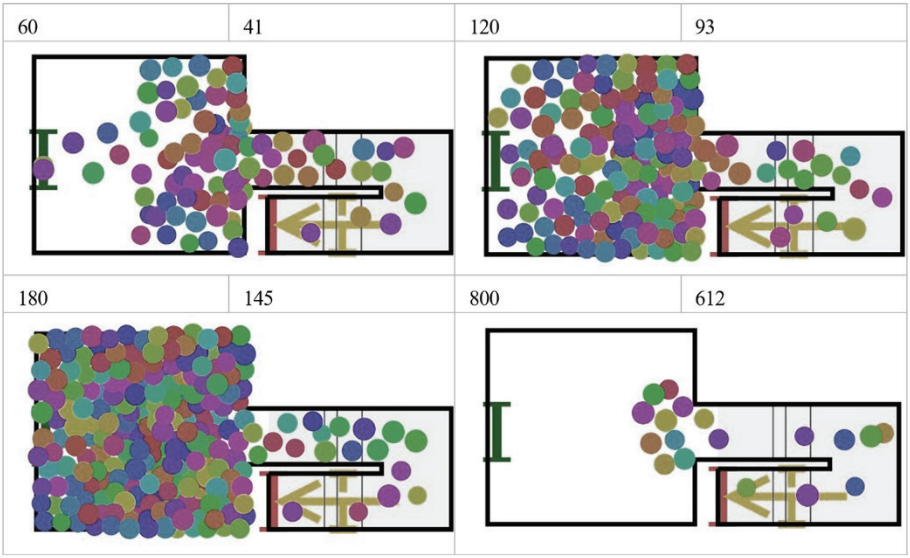


Table 2. Diagram of the description of the tests carried out

Figure 4. Graphical representation of the simulation as the initial room is filled and just before it is emptied

As the room fills with agents, the difference in density of people with people in the stairwell becomes apparent. This is because the crowd pushing against the narrow entrance to the flight of stairs sometimes wedges each other. After slowly squeezing through this constriction, people are able to run down the stairs, which is noticeable in the animation, in the form of a larger radius when turning into the second flight of stairs. It can also be seen in the simulation that individuals moving closer to the centre of the staircase must significantly reduce their speed to ascend the bottom flight, while individuals walking on the outer radius having a longer distance to travel but more room to make the turn often travel faster. In addition, it was assumed that on floor 0, evacuees do not encounter any obstacles and are free to exit the flight of stairs.

Agent-based simulations involve a certain degree of unpredictability, so in order to check the results, two more control trials were also conducted as shown in Table 3 . The differences in specific values compared to test no. 1 did not exceed 3%, with an average difference value of 0.78%.

Measurement time [sec.]	60	120	180	240	300	360	420	480	540	600	660	720	780	840
Number of agents in test no. 2 [unit]	41	93	148	198	254	304	352	402	444	487	529	564	602	630
Number of agents in test no. 3 [unit]	41	93	142	197	246	297	349	398	442	489	527	564	603	630

Table 3. Results of the test no. 2 and 3

In test no. 4, the maximum allowed speed on the staircase was reduced by 20% from 0.98 m/s to 0.78 m/s. This resulted in a decrease in stairwell throughput, but the decrease was at most 3%, with an average value of 2% (Table 4).

Measurement time [sec.]	60	120	180	240	300	360	420	480	540	600	660	720	780	840
Number of agents [unit]	40	92	143	195	246	296	348	396	441	483	522	558	596	624
Decrease in throughput compared to test no. 1	2%	1%	1%	2%	2%	2%	3%	3%	3%	2%	2%	2%	2%	1%

Table 4. Results of the test no. 4

By subjecting individual agents to observations from the crowd, an average travel time of 24.6 s was obtained, which was more than 3 times longer than on an empty staircase. The effective speed was 0.29 m/s.

3. Summary

This paper presents a study on stairwell throughput conducted using agent-based modeling software. The paper suggests a specific method for analysing throughput in a two-dimensional model utilizing the horizontal compound of the users' age-adjusted average stair pace. The study was conducted assuming the reality of evacuation from a building. The throughput determined by the agent-based modeling method can be an important factor determining the size of a building's circulation paths. The following study proposes a method that simplifies the implementation of the analysis expanding the range of potential users. As a result of the simulation, it has been determined that, maintaining the characteristics of the agents described in paragraph 3, the throughput of the investigated staircase is approximately 0.83 people per second. The throughput value is taken as the average value of the middle evacuation period, when the pressure of people is relatively constant as opposed to the initial period, when the bottleneck has not yet been fully sealed, and as opposed to the last period, when the crowd begins to unload. The influence on the possibility of evacuation of people has the width of the staircase run, as the procedural shape of the staircase decreases the average speed at which an agent moves down the stairs. At the same time, the crowd decreases the speed of individuals, because it slows down the movement of pedestrians several times. Although agent-based simulation is characterized by a certain degree of variability, the differences between three consecutive tests were maximum of 3%. this study can be a helpful tool during building design, as it is relatively easy to simulate the throughput of a given staircase, which is affected by factors such as the width and number of stair flights. To perform the study, not much information is required, so the study can be carried out at an early stage of the project. Knowing the capacity of the lowest segment of the staircase and the number of potential users located on all upper floors of the building, it is possible to roughly estimate the amount of time required for people to leave the building. Such information at an early design stage can influence the layout and number of vertical passageways.

4. Further research

The present study focused on the actual staircase in the studied building. Further research can focus on trying to find the correlation in staircase width and identify the optimal values based on agent modeling. Another aspect worth investigating may be the effect of the dimensions of the stair steps on the throughput of the staircase.

References

Bao, Y., & Huo, F. (2021). An agent-based model for staircase evacuation considering agent's rotational behavior. Physica A: Statistical Mechanics and Its Applications, 572, 125923. https://doi.org/10.1016/j.physa.2021.125923

Delcea, C., Cotfas, L. A., Craciun, L., & Molanescu, A. G. (2020). An agent-based modeling approach to collaborative classrooms evacuation process. Safety Science, 121(August 2019), 414–429. https://doi.org/10.1016/j.ssci.2019.09.026

Derksen, C., Branki, C., & Unland, R. (2012). A Framework for Agent-Based Simulations of Hybrid Energy Infrastructures. 1293–1299.

Feng, X., Gao, M., Zhou, Y., & Li, J. (2020). Simulation and optimization of emergency evacuation in gold museum based on AnyLogic. Journal of Physics: Conference Series, 1650(3). https://doi.org/10.1088/1742-6596/1650/3/032080

Fujiyama, T., & Tyler, N. (2004). An explicit study on walking speeds of pedestrians on stairs. 10th International Conference on Mobility and Transport for Elderly and Disabled People, Hamamatsu, Japan, May 2004, January, 1–10. Retrieved from http://discovery.ucl.ac.uk/1243/

Galea, E. R., Sharp, G., & Lawrence, P. J. (2008). Investigating the representation of merging behavior at the floor-stair interface in computer simulations of multi-floor building evacuations. Journal of Fire Protection Engineering, 18(4), 291–316. https://doi.org/10.1177/1042391508095092

Haghpanah, F., Ghobadi, K., & Schafer, B. W. (2021). Multi-hazard hospital evacuation planning during disease outbreaks using agent-based modeling. International Journal of Disaster Risk Reduction : IJDRR, 66, 102632. https://doi.org/10.1016/j.ijdrr.2021.102632

Jiang, Z., & Liu, X. (2012). Numerical simulation study on high-rise student apartment fire evacuation. Advanced Materials Research, 433–440, 3011–3016. https://doi.org/10.4028/www.scientific.net/AMR.433-440.3011

Lee, Y. S., & Malkawi, A. M. (2014). Simulating multiple occupant behaviors in buildings: An agent-based modeling approach. Energy and Buildings, 69, 407–416. https://doi.org/10.1016/j.enbuild.2013.11.020

Ma, J., Song, W. G., Tian, W., Lo, S. M., & Liao, G. X. (2012). Experimental study on an ultra high-rise building evacuation in China. Safety Science, 50(8), 1665–1674. https://doi.org/10.1016/j.ssci.2011.12.018

Pan, X., Han, C. S., Dauber, K., & Law, K. H. (2007). A multi-agent based framework for the simulation of human and social behaviors during emergency evacuations. AI and Society, 22(2), 113–132. https://doi.org/10.1007/s00146-007-0126-1

Qu, Y., Gao, Z., Xiao, Y., & Li, X. (2014). Modeling the pedestrian's movement and simulating evacuation dynamics on stairs. Safety Science, 70, 189–201. https://doi.org/10.1016/j.ssci.2014.05.016

Yuyang, J., & Hongyan, M. (2018). Study on evacuation simulation of medical pension building based on anylogic. Proceedings of the 30th Chinese Control and Decision Conference, CCDC 2018, 00331615010, 2307–2312. https://doi.org/10.1109/CCDC.2018.8407511

Zarboutis, N., & Marmaras, N. (2004). Searching efficient plans for emergency rescue through simulation: the case of a metro fire. Cognition, Technology & Work, 6(2), 117–126. v10.1007/s10111-004-0150-6

Zheng, X., Zhong, T., & Liu, M. (2009). Modeling crowd evacuation of a building based on seven methodological approaches. Building and Environment, 44(3), 437–445. https://doi.org/10.1016/j.buildenv.2008.04.002

CONJUGATED MATERIALITY

Reinstating Material Circularity via Digital Twins

NIPUN GARG [1] and SHENG-YANG HUANG [2]
[1,2] *Bartlett School of Architecture, University College London*
[2] *Architectural Association School of Architecture*
[1] *nipun.garg.21@ucl.ac.uk, 0000-0002-2658-3174*
[2] *ucfnhua@ucl.ac.uk, 0000-0002-2748-5031*

Abstract. Industrial Revolution 4.0 offers an opportunity for the globe to rethink the meaning of building information that breaks the territorial borders of building information systems that are not based project-wise but follow a geopolitical structure. It expands the conventional thought process of being limited to a building to a city/ planetary urbanisation level. As a response to the new urban design theory, the paper posits an approach that amalgamates “Design for Disassembly (DFD)” and “Digital Twins” which have gained traction because of “Circular Economy” and “Industrial Revolution 4.0” respectively, to create an information framework for the urban ecology that focuses on system management rather than project management via “Material Passport (MP) 2.0”. It identifies the gaps within the existing MP and creates a foundational framework for the added information (termed “Material Strategies”) that needs to be a part of MP 2.0 that arise while working across systems by augmenting DFD and Digital Twins via the lens of materials. The material strategies are further investigated through a correlation matrix to understand their interdependency to finally create a JavaScript Object Notation (JSON)-based serialisation of materials to reinstate the material circularity and reduce the carbon emissions that the construction sector accounts for.

Keywords. Design for Disassembly (DFD), Digital Twin, Material Passport (MP), Circular Economy, JavaScript Object Notation (JSON)

1. Introduction

The paper, by positioning itself amidst the "New Materiality" (considering materials as active entities and believing in systems rather than symbols or objects that is line with DFD) and "Cybernetics" (concerned with circular causality or regular feedback, in line with Digital Twins) theories, seeks to create an information framework for the environment based on cross machine interaction and machine intelligence where the role of human becomes secondary. This information framework is primarily based on

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 705-714.

“Material Strategies”, that the paper generates, and which form the foundation of MP 2.0. The paper further highlights their need in order to break from the transcendental notion of working within a single system to work across systems. It then understands their interdependency via matrix analysis and further depicts where these strategies and the parameters associated with them are interjected while working across systems. To cater to the complexity established during this process, the paper attempts at creating a material serialisation that is based on JSON (a computationally sustainable language) for cataloguing and tracking the complex materiality of the building processes. It envisions at reinstating the material circularity within a derailed construction industry by creating a "Digital Material Passport" that uses the digital twin technology to help in reducing carbon emissions at a planetary level.

As the world is entering into an era where systems are run and monitored by machines known as the Cyber-physical systems (CPS), it offers a possibility to remodel the construction industry that considers city and the planet as a mega building project rather than just limiting to a building. The industry's constant emphasis on the recognition and creation of "symbols" (as suggested by the old dialectical materialism by Marx and Hegel) i.e., the overall meaning and the outlook of the building, without giving much importance to the underlying "process" such as the interactions between various materials involved within the same building have rendered the industry less sustainable and linear (consuming 40% and wasting 30% of resources). The result is that materials either end up being downcycled or in landfills when paving way for newer construction thus following the loss of embodied carbon (releasing nearly 11% of the world's carbon into the atmosphere). The need therefore is to shift from the notion of project management (passive perception of materials) to system management (active perception of materials) which "New Materiality" (synonymous with the theory of “New Materialism”) by Manuel DeLanda based on the work of Deleuze and Guattari argues and is addressed by the concepts of DFD and Digital twins to assist in achieving the goal of a "Global Circular Economy".

2. Background

New Materiality brought forward the concept of "active materiality" that categorises systems not only by properties but also by their capacities/ tendencies which can be real without being actual, ontologically defined as "virtual", requiring a catalyst to bring them to actuality (DeLanda, 2015). DFD, which focuses on formation rather than form, treating materials as active entities that are dynamic rather than objects that are constant, within and across systems and emphasises material information as MPs rather than the structure, aligns with this theory but however addresses only a part of it. With the advent of CPS, Digital Twins which has garnered traction as a part of Industrial Revolution 4.0, offers the possibility of perceiving and bringing the virtual tendencies that exist within a material (what was, what is and what can be) to actuality by acting as a catalyst. Industrial Revolution 3.0 revolutionised the physical space by introducing a parallel virtual space in the 20th century with the introduction of computers, the internet, wireless networks, and simulation tools that not only enabled remote cooperation between the virtualized physical assets but also plans and operations more effective and efficient (Tao and Zhang, 2017). Digital twins unveil this shroud of virtual intelligence that is associated with each object and can blur the boundary between the virtual and physical which until now were considered as two separate faces of the same coin. The aim of a digital twin is to create high-fidelity virtual models to imitate the state and the behaviours of each of the physical objects with the ability to evaluate, optimise and predict (Graessler and Pöhler, 2017), which is the reason feedback and information are its essential components to guide physical entities to their optimal states. This description is however analogous with the theory of Cybernetics that recognises the world based on information and control by means of "circular causality" or in simple terms feedback. By playing out multiple scenarios based on constant feedback and avoiding the unintended consequences that are often brought with the well-intentioned changes, digital twins provide a testing bed of experimentation without having to make changes in the real life that can reduce greenhouse gas emissions and cost savings by 50% and 35% respectively according to research carried out by the professional services company EY (Ernst and Young).

Industrial Revolution 4.0 which highlights cross-communication across machines, needs a regular information exchange across the web making good data interoperability necessary. XML (Extensible Markup Language) and JSON are currently the most common data formats that are used for this cybernetic purpose which the paper acknowledges. BIM (Building Information Modelling), the foundation of Digital Twin, uses XML for this purpose. XML is the most widely used format across programmes, people, people and machines for sharing structured information i.e., data, documents, books, transactions, invoices, etc. Industrial Revolution 4.0 which is heavily reliant on data requires significant storage and parsing space. The more complex the data, the more data centres it requires and the more carbon emissions it produces. According to studies, JSON (whose syntax is both human and machine-readable) has proven to be computationally sustainable requiring lesser data storage space when compared to XML which can significantly reduce carbon emissions. The paper, therefore, suggests JSON as a better alternative for information exchange in Industrial Revolution 4.0 which is also the reason why Table 1 shows examples that are JSON formatted BIM data.

Name	Purpose	Focus
vA3C	Plugin exported BIM data from SketchUp, 3DS Max, Revit, Grasshopper, etc. is converted to JSON using a browser-based open-source BIM model viewer vA3C that uses a customised JSON schema definition which focuses exclusively on geometry data while producing documents (vA3C, 2022).	Objects
Autodesk Forge Viewer	Employing its own unique JSON schema definition, it translates 2D and 3D models of more than 70 file formats, including AutoCAD, Fusion360, Revit, etc., to JSON (Forge, 2022).	Objects

GeoJSON	It supports several geometry data types like LineString, MultiString, Point, MultiPoint, Polygon, and MultiPolygon even though its primary use is to serialise geographic data. The only drawback being the term "properties" which is overly general when it comes to complex geometry that does not address building data component specification.	Geographic Information (GIS)
ArchiJSON	It is a protocol for exchanging parameters and data for architectural design that is based on JSON (PyPI, 2021).	Objects
BIMJSON	It is used for exchanging BIM/ MVBs (Minimal Viable BIMs) with the primary goal to design a simple framework that is easy to grasp and has a lot of flexibility for data transfers.	Objects

Table 1. Past attempts

3. Formulation of Material Passport 2.0 (MP 2.0)

3.1. CREATING MATERIAL STRATEGIES – REASSESSING DFD

We start by rethinking DFD approach from the lens of materials which focuses on materials from the outset. DFD strategy "Materials" which suggests the usage of pure, non-toxic, and resilient materials that can sustain multiple life cycles is translated into "Material Selection". The strategy "Service Life" suggests flexibility, adaptability, and temporary structures for easy extraction of materials in translated into "Material Interaction". The "Deconstruction" strategy, which suggests the formulation of a deconstruction plan, environment, and material management is translated into "Material Movement". The strategy "Connections" which suggests visible, accessible, avoiding/ providing dissolvable binds, and mechanical connections to reduce time and minimal damage to materials are translated to "Material Correlation". The last strategy "Standards" which suggests using modular, easily replaceable systems, subdividing a complex structure into components, and using a prefabrication process for assembly/ disassembly to be quick and safe is translated into "Material Interaction". A depiction of the DFD strategies considered according to BAMB and their translation into Material strategies is done in Figure 1.

This is however analogous with the theory of New Materiality which suggests moving away from the transcendental view of the matter to viewing and understanding them in terms of interactions, flows, and processes. A stepwise strategy is established i.e., selected materials for a system are checked for compatibility and then correlated to interact with each other. These correlated junctions can be disassembled to obtain the interacted materials separately which can then be moved and modified to interact with other systems. To be scaled across systems, and reduce time and carbon emissions, many parameters arise when materials from one system are disassembled and modified to be a part of another system. The parameters associated with each of these strategies are established using a correlation matrix in Figure 2.

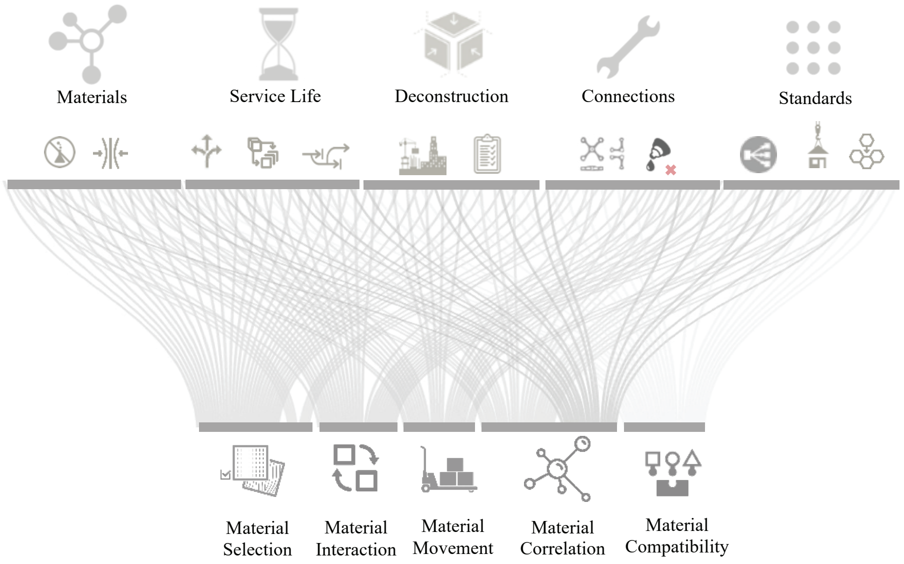


Figure 1. Author’s interpretation of translation of DFD strategies to Material Strategies

Material information as MP is important for creating DFD designs which are catalogues that keep traces of a product's, material's, or system's activities from manufacturing, acquiring, and usage to upkeep for their potential for recovery and make that information easily accessible to the necessary parties throughout their value chain (Luscuere, 2017) for which they require to be updated regularly for their use in subsequent lifecycles. However, the current MPs are addressed only in terms of interaction within a single system while there are several parameters that arise as a result of interactions across systems that can significantly reduce the time taken during the disassembling process and reduce the carbon emissions that are associated with the disassembling process as shown in Figure 3. The correlation matrix in Figure 2, thus helps in creating the foundational framework for the information that needs to be a part of MP 2.0 which could give an added dimension to the existing MP.

The reason for understanding the intensity of materials using the correlation matrix is to reduce the time and carbon emissions when materials are disassembled from one system and are modified to be a part of another system. Figure 3 highlights the complexity involved in this process across two systems which when scaled can increase manifold, which is why the establishment of the correlation matrix is necessary. Another factor that this augmentation has highlighted is the need for the integration of geographical information (GIS) in the current model for tracking the changes that happen to a material. Parameters that are established such as the location of the storage of the extracted materials, time taken, etc. that are related to the strategy of material movement essentially require access to this information. This also highlights why the attempts as highlighted in Table 1 do not work for Industrial Revolution 4.0 as they focus either on objects or geographical information rather their integration.

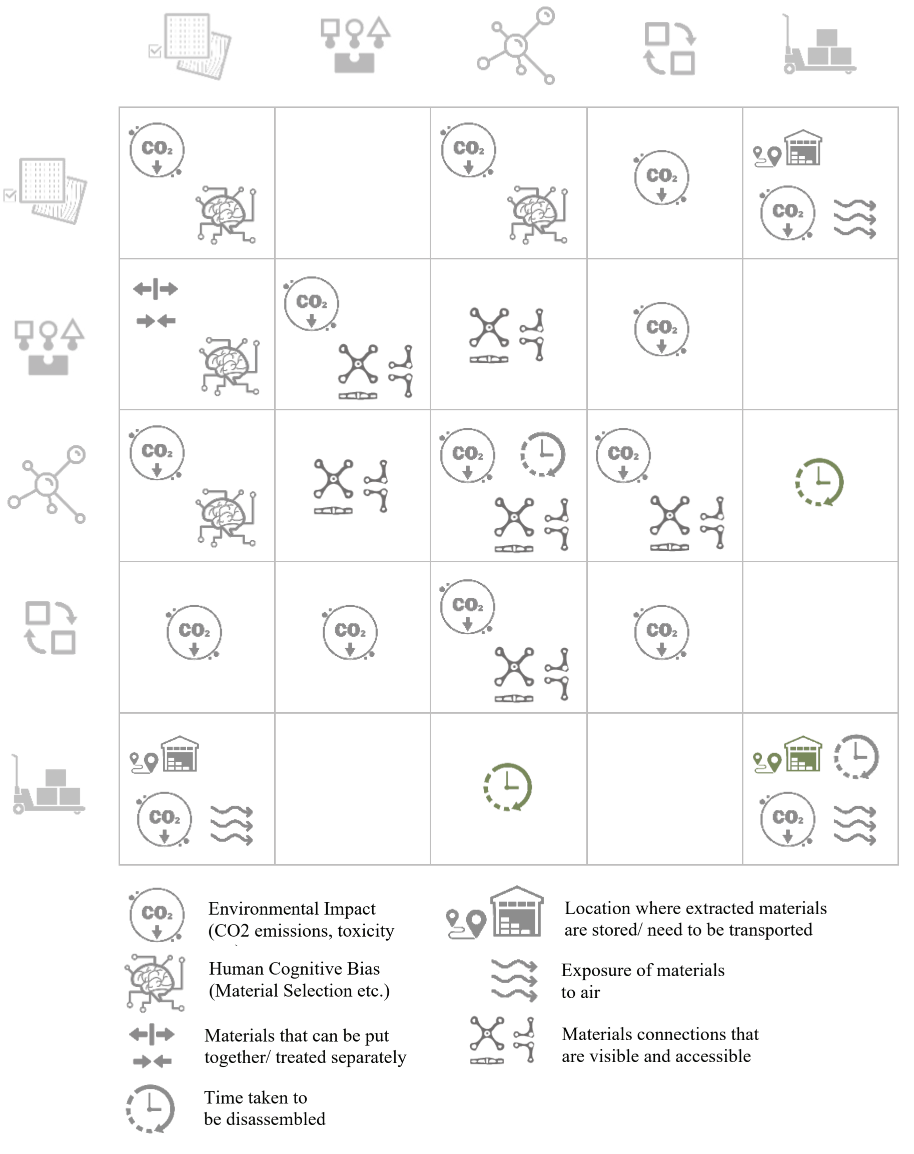


Figure 2. Correlation matrix between Material Strategies and the parameters associated with them

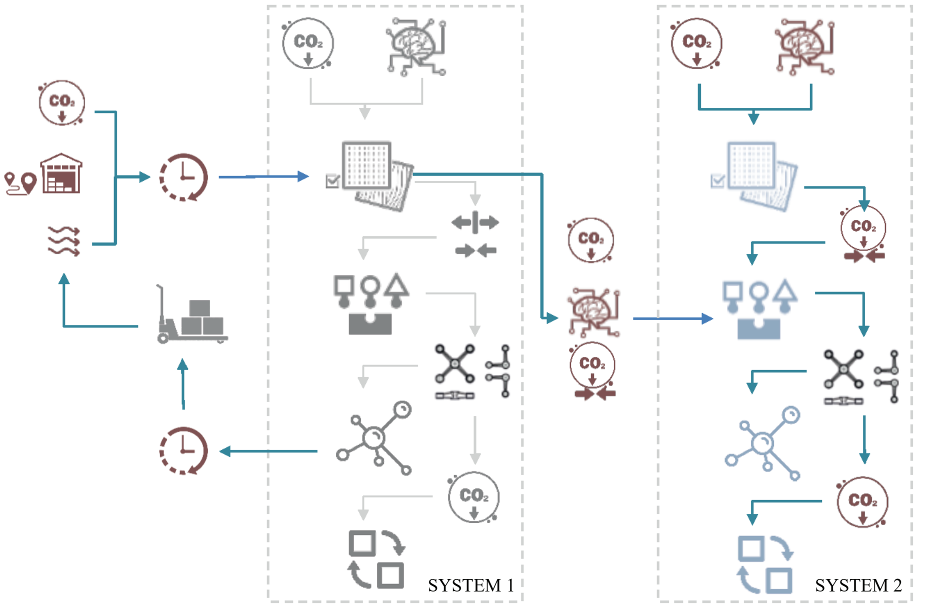


Figure 3. Steps and processes involved during the assembly and disassembly of a system highlighting the material interactions across systems

3.2. ANALYSING THE DIGITAL EXPRESSION OF MP 2.0

In the scenario of MP 2.0, the real-time updating would require information exchange via the web which further would need a comprehensive and integrated system for the exchange of information to be effective. Communication, cooperation, and exchanging data for achieving a common goal are the primary objectives of any system integration with the first and the most crucial step being data interoperability. In the case of BIM which is the base for the creation of Digital Twins for materials for the MP 2.0, this happens with the help of IFC (Industry Foundation Class) which has also been acknowledged as the de facto open BIM standard by more than 20 vendors. For the case of BIM, presently the information is exchanged in the form of BIM objects using the XML format. As stated earlier, JSON has proven to be computationally inexpensive and seems like a better alternative, especially after establishing the complexity that is involved in the creation of MP 2.0. The research envisions a transformed industry where each building/ urban setting has a digital twin at a macro level that can be carefully perceived as a composition of several material twins at a micro level that can catalogue and track the complex materiality of the building processes to rethink the material circularity and can complement DFD and MPs as discussed in this section via JSON (for cross-communication across machines).

3.2.1. JSON/ XML – The Difference that Matters

Although today, XML is one of the most widely used formats for information exchange, JSON proves to be a better alternative. The reason being XML was primarily designed to represent structured documents rather than structured data. The data structure can be visualised as trees that are exchanged across the web, out of which the relevant objects/ elements are extracted. The constant shifting between the tree (XML representation of data) and the object model (internal representation of data in applications) resulted to be more time-consuming. JSON on the other hand was developed simply by reusing the object model of an existing language (in this case JavaScript) that eliminated the need of traversing and parsing through the XML trees and allowed native data structures to directly map the JSON objects.

3.3. DESIGNING THE MP 2.0 SCHEMA

In order to create a global network of materials for cross-communication across various systems via machines, MP 2.0 is a framework of information for the environment that integrates BIM and GIS to create a virtual twin of materials in order to further enhance and complement the "Active Materiality" that DFD has brought to the fore. It is established on the basis of the findings in the above sections so as to reduce global carbon emissions and assist in achieving the goal of the global circular economy. Figure 4 gives a glimpse of how this cross-machine exchange of information might look like across systems.

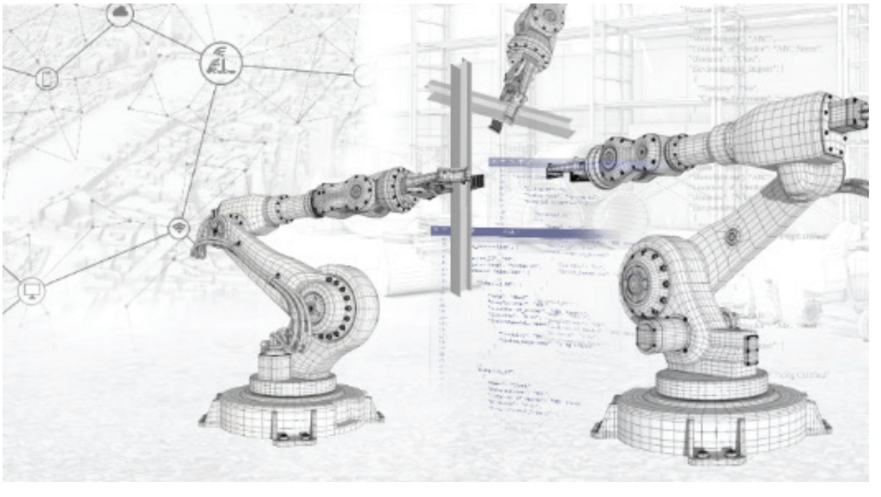


Figure 4. Glimpse of how the new MP 2.0 works across systems

A prototype JSON file has been validated using the JSON validator and formatter (Validator, 2022) while the JSON Schema is validated using the JSON Schema Validator (Newtonsoft, 2022) The entire file can be viewed on the following link – https://github.com/ConjugatedMateriality/MP_2.0. A graphical representation of the JSON design is depicted in Figure 5.

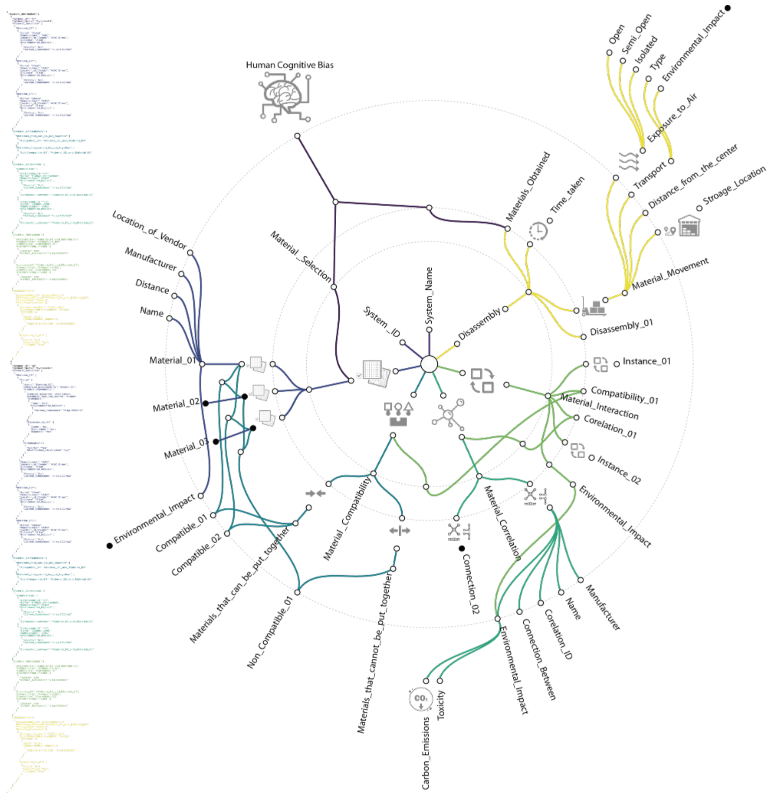


Figure 5. A visual representation of the JSON designed MP 2.0 (Note: Full resolution image can be viewed on the github link)

4. Conclusion

Increased carbon emissions and the potential planetary trauma that is looming as a result of the Anthropocene, have urged the globe to push all sectors, especially the building sector which accounts for more than a third of the world's carbon emissions, to find new approaches that help in attaining "Global Circular Economy". DFD, which focuses primarily on material information in the form of MP, is the recent approach that has been garnering traction to bring this vision to reality. The reason is a notional shift from project to system management and treatment of materials as active entities which is synonymous with the “New Materiality” theory. The same theory which suggests the requirement of catalysts for bringing the virtual tendencies/ capacities associated with the physical material to reality brings the focus to Digital Twins, a term that has been doing the rounds as part of the Industrial Revolution 4.0. The research, by taking cues from the already established theories and approaches, proposes MP 2.0 which augments the DFD strategy and integrates it with the Digital Twin technology to solve a part of the "Global Circular Economy" puzzle. “Material Strategies” by translating DFD strategies are established and the parameters associated with them while working across systems are depicted via a correlation matrix in the paper. A case scenario between two systems is shown to depict the complexity that arises when these parameters are interjected. Industrial Revolution 4.0 which is marked by CPS comes to aid this complexity. It requires cross-communication across machines and since the role of humans becomes secondary, data interoperability becomes the most crucial step for exchanging information across the systems. The research suggests JSON as a better alternative in comparison to XML for this purpose. It has proven to be computationally sustainable in handling and parsing complex data. In the process of establishing the parameters associated with "Material strategies", the paper also highlights why the past attempts that either focus on BIM (the backbone of Digital twins) or GIS need integration. The research attempts to create a JSON prototype addressing all these findings. MP 2.0 which is ever evolving with new materials and information, not only the material circularity is restored but also the carbon emissions are reduced that are involved as a part of this exchange of materials across systems. It is an information and knowledge framework for urban ecology that extends the capability of an existing MP i.e., from a building to consider the earth as a mega building project. It envisions a transformed construction industry where the eternal essence of materials coexists within the transient urban ecosystem something which closely resembles the principle of nature where one system's trash becomes feed for another.

References

DeLanda, M. (2015). The new materiality. Architectural design, 85(5), 16-21, https://doi.org/10.1002/ad.1948.

Forge, A. (2022). AutodeskForge.com. Retrieved 25/07/22 from https://forge.autodesk.com/

Graessler, I., and Pöhler, A. (2017). Integration of a digital twin as human representation in a scheduling procedure of a cyber-physical production system. 2017 IEEE international conference on industrial engineering and engineering management (IEEM), doi: 10.1109/IEEM.2017.8289898.

Grieves, M. (2016). Origins of the Digital Twin Concept. https://doi.org/10.13140/RG.2.2.26367.61609

Herbert, G. (1978). Pioneers of prefabrication: the British contribution in the nineteenth century (Vol. 19). Johns Hopkins University Press.

Luscuere, L. M. (2017). Materials Passports: Optimising value recovery from materials. Proceedings of the Institution of Civil Engineers-Waste and Resource Management, DOI:10.1680/jwarm.16.00016.

Merrild, H., Jensen, K. G., and Sommer, J. (2016). Building a circular future. GXN.

Newtonsoft, J. S. V.-. (2022). Jsonschemavalidator.net. Retrieved 25/07/2022 from https://www.jsonschemavalidator.net/

PyPI. (2021). Archijson. https://pypi.org/project/archijson/

Rifkin, J. (2011). The third industrial revolution: how lateral power is transforming energy, the economy, and the world. Macmillan.

Tao, F., and Zhang, M. (2017). Digital twin shop-floor: a new shop-floor paradigm towards smart manufacturing. Ieee Access, 5, 20418-20427.

vA3C. (2022). Va3c.github.io. Retrieved 25/07/2022 from https://va3c.github.io/

Validator, J. F. a. J. (2022). Jsonformatter.org. Retrieved 25/07/22 from https://jsonformatter.org/

Digital Heritage, Representation and Visualisation

EXPLORING THE APPLICATION OF THE DIGITAL GAMIFICATION MECHANISMS TO THE EXPERIENCE OF PHYSICAL ARCHITECTURAL EXHIBITIONS

WEIQIONG LI[1], TIANTIAN LO[2], and XIANGMIN GUO[3]
[1,3]*Harbin Institute of Technology(Shenzhen)*
[2]*The Hong Kong Polytechnic University*
[1]*weiqiongli221@gmail.com,*
[2]*skyduo@gmail.com,*
[3]*24904404@qq.com*

Abstract. This study aims to respond to the 'human-centred' theme of digital heritage and visualisation by exploring a new approach to applying gamification mechanisms to design physical architectural exhibitions. This paper analyses the current exhibition's gamification design in three parts-core drivers, defining characteristics and development models. Then constructs a design model for "digital gamification". The history museum of Harbin Institute of Technology (Shenzhen) is selected as an example to conduct an empirical investigation. Finally, future experiments are proposed to evaluate the design process's effects on improving the platform's design. It is expected that the demonstration of this study will enrich the exploration of the application of the emerging design method of digital gamification mechanism in exhibition design. On the one hand, it attempts to construct the relationship between the influence of digital gamification mechanisms on the tangible and intangible information in the audience's cognitive space, thus providing new ideas for designing cultural experiences in future exhibition spaces. On the other hand, it gives new vitality to the exhibition design and enhances the audience's motivation to interact, which helps to expand cultural communication's influence.

Keywords. Exhibition Space, Experiential Mechanism, Digital Interaction, Gamification, Extended Reality

1. Introduction

In the new era, the transformation of the function of exhibition space has facilitated the exploration of innovative ideas for its design. Traditionally, the primary role of exhibition space was to collect and research and the collection is the centre of its attention. As a result, the spatial planning and design often differentiate and cluster exhibits according to their country, theme, age and other attributes. Spatial elements such as lighting and scale serve to give a better picture of the exhibits in their entirety. Visitors are also encouraged to follow the design routes to view the exhibits. However,

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 717-726.

as the social roles and functions of exhibition space have gradually diversified, maximising social value has become a primary objective in the design of exhibition spaces.

Recent research trends have found that entertainment and games can support the learning and educational mission of exhibition space as a place of informal learning, enhancing the cognitive efficiency and experience of people in the space. On the one hand, digital interactive technologies are used to enrich the presentation of cultural information and increase the sensory engagement of the audience. On the other hand, playful elements stimulate the audience's behavioural motivation and promote spontaneous learning and exploration based on personal interests in the spatial experience. However, gamification mechanisms are often combined with virtual technologies for building online platforms for displaying interactive experiences, with less research on the combination and application of physical environments.

For this reason, this paper uses digital technology and gamification as tools to explore a new exhibition experience design model for visitors to interact with space offline. Firstly, based on the literature study, we analyse the gamification design in current exhibition design applications regarding core drivers, defining characteristics and development models. Combine extended reality technology and gamification mechanisms to build a new design method to explore interactively in natural spaces. The design model of "digital gamification" was developed. Based on the design model, an empirical study of the "Objective-Analysis-Design" process was completed by taking the School History Hall of Harbin Institute of Technology (Shenzhen) as an example. The digital gamification design guides and stimulates audience interaction and continuous participation, thus enhancing the audience experience and promoting their participation and sharing (Figure 1).

Figure 1. Digital gamification mechanisms to enhance the exhibition experience

2. Overview of gamification's concept and its application

The definition of 'gamification' varies somewhat between scholars. However, Deterding provides a more widely accepted concept of gamification as a mechanism for using game design elements in non-game contexts to improve user experience and engagement(Deterding, S., 2011).

Scholars cite different grounded theories from cognitive and social psychology to reveal the effects of gamification, with self-determination theory (SDT) and mind-flow theory being the most commonly used theories. Self-determination theory is commonly used in gamification to derive the effects of game design (Wee, S. C., 2019) . It analyses human behaviour in terms of its intrinsic and extrinsic motivational components. Although the positive effects of extrinsic motivation are short-term, they can be combined with, or even transformed into, intrinsic motivation when fully internalized or conscious(Ryan, R. M., 2020) and thus enhance the user's perception of experience and behavioural autonomy more sustainably. Flow is a psychological concept that means "the experience of forgetfulness that occurs when people complete a difficult task that matches their skills" (Csikszentmihalyi, M., 1991). When there is a balance between personal skill and challenge, people enter a state of mind flow where they are highly focused, and their behaviour and consciousness are integrated. Clear goals and immediate feedback in the game elements can positively affect the maintenance of this state.

The concept of gamification has formed a complete theoretical research framework in terms of validity, game elements, intrinsic motivation and impact assessment. It is currently focused on specific application strategies and design frameworks. In the exhibition design, specific applications of gamification mechanisms include electronic guided tours, digital interaction of cultural heritage, virtual exhibition galleries, awareness-raising and education, and cultural exchange. Interactions include fully online virtual game interactions, virtual interactions based on extended reality technologies (VR/AR/MR) and fully offline physical prop interactions. In terms of interactivity, the amount of information presented, and the enhanced perception of the physical environment (Figure 2), the interaction forms based on extended reality technology are better able to enhance the audience's connection to the actual environment while increasing the fun of the experience. The physical presence and atmosphere of the actual environment enhance the embodied nature of the audience's experience. This paper, therefore, explores the mechanics of digital gamification based on extended reality.

Interaction forms	Interestingness of the experience	The amount of information presented	Perceived enhancement of the physical scene
Online interaction	■■■■■	■■■■■	■
XR interaction	■■■■	■■■■	■■■■
Offline interaction	■■	■■■	■■■■■

Figure 2. Comparison of interaction forms of gamification mechanisms

3. An application framework of digital gamification platforms in the design of exhibition spaces

3.1. THE DRIVERS AND DEFINING CHARACTERISTICS OF GAMIFICATION

Based on theoretical research, Jane summarised four defining characteristics of gamification: goals, rules, feedback systems and voluntary participation (McGonigal, J., 2011). (1) goals are used to focus the user's attention on a point so that a specific outcome is achieved, i.e. to provide the user with a purpose for their behaviour. (2) rules are used to guide the player into the story situation set by the game. By setting certain constraints to limit the user's goal attainment, the user gains a sense of achievement and engagement in overcoming obstacles. However, the difficulty setting of the rules needs to consider the individual ability differences of the participants to avoid forming a participation threshold. (3) the feedback system can prompt the difference between the player's current progress and the overall goal, which helps the user to control their behaviour during the experience. (4) voluntary participation means that the player can consciously and voluntarily choose the game tasks or goals. These four factors - goals, rules, feedback systems and voluntary participation - work together to achieve the four primary purposes for which users play the game - to be satisfied with their work, to experience success, to build social interaction and to participate in ambitious endeavours (Figure 3). In line with the laws of behavioural psychology theory, the user experience is enriched, enhancing the game's appeal and the user's willingness to participate. This fits with the objectives of the exhibition design experience that this study aspires to achieve and is also applicable to the design of this digital gamification mechanism.

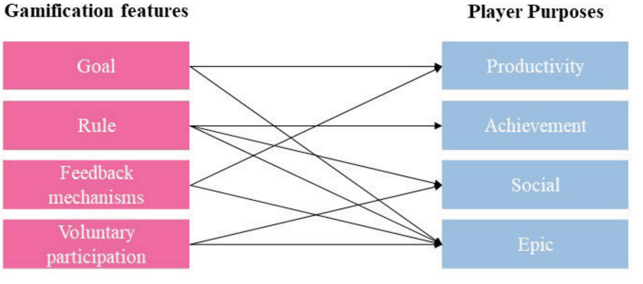


Figure 3. Gamification defining features and player purpose

3.2. GAMIFICATION DEVELOPMENT MODEL

The MDA model is widely accepted for explaining game design from a systems perspective. It divides the game system into three distinct components: Mechanics, Dynamics and Aesthetics. Mechanics are the underlying logic at the algorithmic and data level the designer has set up for the game. Dynamics is the interaction between user behaviour and the game. It serves as the interface between the designer's mechanics and the aesthetics of the player's experience, on the one hand showing the player the feelings the designer wishes to convey and, on the other hand providing user feedback to the designer to make adjustments to the game mechanics. Aesthetics, on the other hand, describe the emotional changes that occur as the user moves through the dynamic process and include eight core descriptions: Sensation, Fantasy, Narrative, Challenge, Fellowship, Discovery, Expression, Submission(Hunicke, R., 2004). The MDA model enables a two-way interaction between the designer's goals and the user's experience, which constitute the game's use and entertainment value.

As a model with a high degree of abstraction, MDA reflects the underlying logic of gamification design but is not a sufficiently concrete and clear approach suitable for guiding practical design. It is a multidisciplinary process involving psychology, design and computing. How can this series of processes be effectively linked? Based on this question, Benedikt reviewed the literature on gamification design methods. They found that most gamification developments follow seven similar main stages: "project preparation - pre-analysis - gamification conception - concrete design - platform building - evaluation of success - monitoring and management" (Morschheuser, B., 2018).

On this basis, combined with the characteristics of the exhibition building space and the above analysis, the digital gamification design process built in this paper is shown in the following diagram

(1) Objectives. Collect the objectives of each stakeholder. Validate the reasonableness of the objectives under the basic principle of achieving mutual benefits for the organisational side and the stakeholders. As multiple objectives can blur the user's attention in the gamification experience, 1-2 objectives need to be selected as the most worthwhile to achieve based on the priority of the objectives.

(2) Analysis. Analysis refers to sorting out and analysing existing conditions within the space. Firstly, it summarises the tangible information (spatial scale/function, exhibits) and intangible information (stories, legends, customs, etc.) that needs to be conveyed within the exhibition design; secondly. It builds a logical network of relationships between all the information and classifies the level of importance according to its cultural value.

(3) Design. Using the rules as a thread, string together the ideas from the first two stages to form an abstract conceptual prototype. The prototype is refined through feedback.

(4) Development. Development can be divided into two phases: content development and technical development. Content development enriches the abstract game prototype into a complete interactive script guided by the objectives defined in the early stages and designs appropriate plot contexts and gamification mechanisms. At the same time, combined with real space, the five elements-"spatial nodes - cultural information - plot settings - interaction modes - gamification elements" are linked to building an implementable script content.

After content development, technical development begins, including asset library creation - application development - interaction deployment. Firstly, the creation of the asset library includes 2D assets (UI interface, icons) and 3D assets (live scans, tasks, actions, effects, digital exhibits). Secondly, a gamification application is developed using the Unity platform, guided by a script. Finally, the application is imported into XR devices such as Hololens to complete the interactive deployment and achieve an immersive experience for visitors.

(5) Evaluation. Quantifying the user experience in terms of both objective and subjective data. Objective data includes eye movement data (visual heat map, visual dwell time, visual transition state) and behavioural data (walking trajectory, dwell time, interaction behaviour). Subjective data include questionnaires, interviews and personal perception mapping etc.

Through the evaluation of subjective and objective data, on the one hand, the effect of the digital gamification mechanism on visitors' experience enhancement and knowledge learning can be derived through their subjective experience feelings and mapping. On the other hand, grasping visitors' cognitive behaviour patterns in the process of gamification experience can provide data reference for the subsequent optimisation and improvement of the digital gamification mechanism.

4. An empirical exploration of digital gamification platforms

4.1. OVERVIEW OF RESEARCH SUBJECTS

The History Hall of Harbin Institute of Technology (Shenzhen) is located inside the Shenzhen Campus of Harbin Institute of Technology in Nanshan District, Shenzhen, Guangdong, China. It is a permanent thematic exhibition intended to introduce the long history of the university, its structure, its work and its outstanding research achievements to the university's staff, students and foreign visitors.

The exhibitions in the exhibition hall are rich in content, spanning a wide range of periods. At the same time, the exhibition information is presented mainly in the form of graphic panels, partially supplemented by digital media technologies such as panoramic video roaming and interactive projections. People often need to visit and learn on their own in the exhibition hall without a conventional guided tour. There is much room for improving their cognitive experience, so this space is an object of empirical evidence for exploring digital gamification mechanisms.

4.2. APPLICATION OF DIGITAL GAMIFICATION MECHANISMS

4.2.1. Objective

The nature of the exhibition is a single propaganda function, with a clear and fixed theme and a small number of stakeholders, including the school administration, staff and students, and external guests. Therefore, the museum aims to introduce the history and achievements of the school in a systematic way and to promote the school spirit.

4.2.2. Analysis

By sorting out the information within the space, a logical relationship diagram of the exhibition content can be constructed from two dimensions: the object to which the information belongs and the information hierarchy (Figure 4). After overlaying the different coloured information layers with the current flow of the exhibition, it can be found that the current flow generally follows the sequence of information narration from macro to micro, from cause to effect. However, there is still a localised mix of colour blocks in the overall space (Figure 5). This reflects the intersection of different layers of information in the arrangement of the existing spaces (Figure 6). It also reflects, to some extent, the difference between the traditional, object-centred layout logic and the human perception-centred logic of gamified layout.

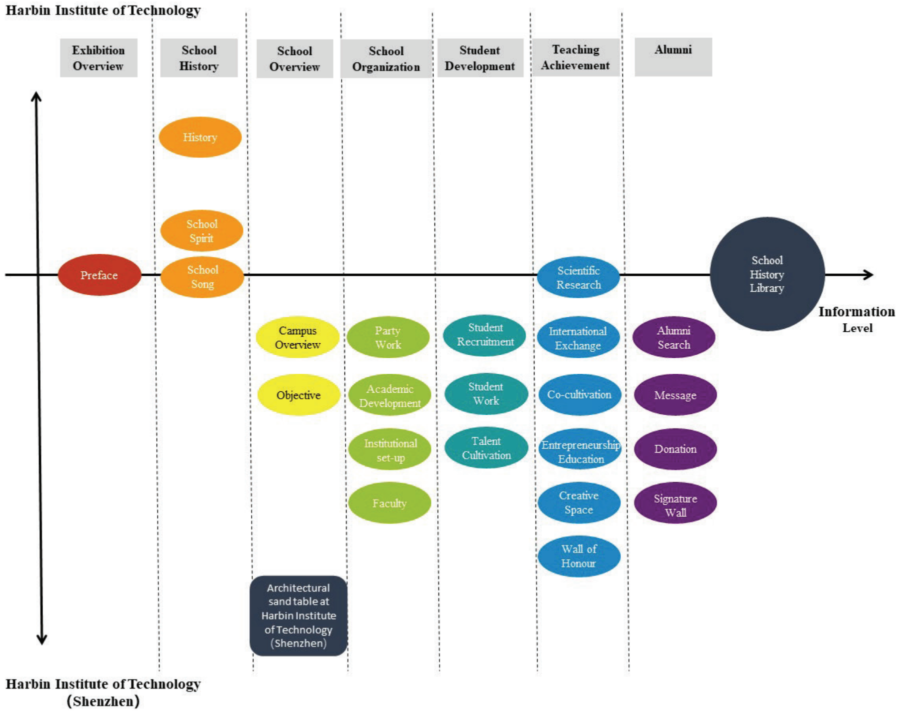


Figure 4. A logical diagram of the exhibition content

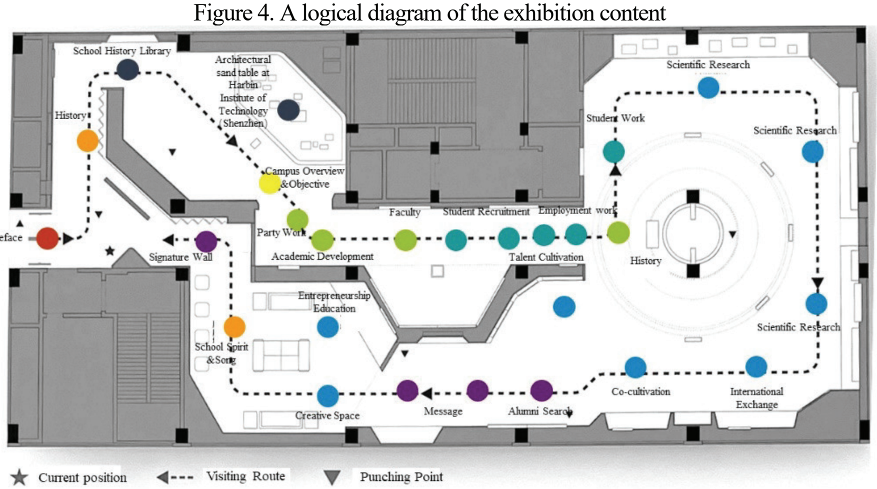


Figure 5. Projection of gamification logic in real physical environment

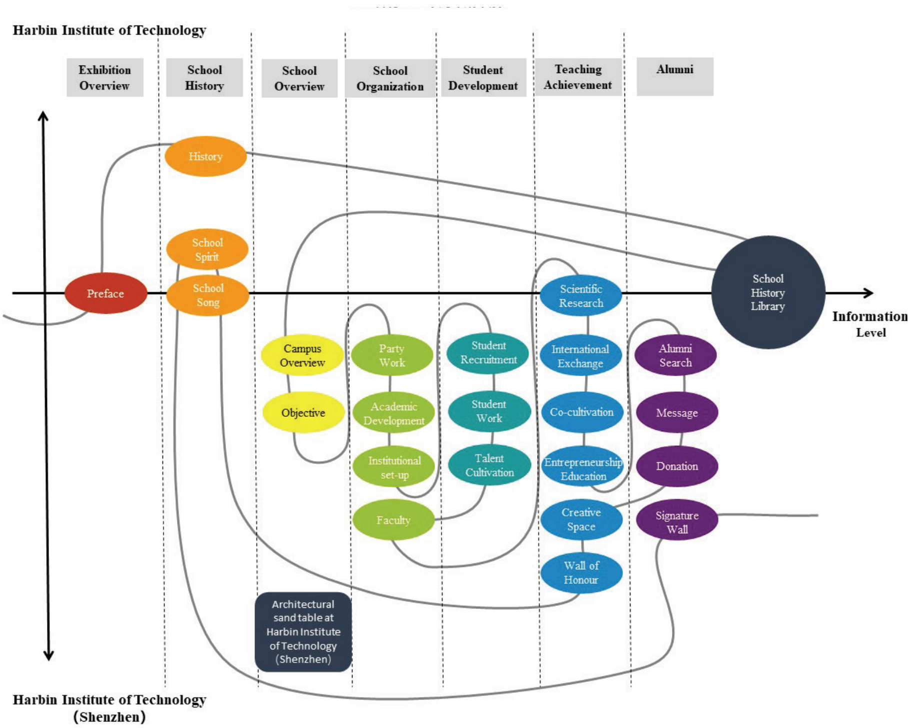


Figure 6. Current layout logic

4.2.3. Design

Considering the limited cognitive energy of the audience, it is not easy to experience the completion of all sections of the exhibition content. It is, therefore, necessary to

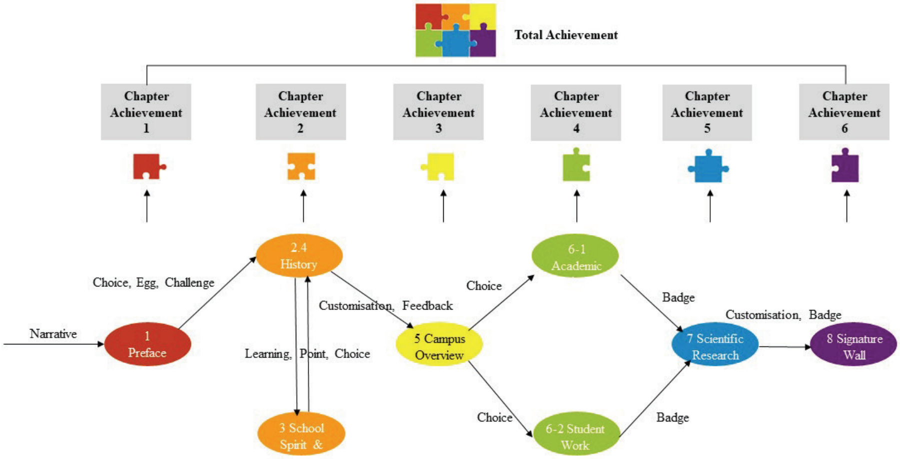


Figure 7. Prototype diagram for gamification design

classify information into two categories, "must complete" and "choose to complete", based on (1) effectiveness in terms of goal attainment and (2) importance of the information. "Must-complete" refers to the storyline experience that the user must participate in during the visit. "Opt-in" means the user can choose at least one of the same type of episodes to participate in and then move on to the next stage of the experience. After completing the information segmentation, appropriate gamification elements are selected to link the information content and build the game prototype diagram (Figure 7).

4.2.4. Content development

Considering that the exhibition's objective is to introduce the school's history and spirit systematically, the plot of this digital gamification mechanism is set in the context of a commission that "I" receive from a person who has lost his memory. "I" need to go back in time and help the amnesiac to recover his lost memories. The game's goal is to answer two questions posed by the amnesiac: 1. Who is he? ; 2. What did he do?

The game mechanic is: visitors follow the cursor to the plot points in the space and complete the plot exploration and tasks at each station to obtain the corresponding "memory fragments". The "memory shards" completeness is linked to completing the quest at that station. Once all the "fragments" have been collected, the visitor can retrieve the complete memory and complete the commission (Figure 8).

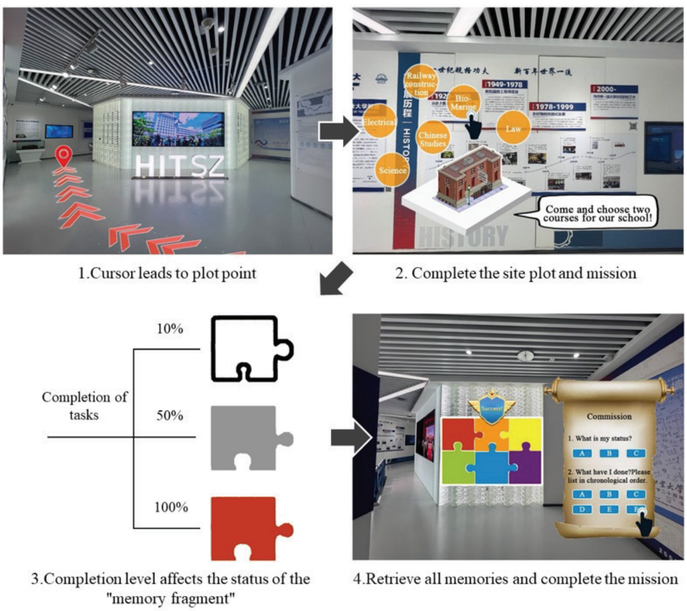


Figure 8. Diagram of the game mechanic

The plot points and the flow of the school history museum were identified based on the previous prototype game map. The design of the interactive script is based on this.

5. Conclusion

Theories on how gamification mechanisms affect people's motivation, behaviour and learning in different contexts are well documented. However, research into the integration of digital gamification mechanisms with the design of scenic experiences in natural exhibition spaces is still at an exploratory stage.

Based on the basic model and decisive characteristics of gamification, this paper constructs a digital gamification design model. In the specific application of the History Museum of HIT (Shenzhen), the content layout will show different logical states under the gamification design and the traditional design mechanism.

However, as the application process of "design-develop-implement-evaluate" for digital gamification mechanism is a longer-term process, this paper has only explored it to the content development stage. Further empirical research in real-life spaces is required to evaluate the effectiveness and impact of this design approach. In the future, the design approach can be improved and refined in light of practical experience. It is also necessary to determine the effectiveness of specific gamification elements in depth. This will enable the most appropriate elements to be selected for application in the design of future exhibition spaces, further enhancing the positive effects of digital gamification mechanisms.

Acknowledgements

This research is supported by the following funds: National Foundation for Philosophy and Social Science Late Grant Project (No. 19FXWB026); Science and Technology Plan Project of Guangdong Province (No. 2017A020220011).

References

Csikszentmihalyi, M. (1991). Flow. København: Munksgaard.

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: defining" gamification". In Proceedings of the 15th international academic MindTrek conference: Envisioning future me-dia environments (pp. 9-15). https://doi.org/10.1145/2181037.2181040

Hunicke, R., LeBlanc, M., & Zubek, R. (2004). MDA: A formal approach to game design and game research. In Proceedings of the AAAI Workshop on Challenges in Game AI (Vol. 4, No. 1, p. 1722).

McGonigal, J. (2011). Reality is broken: Why games make us better and how they can change the world. Penguin.

Morschheuser, B., Hassan, L., Werder, K., & Hamari, J. (2018). How to design gamification? A method for engineering gamified software. Information and Software Technology, 95, 219-237. https://doi.org/10.1016/j.infsof.2017.10.015

Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future direc-tions. Contemporary educational psychology, 61, 101860. https://doi.org/10.1016/j.cedpsych.2020.101860

Wee, S. C., & Choong, W. W. (2019). Gamification: Predicting the effectiveness of variety game design elements to intrinsically motivate users' energy conserva-tion behaviour. Journal of environmental management, 233, 97-106. https://doi.org/10.1016/j.jenvman.2018.11.127

INTEGRATING HERITAGE PRESERVATION AND CITY DEVELOPMENT

A Design Framework for Digital Interactive System Based on Augmented Reality for the World Cultural Heritage Grand Canal

YUE YING ZHANG[1], HAO RAN WANG[2] and KE ZHU[3]
[1,2,3] *Southeast University, School of Architecture*
[1]*Zyy1053160151@gmail.com, 0000-0002-6892-8214*
[2] *w840604772@163.com, 0000-0002-9288-3933*
[3]*220200027@seu.edu.cn, 0000-0002-7490-9247*

Abstract. ICOMOS has pointed out that digital technologies have become important ways to protect heritages. Existing researches focus on the digital reconstruction of heritage in virtual spaces. Less attention is paid to the utilization of heritage entities and the potential for virtual heritage interaction to present complex values of heritages. Augmented reality (AR) can integrate text, images, and models into digital information to add to heritages in real sites. This paper takes the Grand Canal， a great masterpiece of Chinese civilization and even human history, as a research object. With the acceleration of the urbanization process and the change in transportation patterns, the contradiction between the preservation of it and urban development has become increasingly prominent. Based on the analysis of heritage values, this research reconstructs digital models of heritages and develops interactions with them in the Unreal Engine, relying on mobile AR to create a novel cultural landscape through a combination of virtual and real Grand Canal. This research contributes to improving the material and cultural living standards of citizens by integrating heritage preservation and urban design through the design and development of this digital system.

Keywords. Digital Heritage Preservation, Augmented Reality, Heritage Value, The Grand Canal, Cultural Landscape.

1. Introduction

As a witness to history, cultural heritage is a precious non-renewable resource. The display and interpretation of heritage is an important way to enhance the public's awareness of the value of heritage itself and thus promote heritage preservation(Neil A. Silberman, 2008). With flexibility and information integration, various digital technologies have been used to display. Nowadays, specialists in architecture and other relative areas have seen heritage as a kind of social infrastructure. However, through

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 727-736.

rapid urbanization, the contradiction between the preservation of architectural heritage and the development of urban space has gradually become prominent(Mason, 2006). Based on the principle that an augmented reality(AR) system can superimpose virtual content on real scenes, this paper tries to resolve the contradiction through a well-designed and developed AR system.

Figure 1 shows the Wusha Shipyard alongside the Grand Canal, one of the greatest masterpieces of hydraulic engineering in human history(UNESCO, 2014), as the research object. In the past, many shipyards were built on the riverside, including the Wusha Shipyard. Those shipyard architectures and shipbuilding techniques also became part of the civilization of the Grand Canal. In modern times, however, the waterway could no longer accommodate the manufacture and transportation of large vessels. Wusha shipyard gradually fell out of use. Nowadays, as a registered historic building, the reconstruction of the Wusha Shipyard and its surroundings is strictly restricted by regulations. This paper attempts to introduce AR into the urban design of this historic area, and discuss the process to design and develop a novel cultural landscape combining virtual history and real space through an in-situ digital system.

Figure 1 Photo of Wusha Shipyard

2. Related Research and Works

Augmented reality has been widely used in various disciplines including heritage display and urban design. Review and evaluation of these cases contain the specific technology used, real scene augmented, and virtual contents generated.

2.1. AR SYSTEM FOR HERITAGE PRESERVATION

The concept of information augmented on heritages for the exhibition has been around for a long time. In the Heidentor ruins in Australia, its missing forms are complemented by setting up transparent acrylic panels with an intact outline at fixed points (Ledermann and Schmalstieg, 2003). Through the applications of AR, IntraCom developed an electronic guide system for monuments that provided on-site restoration of ancient Greek monuments(Gleue and Dähne, 2001). There also has been a restoration of the mottled and mutilated frescoes on the vault of the Basilica in Brasilia(Portal´ es et al., 2009). These researches focus on recreating the physical forms or materials that heritage used to be. However, it is difficult to interpret the rich value of the heritage by simply recreating its physical components. Studies have also been conducted to superimpose virtual living scenes of the past on real space(Papagiannakis et al., 2005). This enabled audiences to gain a deeper understanding of the culture embedded in the architectural heritage.

These in-situ AR systems for heritage preservation are still in academic research and have not been widely used. However, most of the applications in practice take specific objects in reality like postcards as AR markers instead of heritage itself. Researchers have already developed interactive virtual heritage in AR applications through game engines such as Unity3D to represent internal structures of the architectural heritage(Nagakura and Sung, 2017) or the design and construction process(Niblock et al., 2022). Though these practices dismiss a possible connection between real and virtual components for heritages, their concentration on interaction with virtual objects can truly increase users' understanding of history(Kapell and Elliott, 2013). As values of heritage are multi-dimensional, it is difficult to represent the rich culture that heritage contains. As a result, a relatively integrated and comprehensive system for heritage presentation is urgently needed.

2.2. AR SYSTEM FOR URBAN ENVIRONMENT

Real environments superimposed by fantastic virtual have been imaged in many pieces of research. Keiichi Matsuda proposed a city in augmented space containing compound functions like entertainment, guidance, advertisement, communication, and so on(Matsuda, 2010). Nowadays, with the development of both software and hardware, it is possible for real-time data transmission and rendering of massive images, which are important in an AR system for the urban environment. Therefore, numerous applications have been released on the market and come into citizens' daily life.

The accessibility and interactivity of mobile AR can enable a closer relationship between the public and urban environment including heritages(Chandini Pendit et al., 2014). Digital platforms can open up various virtual content to people with different identities, and promote heritage conservation and urban development through their systemic co-construction in a further step. Hence, some governments have regarded the AR system for heritage preservation as a part of smart city projects (Batchelor and Schnabel, 2021).

3. Research Content and Methodology

There are complex factors about heritage preservation and urban design to consider, therefore, it is necessary to identify specific aspects of design contents. Meanwhile, as a digital project, this research focuses on not only the design framework but also the technical development process of the AR system.

3.1. DESIGN ASPECTS

According to the characteristics of AR, two levels of design were required primarily. One is the real scene chosen to be augmented, and the other is corresponding virtual content superimposed on reality. That is, the virtual content should be tightly relative to the theme of the real scene, instead of a fantastic but irrelative one. It is challenging that the selection of the most representative real space is decided by the value of heritage, as well as its augmented virtual content, which requests a comprehensive understanding of heritage and its culture contained.

Furthermore, this paper focuses more on the processing of diverse virtual content than real space. For heritages like the Wusha Shipyard, there are different forms of resources to present, including introductive texts, documents or literature, drawings or photos, media like videos or audio, and so on. It is difficult to use suitable kinds of resources corresponding to the real scene. Accordingly, the integration of different kinds of resources is also important, as they can make a joint effort on enriching reality. Interactive operations of these virtual objects should also be taken into consideration to help users operate easily and understand the culture interestingly.

3.2. TECHNICAL DEVELOPMENT

The system was designed in traditional methods, and the resources were researched and processed in both real and virtual terms. At the same time, virtual construction is modelled for the architecture of the Wusha Shipyard and its surroundings in 3D software like Rhino.

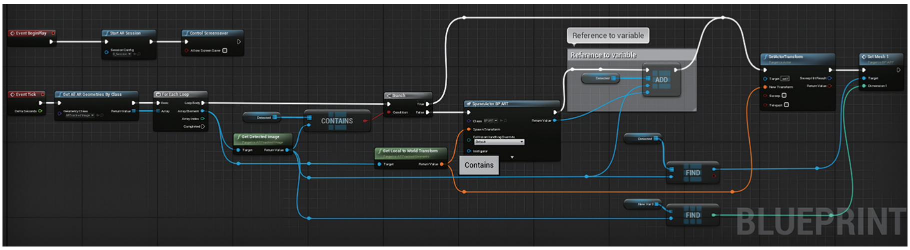


Figure 2 Blueprint for matching virtual object with the real scene

Although most AR systems are developed using Unity3D and ARKit, this paper tries Unreal Engine and ARCore for convenience and free accessibility. Unreal Engine is a game engine for developing systems with immersive and interactive manipulation. Its capability of 3D rendering can realistically reconstruct virtual objects. Most importantly, its visual programming named Blueprint greatly reduces the technical threshold and provides the possibility for an architect to engage in programming. Thus it is likely to meet the development needs of multi-disciplinary collaboration. In this research, the resources collected are imported into Unreal Engine 4.6.2 to render and make interactive manipulation programming through Blueprint shown in Figure 2. 3D models can be imported through a plugin called Datasmith. This version of Unreal Engine has inserted plugins for both ARkit and ARCore to develop and output.

4. Design Framework and Contents

Faced with the complex composition and surrounding environment of the Wusha Shipyard, the design of the system framework and specific concepts of scenes are based on its heritage values. Furthermore, this paper designs the presentation methods and interactive operations according to the characteristics of the contents.

4.1. THEME PLANNING

We first analyze the value of the Wusha shipyard as a heritage in three dimensions: historical, scientific, and artistic. Figure 3 shows the process of corresponding the intangible value of Wusha Shipyard to specific tangible objects in the real environment.

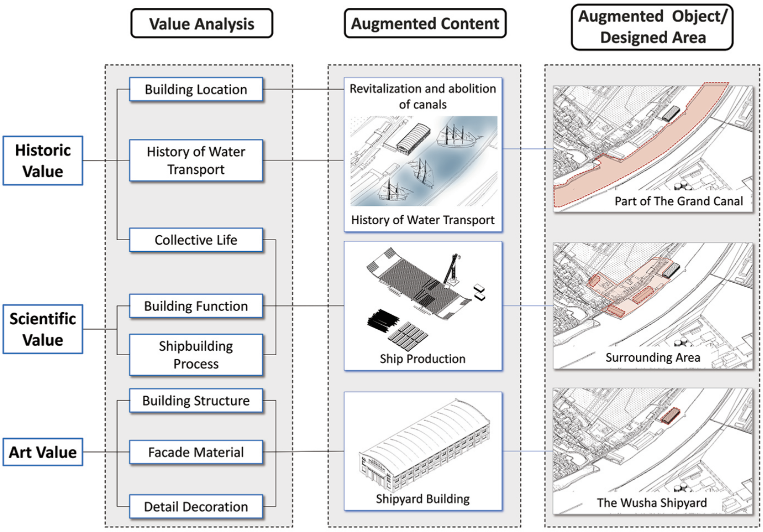


Figure 3 Exhibition Theme and Function Planning

Artistic value can be shown by the specific structural or detailed forms of the Wusha Shipyard architecture. Its main structure is an old-style row frame with a steel truss roof. In the north and south elevations is its red clay brick façade with a grey clay brick base. Rhythmic row-frame columns and window openings form the main image of this architectural heritage. Also, large windows and embellishments on the gable walls reflect the typical style of factory buildings in the local area at the time.

For its scientific value, the shipyard and its surroundings can reflect the technical processes of the shipbuilding industry. Its large span structure required for shipbuilding, the launching of the shipyard (now destroyed) in old historical photographs, and the scenes of the life of the shipyard workers, can all provide information for the study of modern ship production activities in the factory.

Compared with the artistic and scientific values carried by tangible material objects, historical values are more abstract. They are analyzed in terms of two aspects. One is Huai'an's existence as the hub of canal transport throughout the entire route of the Grand Canal. The other is the rise and fall of the Li Canal, a part of the Grand Canal, along with the construction of the shipyard. This part is overarching in the system, and the presentation of this civilization is intended to use the river as an augmented object.

4.2. AUGMENTED SCENE AND INTERACTION DESIGN

The image recognition used in AR systems is a marker-based technology, requiring only a limited amount of the user's field of view (FoV) to be rendered with computer-generated graphics(Kasapakis et al., 2016). Based on the above themes, Figure 4 shows suitable scenes and perspectives to be recognized in reality that can present the corresponding heritage values for the specific design. As for the values of Wusha Shipyard displayed mainly in the virtual system, real space can be planned for other functions under the upper planning.

Information that can be augmented is diverse. It used to be a simple overlay of text, 2D images, and videos, which are much easier to recognize. Interaction especially with 3D models takes more time and effort but can make visitors easy to understand. Figure 4 also shows the virtual recreated scenes and the interactive operations with each kind of resource.

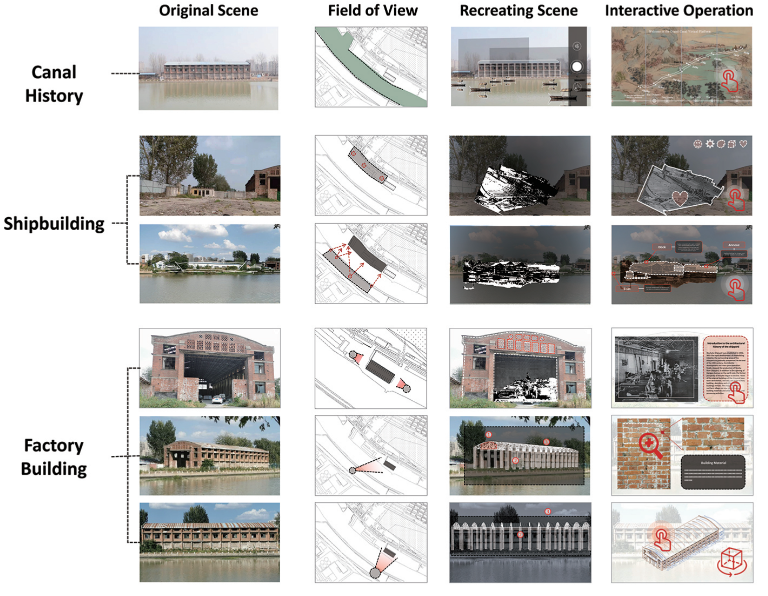


Figure 4 Specific Scenario and Interaction Design

4.2.1. Chronological overlay of canal history

The existing canal is no longer used for transport and production functions. In the augmented interface, the virtual Grand Canal shows a history dominated by canal transportation. This system is planned to restore the scenes of boats traveling on the Grand Canal as shown in Figure 5, supplemented by text and historical maps to introduce the history. In the upper planning, it is designed as an ecological landscape river. Therefore, it is mainly retained for water purification and treatment.

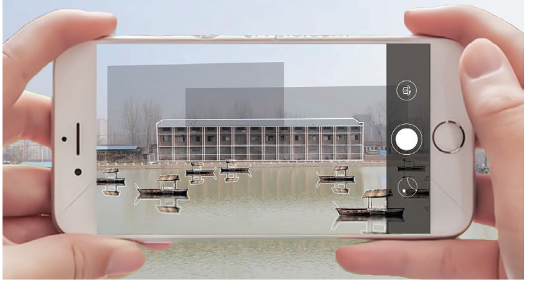


Figure 5 scenes of boats traveling on the Grand Canal

As the beginning parts of the application, these augmented scenes of different dynasties can be switched through timeline buttons to show the rich history. Interactive operations in this part are basic with relevant text and introductory notes.

4.2.2. Historical representation of industrial production

This session deals with the procedures of shipbuilding and the life of staff in the factory in the 1960s. A combination of texts, historical photographs, and videos is used to show residents' memories at that time. In the past, the shipbuilding process was mainly conducted outdoors, where the dockyard and the platform no longer exist today. The system recreates the equipment and structures required for the shipbuilding in virtual space based on the original layout in reality.

The interactive experience is only available inside the site and integrated with real urban space. In this area, parts of the ship to be constructed and assembled in digital forms are scattered at the starting point and can be put down in the right location virtually. This interesting interaction acts as a guidance system to instruct visitors to walk following the construction process. Through in-situ augmented experience on a real scale, visitors can acquire a relatively authentic understanding of the shipbuilding process and outdoor space around the shipyard.

Meanwhile, the real space can be used for other functions to serve citizens or visitors instead of the exhibition, as the information about Wusha Shipyard is presented virtually. This area in reality is designed as waterfront green space, setting hard court leisure spaces at locations of the best view to imply there exist interactions.

4.2.3. 3D enhancement of shipyard architecture

The display of the shipyard buildings focuses on architectural details. As the shipyard architecture is preserved well in reality, the virtual presentation of the shipyard is mainly aimed at the urban space distant from it. In this way, the attractiveness of the whole area surrounding Wusha Shipyard can be enhanced. Following interactive operations are proposed for visitors' better understanding of this 3D architecture.

(1) Extraction of structural elements: In real space or photography, the steel frame structure with large and dense trusses is often hidden beneath the skin of the building. In this system, the AR system can generate and render the virtual steel frame structure of the shipyard as a virtual model and make it visible above the skin of the building.

(2) Enlargement of the facade details: Compared to the huge volume of the shipyard building, small-scale facade decorations are invisible to the naked eye. Highlighting details of the window and decorations on the digital interface suggests they can be interacted with. Clicking on the highlighted section allows users to zoom in and see these details of the construction more clearly.

(3) Rotation of the building model: In reality, the grant size of the building also makes it difficult to see all angles of it from limited positions. However, a virtual building model can make different angles of the roof, foundations, opposite sides, etc. visible through operations such as rotation.

5. Discussions

As is shown in Figure 6, this research finally proposes a framework for the design and development of an AR interactive system that combines the historic values of architectural heritages and the development of urban spaces. Through this framework, heritage can be preserved without overexploitation. Meanwhile, it can be made full use of as social infrastructure in urban design, both in virtual and real spaces through AR and interactive operations. The effects of the system should be evaluated from two aspects including heritage preservation and urban development respectively. There also exist some limitations to break through in future research.

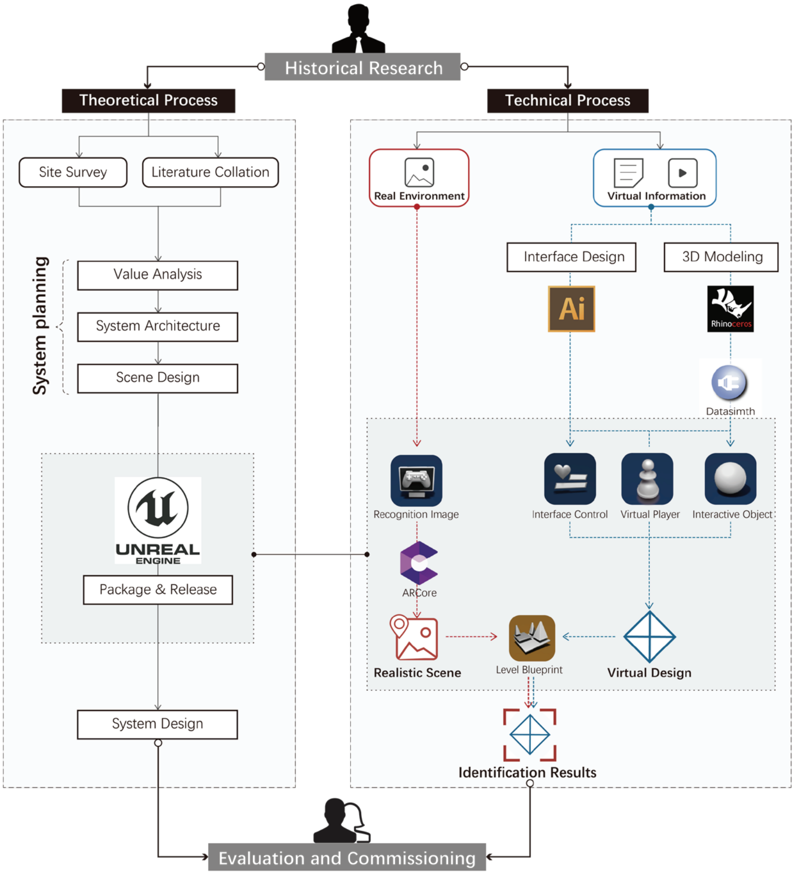


Figure 6. Diagram for design and development process

5.1. INNOVATION IN HERITAGE PRESERVATION

This in-situ AR system can complement and enrich heritage entities with intangible cultural contents that is difficult to present. This leads to minimal disturbance to the heritage entities, as the extent of reconstruction of existing heritages and their surroundings has always been controversial.

Based on heritages in reality, the system also works as a digital platform to assemble different types of information that can document and display the values of heritages. Visitors can get access to the information in an organized approach through interactive operations with virtual objects, strengthen their understanding of heritage, and enhance their awareness of heritage preservation.

5.2. INFLUENCE ON CITY DEVELOPMENT

The system ensures minimal disturbance to the local landscape by making full use of heritage as a kind of social infrastructure. As most cities along the canal are saturated with construction, there are few undeveloped lands. Restoring the historical buildings through AR can save the land and reduce the development cost and the disturbance to the current life of residents.

In addition, the combination of virtual scenes and real sites creates a novel kind of urban landscape. Citizens can experience both fantastic historic circumstances and realistic present sceneries, which enriches the spatial dimension of the city.

The role of the digital economy cannot be underestimated. The application of the system can inject fresh energy into the city's tourism as well, and enable visitors to gradually expand their knowledge of heritage as a whole. As The Grand Canal was a cross-regional link for economic and cultural exchange in ancient times, this system will reinvent this function. It can transcend the limits of time and space to enable the exchange and interaction of different people.

More than just an exhibition of historical information, the system has the potential to work as a part of a smart city system, as the content of the virtual system can be expanded without restriction. The easy access and operation also facilitate the participation of all levels of people and the variety of activities.

5.3. LIMITATIONS AND FURTHER WORK

There still exist several limitations in both research and the system. The research focuses mainly on the design of virtual aspects, while the real space can also be used to suggest the interaction of virtual objects through forms of landscape or architectural design such as paving. Thus, the design of the real urban environment is equally prominent, and the approaches to it will be further researched.

The current AR system is under experiment, so there still exist some technical problems to solve. For example, the obstacles in the field of view can negatively affect the recognition of a specific scene, along with users' virtual experience. Also, the effects of this augmented system request to be examined more scientifically through in-situ experiments and user interviews.

6. Conclusion

Faced with the contradiction between heritage preservation and city development, the design framework of an augmented system is based on diverse values of heritage. All these values, including historic, scientific, and artistic ones, along with contemporary values proposed recently, correspond to social awareness of history and willingness for urban life. Augmented reality provides an opportunity to convey and present these values to the public. Through the superimposing of virtual objects on real scenes and interactive operations, this AR interactive system is prospected to work on integrating the two factors and improving both of them.

References

Batchelor, D., & Schnabel, M. A. (2021). *Opportunities and Recommendations for Local Governments Delivering Smart Heritage*. 749–758. Hong Kong. https://doi.org/10.52842/conf.caadria.2021.2.749

Chandini Pendit, U., Bahrin Zaibon, S., & Aida Abu Bakar, J. (2014). Mobile Augmented Reality for Enjoyable Informal Learning in Cultural Heritage Site. *International Journal of Computer Applications*, 92(14), 19–26. https://doi.org/10.5120/16077-5286

Gleue, T., & Dähne, P. (2001). Design and implementation of a mobile device for outdoor augmented reality in the archeoguide project. *Proceedings of the 2001 Conference on Virtual Reality, Archeology, and Cultural Heritage - VAST '01*, 161. Glyfada, Greece: ACM Press. https://doi.org/10.1145/584993.585018

Kapell, M., & Elliott, A. B. R. (Eds.). (2013). *Playing with the past: Digital games and the simulation of history*. New York: Bloomsbury Academic.

Kasapakis, V., Gavalas, D., & Galatis, P. (2016). Augmented reality in cultural heritage: Field of view awareness in an archaeological site mobile guide. *Journal of Ambient Intelligence and Smart Environments*, 8(5), 501–514. https://doi.org/10.3233/AIS-160394

Ledermann, F., & Schmalstieg, D. (2003). Presenting Past and Present of an Archaeological Site in the Virtual Showcase. *4th International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage*, 1–6.

Mason, R. (2006). Theoretical and Practical Arguments for Values-Centered Preservation. *CRM-Washington*, 3.

Matsuda, K. (2010). *DOMESTY/CITY：THE DISLOCATED HOME IN AUGMENTED SPACE*.

Nagakura, T., & Sung, W. (2017). AR mail from Harbin. *ACM SIGGRAPH 2017 VR Village*, 1–2. Los Angeles California: ACM. https://doi.org/10.1145/3089269.3089274

Neil A. Silberman. (2008). ICOMOS Charter for the Interpretation and Presentation of Cultural Heritage Sites. *International Journal of Cultural Property*, 15(4), 377–383 https://doi.org/10.1017/S0940739108080417

Niblock, C., McGuire, L., Harding, J., Zillner, G., Hamill, C., & Whitney, A. (2022). An augmented and interactive exhibition of an archived model for Frederick Kiesler's Endless House, 1959. *Frontiers of Architectural Research*, 11(6), 993–1006. https://doi.org/10.1016/j.foar.2022.04.002

Papagiannakis, G., Schertenleib, S., O'Kennedy, B., Arevalo-Poizat, M., Magnenat-Thalmann, N., Stoddart, A., & Thalmann, D. (2005). Mixing virtual and real scenes in the site of ancient Pompeii. *Computer Animation and Virtual Worlds*, 16(1), 11–24. https://doi.org/10.1002/cav.53

Portal´es, C., Lerma, J., & PÃ, C. (2009). Photogrammetry and augmented reality for cultural heritage applications. *The Photogrammetric Record, 24*(128), 316–331. https://doi.org/10.1111/j.1477-9730.2009.00549.x

UNESCO. (2014). The Grand Canal. Retrieved December 12, 2022, from UNESCO World Heritage Centre website: https://whc.unesco.org/en/list/1443/

INTERPRETING GENDER DIFFERENTIATION IN URBAN CONSUMPTION PLACES BASED ON THE PREFERENCE LEVEL OF SPATIAL PERCEPTION

A Case Study of Four Commercial Areas in Beijing

XIAOYI ZU[1], CHen GAO [2,3]* and ZHIXIAN LI [4]
[1]*School of Architecture, Tsinghua University, Haidian District, Beijing, China 10084;*
[1]*zuxy22@mails.tsinghua.edu.cn, 0000-0002-4774-1194*
[2]*Leibniz Institute for Research on Society and Space (IRS), Economy and Civil Society, Flakenstraße 29-31, 15537 Erkner, Germany;*
[3]*Geography Department, Humboldt-Universität zu Berlin, 12489 Berlin, Germany*
[2,3]*chen.gao@leibniz-irs.de, 0000-0003-0576-0060*
[4]*College of Environmental Design, University of California, Berkeley, CA 94720, USA;*
[4]*zhixianli@berkeley.edu, 0000-0003-2649-8137*
**Author to whom correspondence should be addressed.*

Abstract. Taking Street view map and Random forest model as the applications and four consumption places in Beijing as case studies, this study proposes a method that maps spatial perception preferences at the local scale to the global scale by following steps: Firstly, download street view images of consumption places from BaiduMap API, then combined the preference of the local street view images scores by the volunteers of both genders and the proportion of visual elements in the images, predicted the preference level of case areas at the global scale by the Random forest model, and finally, through FCN model and sDNA model, fully revealed the gender differentiation phenomenon of consumption places at the image, function and location contents. The results indicate that both genders have a preference for places of Catering function. Besides, females generally prefer consumption places with more conspicuous signboards, greening and better spatial design quality, and have clear pre-determined consumption targets; males generally prefer consumption places with more conspicuous columns, smaller signboards, and have less demand for the spatial design quality of consumption places.

Keywords. Random Forest Model,FCN Model, Built Environment, Consumption Places, SDNA Model

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 737-746.

1. Introduction

With the advancement of urbanisation and the improvement of the economic level, the development mode of Chinese cities has changed from traditional production-oriented to consumption-oriented (He and Lin, 2015). The spatial quality and urban distribution of consumption places can have a significant impact on residents' public activities. Through this view, extensive researches have been conducted on the spatial quality indicator of urban consumption places (Liu et al., 2020; Chen and Wang, 2009). However, the above studies are mainly limited to the planning and function level and lack the description of residents' subjective spatial perception preferences; in addition, these studies regard residents as a whole group, ignoring the influence of gender factors on consumption preferences. Therefore, it is necessary to explore a subjective perception-based measurement method to comprehensively describe the spatial preference patterns of residents' consumption behaviours and the gender differentiation phenomenon in consumption places.

Street view map and Random forest model provide a possibility to implement the above approach (Immitzer et al., 2012; Long, 2016). The former provides a vast amount of urban street view images, which can display the spatial scene images of urban areas; the latter can predict the scores of the full sample based on the scores of a test sample. With these tools, volunteers of both genders only need to score a small sample of street view images, and then the method can predict scores for the whole research area. Therefore, this paper will be unfolded from a local street view image study to a global spatial distribution study, and try to propose a quantitative evaluation method based on subjective spatial perception preference, then reveal the gender differentiation phenomenon of urban consumption places.

2. Methodology

2.1. RESEARCH FRAMEWORK

The research framework consists of four phases: In the first phase, the POIs (points of interest) of all consumption places in the study area are acquired, and the street view images of them are downloaded through the Street View Map API. In the second phase, the trained FCN (Fully Convolutional Networks) model is applied to conduct semantic segmentation of building facade elements in street view images to establish a quantitative measure of their visual characteristics. In the third phase, volunteers of both genders are invited to score the perceived preference level of sampled street view images, then the Random forest model is trained, and the scoring samples with higher accuracy are selected to predict the gender preference scores of all selected consumption places. In the fourth stage, we distinguish the gender preference points and analyse the gender differentiation phenomenon through their quantitative measurement.

2.2. STUDY AREAS AND DATA SOURCES

The study areas of this paper include four well-renowned commercial areas in Beijing: Dongsi, Xisi, Di'anmen and Qianmen (As shown in Figure 1). The reasons for selection are as follows: Firstly, the commercial spaces in these areas have been developed for a long period, and accurate POIs of consumption places can be comprehensively obtained. Secondly, these areas are located in Beijing's historic preservation area, and the buildings are mostly one-story courtyards, so the obtained street view images can directly exhibit the visual characteristics of the corresponding consumption places.

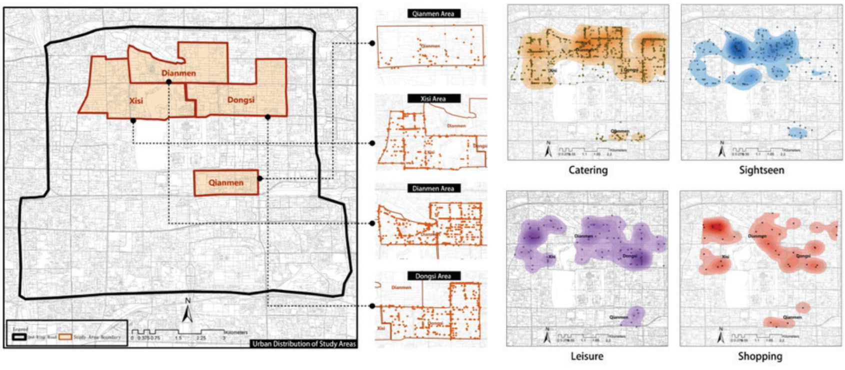


Figure 1. Distribution and kernel density of the study areas

According to the four phases in the research framework, the data to be acquired in this paper include the data on commercial district boundary and surrounding road network, the data on consumption places' POIs within the study area, and the data on street view images of POIs. The data sources are as follows: 1) Commercial district boundary and surrounding road network data obtained through the OSM website (2022). 2) Consumption places' POIs data obtained through the Peking University Geodata platform (2022). 3) Street view image data is downloaded through Baidu Street View Map (2022) (As Figure 2 shows). After the data acquisition, the first two types of data were performed in the Arcmap platform and were classified according to four functional categories of Leisure, Catering, Shopping and Sightseeing, with 14 subcategories (Table 1).

Table 1. Categories of the Functions of Selected POIs

Category	Subcategory	Number of POIs	Percentage of Dataset	
Leisure	Entertainment venue	54	2.89%	5.40%
	Beauty salon	25	1.34%	
	Massage parlour	10	0.53%	
	Performance center	12	0.64%	
Catering	Chinese formal dining	1095	58.56%	77.27%
	Bar&Pub	114	6.10%	
	Western formal dining	110	5.88%	
	Fast food restaurant	99	5.29%	
	Cafe&Tea	27	1.44%	

Shopping	Wholesale markets	6	0.32%	2.30%
	Supermarket	34	1.82%	
	Grocery store	3	0.16%	
Sightseeing	Scenic Area	270	14.44%	15.03%
	Park&square	11	0.59%	

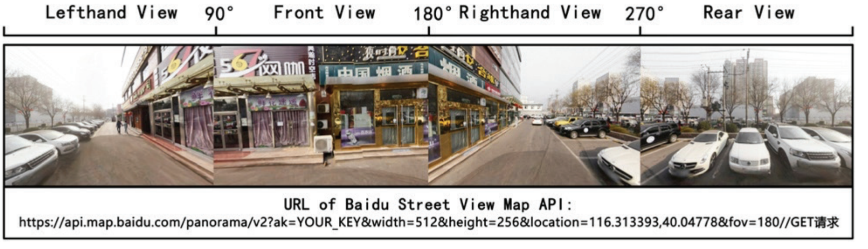


Figure 2. Sample images of Baidu Street View Map

2.3. TOOLS

2.3.1. FCN MODEL

FCN model was selected to quantify the visual features of street view images (Figure 3). The model has a high image segmentation accuracy and can effectively identify various types of facade elements in street view images and reveal their proportions in the images (Huang et al., 2020). Prof. Q. Guan's team at the China University of Geosciences trained FCN with ADE_20K training dataset and obtained accurate segmentation accuracy (0.8144), which was packaged as an open-source software, through which the image semantic segmentation steps in this paper will be implemented (Yao et al., 2019).

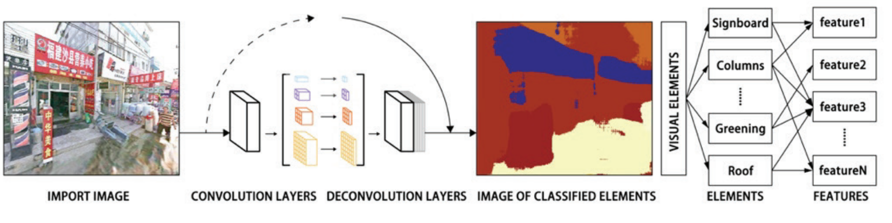


Figure 3. Mechanisms of image semantic segmentation by FCN

To accurately describe the visual characteristics of consumption places, the study selected roof (corresponding to the ceiling tag), column (corresponding to the column; pillar tag), signboard (corresponding to the signboard; sign tag), and greenery (corresponding to the plant; flora; plantlife, grass, flower tag) categories of visual elements, calculate the proportion of each category in the image, and divide them by the proportion of the building facade (corresponding to the building; edifice tag) in the image to describe the proportion of each element in the form of the facade, and then quantitatively measure the visual characteristics of the consumption place.

2.3.2. RANDOM FOREST MODEL

Random forest model is an unsupervised machine learning model which obtains local sample features by random sampling and performs predictive classification of the full domain scores (Breiman, 2001). Its application steps are: 1) draw a few samples from all samples with playback, 2) randomly select several features among all the features of the samples to score and playback, and build a decision tree between the features and the samples by multiple random sampling, 3) repeat the first and second steps several times to generate multiple decision trees, 4) add up the predictions for all the decision trees and assign predicted scores to all samples. The predicted value is calculated as in Eq. (1):

$$\mathrm{RF}(x) = \frac{1}{B}\sum_{i=1}^{B} T_{i,\, z_i}(x) \qquad (1)$$

$\mathrm{RF}(x)$ in Eq. (1) is the predicted value of the Random forest, B denotes a total of B decision trees, z_i denotes the training set used for the i-th decision tree, which is the subset obtained from all training sets sampled by ranks and columns using the bagging method, and $T_{i,\, z_i}$ is the number of training rounds.

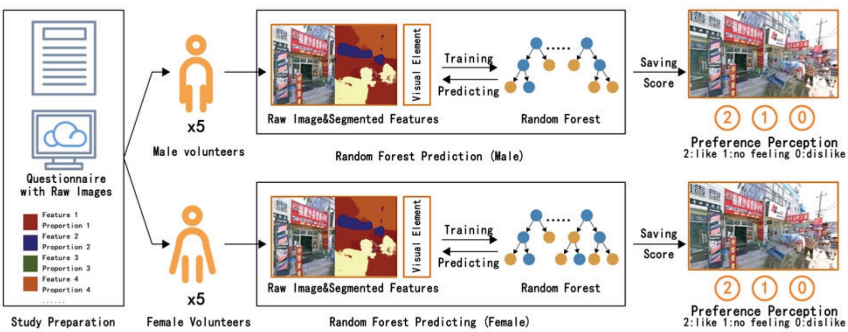


Figure 4. Mechanisms of Random forest model

As shown in Figure 4, 250 street view images were randomly sampled, and then ten male and ten female volunteers scored the sampled images according to the principle of '2-like, 1-no feeling, 0-dislike'. Finally, the training data set was constructed based on the proportion of the four types of visual elements and the scoring results. After the scoring stage, Random forest model was trained by the fit() method in Python, and the predict accuracy was calculated by the score() method. Then the samples with higher accuracy were selected to predict the preference levels of consumption places in the whole study area.

2.3.3. sDNA MODEL

sDNA (Spatial Design Network Analysis) is a spatial analysis tool developed by the Cardiff School of Planning & Geography and the Sustainable Places Research Institute

(SPRI). This tool can quantify the spatial attractiveness and accessibility attributes of a given segment Map. The location characteristic to be described in this study is the spatial attractiveness of male and female preference points in the road network system they are located in, so the NQPDA (Network Quantity Penalized for Distance, Angular parameter is selected, which is calculated as shown in Eq. (2):

$$\mathrm{NQPDA} = \sum_{y \in R_x} \frac{P(y)}{d_\theta(x,y)} \tag{2}$$

In Eq. (2), $P(y)$ denotes the angle-based weight of a node y within the search radius R and $d_\theta(x, y)$ is the shortest topological distance from node x to node y. The larger the value of NQPDA, the stronger the spatial attractiveness of the selected point, and vice versa. In this study, since the mode of excursions in the study area is walking, a walking preference distance of 1500m for commercial activities was selected as the search radius to measure the NQPDA values for each gender preference point.

3. Experiment and result

3.1. IMAGE SEMANTIC SEGMENTATION RESULTS OF FCN MODEL

According to the research framework, a total of 6694 street view images of consumption places were obtained through the Baidu Street View Map API (each point obtained four images of 0°, 90°, 180° and 270°). Firstly, the FCN model is applied to conduct semantic segmentation of all acquired street view images, and then the segmentation data from four angles are summed up to count the proportion of four types of visual elements in the building facade to quantify and measure their visual characteristics. The percentages of the four types of visual elements in all points are shown in Table 2.

Table 2. Proportion of selected elements on the building façade

	Signboard	Greening	Roof	Columns
Proportion	14.75%	18.16%	11.32%	11.48%

3.2. PREDICTION RESULTS OF RANDOM FOREST MODEL

According to the research framework, 250 street view images were taken, and ten volunteers of each gender scored the preference level in the form of an online questionnaire. Then the samples were input into Random forest model to calculate their predict accuracy. The results are shown in Figure 5.

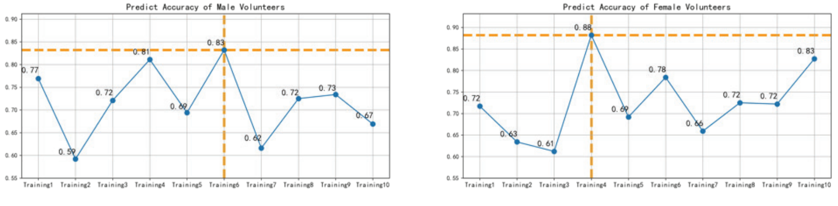


Figure 5. Predict accuracy of Random forest model training

In terms of gender, the predict accuracy of the 6th male training sample peaked as high as 0.83, so the regression results of this sample are taken to predict the male preference levels of all consumption places. The predict accuracy of the 4th female training sample peaked as high as 0.88, and the regression results of this sample are taken to predict the female preference levels of all consumption places.

3.3. ANALYSIS BASED ON FCN AND RANDOM FOREST MODEL

3.3.1. GENDER DIFFERENTIATION AT IMAGE CONTENT

Based on the segmentation and the prediction results of the full domain preference situation, the consumption places are grouped according to the preference level, and the proportion of different visual elements in the building façade is counted by gender. The results are shown in Figure 6 and Table 3.

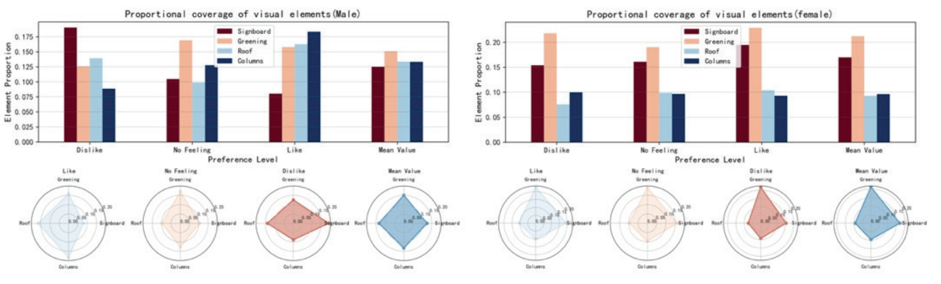


Figure 6. Proportion of visual elements in the building facade

Table 3. Proportional of visual elements of each preference levels

Male				
Preference Level	Signboard	Greening	Roof	Columns
Dislike	15.41%	21.79%	7.56%	9.27%
No Feeling	16.11%	18.99%	9.88%	9.68%
Like	19.47%	22.89%	10.42%	9.97%
Mean Value	17.00%	21.22%	9.29%	9.64%
Female				
Preference Level	Signboard	Greening	Roof	Columns
Dislike	19.04%	12.59%	13.92%	8.86%
No Feeling	10.47%	16.93%	9.88%	12.78%
Like	8.03%	15.78%	16.29%	18.33%
Mean Value	12.51%	15.10%	13.36%	13.32%

The radar chart distributions of females' predicted results are relatively consistent at different levels of preference, with the greening element accounting for the most (mean value of 21.22%), followed by the signboard element (mean value of 17.00%) and roof and columns element always accounting for a smaller proportion (mean values of 9.29% and 9.64% respectively), indicating that the first two elements generate the main attraction in the females' consumption. In addition, the increasing trend of the signboard element preference level indicates that the greater the proportion of this element in the facade, the more likely it attracts females. In contrast, there was a significant difference in the proportion of visual elements at different preference levels for male predicted results, with the proportion of signboard element decreasing significantly (19.04% to 8.03%) and the proportion of columns element increasing significantly (8.86% to 18.33%) as the preference level increased. While the percentages of the other two elements did not show a significant pattern of change, their attractiveness to men declined as the proportion of the signboard element increased.

3.3.2. GENDER DIFFERENTIATION AT FUNCTION CONTENT

Based on the results of the full domain prediction for both genders, the mean value of the preference level was counted in all points by function (the results are shown in Figure 7). When the mean value is greater than 1, the corresponding function category is preferred, and vice versa, not preferred.

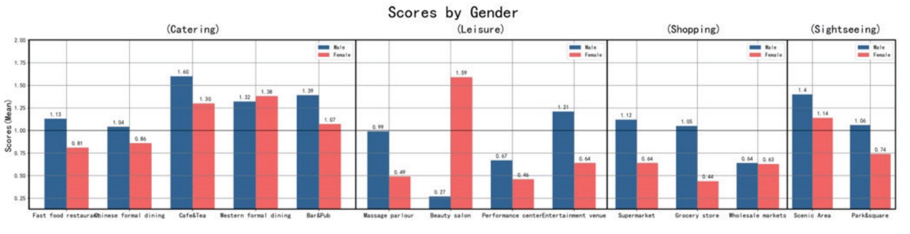


Figure 7. Mean values of the preference levels of all points

According to the results, the mean scores of the Catering are higher than those of the Leisure, while those of Shopping and Sightseeing are lower, and the mean preference scores of males are higher than those of females. In terms of detailed categories, males' preference scores in the Catering are higher than 1, especially in the two subcategories of Cafe&Tea and Bar&Pub, which favour the conversation atmosphere, with mean scores as high as 1.60 and 1.39, while females are not interested in Fast food and Chinese formal dining, and are more interested in Western formal dining and Cafe&Tea, which have better spatial design quality and are more popular, with scores fluctuating around 1.35. In terms of Leisure, males only preferred Entertainment venue (mean score of 1.21), while females preferred Beauty salon (mean score of 1.59). In terms of Sightseeing, females only preferred the Scenic area, while males preferred both subcategories.

Since the study areas are located in the Hutongs of the old city of Beijing, there are no modern shopping malls or shopping districts, and the spatial design quality of other shopping places is poor, so the preference scores of females for Shopping are low (less than 0.65), while males have lower requirements for spatial quality and prefer Supermarket and Grocery store which are close to their daily life (value greater than one).

3.3.3. GENDER DIFFERENTIATION AT LOCATION CONTENT

The POIs with the gender preference (score of 2) were labelled in the Arcmap platform, and then the NQPDA values were also calculated. The spatial distribution of the gender preference points was marked on the map by kernel density, and the distribution of NQPDA values was presented in the form of a box plot. The results are shown in Figure 8.

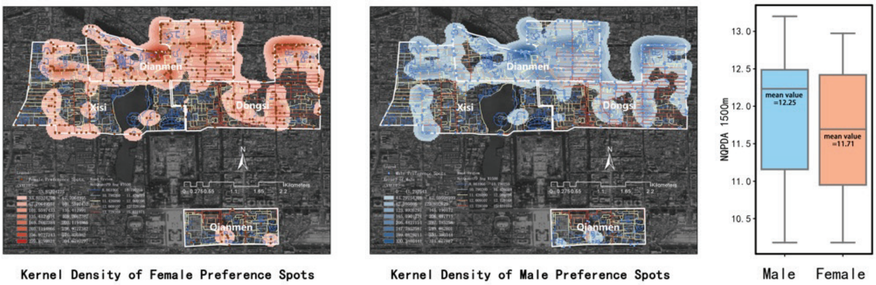


Figure 8. Kernel density of sites by gender preference

The mean value of male preference points is significantly higher than females (12.25 and 11.73). Compared with the kernel density map, the consumption range of both genders is more consistent at the macro level, with a higher density in the Di'anmen area and a lower density in the Qianmen area, where both Xisi and Dongsi are distributed along the main traffic axes in the region. However, microscopically, male preference points are concentrated in roads with higher NQPDA, indicating that their consumption behaviour is mainly guided by the spatial attractiveness of consumption places without a strong pre-determined purpose, while female preference points show a more uniform distribution, indicating that their consumption behaviour has a strong pre-determined purpose and does not comply with the spatial attractiveness.

4. Conclusion

Through the FCN, Random forest and sDNA models, this paper analyses the gender differentiation phenomenon of consumption places by taking four commercial areas of Beijing as examples. The results indicate significant gender differences in the three contents: both genders prefer Catering consumption. Females generally prefer to choose places with more conspicuous signboards, better greening, and better spatial design quality and have a pre-determined purpose of consumption, with less demand for the spatial attractiveness of consumption places. Males generally prefer to choose places with more conspicuous columns, smaller signboards, and more robust spatial attractiveness, with less demand for spatial design quality. Males prefer daily and communication places, and females prefer high-quality exquisite spaces.

Generally, this study proposes a subjective perceived preference regression model based on street view images that can accurately, efficiently, and quantitatively evaluate residents' preferences for consumption places at the urban scale, and guide the optimisation of gender equity in old commercial areas. It has a high measurement efficiency and broader application scope. However, this study is still limited in that the selected cases are all old commercial areas in the main city of Beijing, lacking attention to other urban areas of the city, and lacking detailed optimisation strategies. This is also a focus for future research.

References

Baidu Street View Map. (2022). Street View Images. Retrieved November 5, 2022, from https://quanjing.baidu.com/

Breiman, L. (2001). Random Forests. Machine Learning, 45(1), 5-32.

Chen, Y. H., & Wang, Y. Q. (2009). Approach to the Development of Traditional Urban Commercial Area:A Case Study of Xiamen. Arid Land Geography, 32(1), 152-158.

He, S. J., & Lin, G. C. S. (2015). Producing and Consuming China's New Urban Space: State, Market and Society. Urban Studies, 52(15), 2757-2773. https://doi.org/10.1177/0042098015604810

Huang, H. S., Lan, Y. B., Yang, A. Q., Zhang, Y. L., Wen, S., & Deng, J. Z. (2020). Deep Learning Versus Object-based Image Analysis (OBIA) in Weed Mapping of UAV Imagery. International Journal OF Remote Sensing, 41(9), 3446-3479. https://doi.org/10.1080/01431161.2019.1706112

Immitzer, M., Atzberger, C., & Koukal, T. (2012). Tree Species Classification with Random Forest Using Very High Spatial Resolution 8-Band WorldView-2 Satellite Data. Remote Sensing, 4(9), 2661-2693. https://doi.org/10.3390/rs4092661

Liu, Y., Yang, L. C., & Chau, K. W. (2020). Impacts of Tourism Demand on Retail Property Prices in a Shopping Destination. Sustainability, 12(4). https://doi.org/10.3390/su12041361

Long, Y. (2016). Redefining Chinese City System with Emerging New Data. Applied Geography, 75, 36-48. https://doi.org/10.1016/j.apgeog.2016.08.002

OSM website. (2022). Commercial District Boundary and Surrounding Road Networks. Retrieved November 5, 2022, from https://www.openstreetmap.org/

Peking University Geodata platform. (2022). Consumption Places POIs. Retrieved November 5, 2022, from https://geodata.pku.edu.cn

Yao, Y., Liang, Z. T., Yuan, Z. H., Liu, P. H., Bie, Y. P., Zhang, J. B., . . . Guan, Q. F. (2019). A Human-machine Adversarial Scoring Framework for Urban Perception Assessment Using Street-View Images. International Journal of Geographical Information Science, 33(12), 2363-2384. https://doi.org/10.1080/13658816.2019.1643024

SPATIAL ADAPTATION RULES OF EXTENSION BUILDING IN TRADITIONAL ARCHITECTURE

Case Study Of Batak Toba Houses And Their Extensions In North Sumatera, Indonesia

JONATHAN YOAS[1] and RIZAL MUSLIMIN[2]
[1,2] *Sydney School of Architecture Design and Planning, The University of Sydney, NSW, Australia.*
[1] *jonathan.yoas@sydney.edu.au, 0000-0001-9500-6533*
[2] *rizal.muslimin@sydney.edu.au, 0000-0002-1031-6135*

Abstract. This study evaluates the spatial configuration of existing traditional houses and their extension buildings to identify adaptation patterns applied by current generation occupants. As preserved by the community for their sacred values, transformation in traditional houses might seem slow, and an attachment to a new building could be disruptive. However, by tracing how the spatial configuration is being preserved, varied, and changed through computational modelling, this research shows that the space adaptation is more subtle than expected. The study formulated rules for traditional houses and their extension transformation using space syntax and shape grammar analysis. The rules simulation highlights consistent degrees of adaptation and evolution patterns in the four villages that were observed that were not immediately apparent visually. The adaptation pattern can be used as a reference for heritage conservation and future design precedence.

Keywords. Batak Toba, Traditional Architecture, Space Syntax, Shape Grammar, Digital Heritage

1. Introduction

Traditional houses typically have limited capacity to accommodate growing occupancy and current needs. Their spatial adaptation juggles sacred social and cultural values, economic capacity, profane demand, and the technical ability to knit them all together in bounded, given sites. As observed in the continuum of case study houses, design circumstances are challenging to reconcile. Recognising contributing characteristics in space arrangement remains critical for the sustainable regeneration of houses.

Our studies seek to identify intrinsic patterns (or genotypes) in traditional houses and extension building space configurations and devise rules for regenerating new configurations with similar characteristics. The rule creation employs a combination of space syntax and shape grammar approach.

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 747-756. © 2023 and published by the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA), Hong Kong.

Space syntax focuses on the structure and relationship of functional spaces by using justified graphs to represent networks of spaces, including those without permanent partition, and not solely rely on physical representation as in house plans diagrams.

Combining the intrinsic patterns derived from space syntax with shape grammar in this study offers a set of rules representing spatial adaptation for novel spatial arrangement regeneration.

2. Methods

Space syntax highlighted the study of networks and configuration of spaces intertwining with social characteristics of a place in what he termed "genotypes" (Hillier, 1984). By modelling house plans into graphs, with vertices representing rooms or areas accommodating activity anchored to a space, usually the yard, they represent how the space is structured (Hillier, 1984). The justified graph illustrates a living system by the inhabitants of a house. In a traditional community, this translates to how occupants live their lives as a group or their consensual lifestyle.

The research measured the houses' configurational characteristics, following space syntax mannerism, their depth (L), total depth (TD), mean depth (MD), relative Asymmetry (RA), and Integration (i) values, defined by the following expressions:

$$TD = \sum(0 \times nx) + (1 \times nx) + \cdots (x \times nx) \quad (1)$$

$$MD = TD/(K - 1) \quad (2)$$

$$RA = 2 \times (MD - 1)/(K - 2) \quad (3)$$

$$i = 1/RA \quad (4)$$

With:

K *Number of nodes (functional space)*

L *Depth, distances from carrier (exterior: yards or "alaman")*

TD *Total Depth*

MD *Mean Depth*

RA *Relative Asymmetry*

i *Integration*

Values of depths: number of depths, total depth, and mean depth *(L, TD, MD)* evaluates depth or distance between vertices to the anchor point. A 'deep' configuration, demonstrated by a higher value of *L*, means that a house dweller will have to go through more rooms before reaching a room in a higher depth index. An integration value *(i)* displays a connection to vertices indicating their commonality. The more 'public' a space, the higher the *(i)* value (Hillier, 1984). Patterns are achieved by comparing graph sets of traditional houses with extensions as a more recent adaption to the inhabitant's communal way of space use.

Rule-based analysis on vernacular architecture has been widely studied with Shape Grammar: inquiries of traditional Turkish and Taiwanese dwellings (Cagdas, 1996; Chiou & Krishnamurti, 1995) highlighted shape grammar applications by creating rules for plan generation. In Chinese construction rules, Li (Li, 2001) explored the use of grammar for educational purposes. This research employs space grammar to utilise characteristic spatial arrangement values as rules for configurational regeneration.

2.1.1. Case Study: Traditional Batak Toba Houses and Their Extensions

As a case study, we surveyed Batak Toba traditional houses and their extensions, spread across the Samosir and Toba Samosir regent, North Sumatra, Indonesia. The house type has been sufficiently documented from various perspectives (Boer, 1920; Sargeant & Saleh, 1973; Waterson, 2010). After the region's inclusion in Indonesia's top ten priorities for tourism development, it has been experiencing major construction work and revitalisation of traditional villages.

The study selected eighteen houses from four villages for our fieldwork observation. The number of villages represents a distribution of Batak Toba settlements around Lake Toba and Samosir island, following their migration path (Modigliani, 1892; Situmorang, 2009). To date, traditional houses in the region, some reaching more than one hundred years old, are still inhabited mainly with extensions almost immediately connected at the back. Their choice of material and construction method vary between villages, from timber, masonry, or hybrid. Most of these constructions are fragments or resemble almost nothing to the originals.

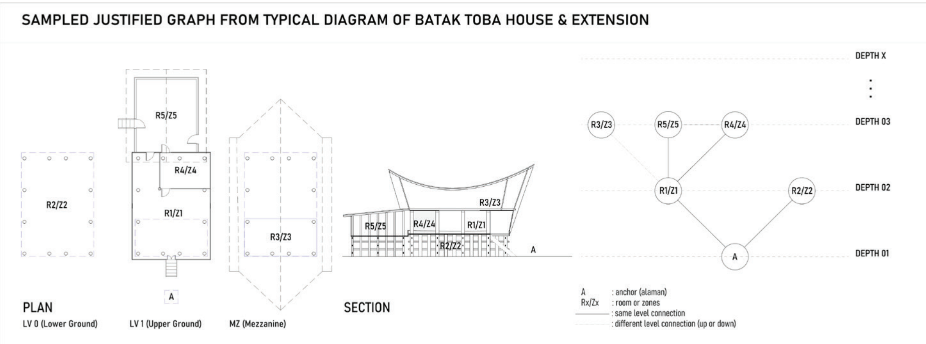


Figure 1 Sampled Justified Plan Graph of a Typical Batak Toba House and its Extensions.

2.2. SPACE SYNTAX ANALYSIS FOR CONFIGURATION PATTERNS

During fieldwork, this study created measured drawings of plans of existing traditional houses and their extensions. House plans are then modelled into a justified graph to observe their graph shapes, depth, and the character of vertices and edges in the graph (Figure 1). In a justified graph, the number of vertices represents the number of functional spaces in a configuration. Functional space is a whole or portion of a bounded space with designated activity to be used with or without a physical boundary.

2.3. RULE-BASED FORMAT FOR ANALYSING CONFIGURATION DEVELOPMENT

Findings from the field will be represented in rule format A $\rightarrow$ B, which consists of two labelled shapes, in this study, justified graphs labelled A and B. We use this rule format to modify a given shape graph into a new one, forming a network of space configurations. Rules can be applied to initial shape graphs and labelled shape graphs resulting from them. The process is iterative to achieve desirable space arrangement.

3. Patterns from Graph Comparison

3.1. PATTERN IDENTIFICATION

Our analysis of patterns identifies consistency in the case studies' justified graphs, as seen in the space measurement values. Here, graphs were redrawn and simplified to highlight regularity. Comparable identification steps were applied for original graph sets (Figure 2) and extension graph sets (Figure 3).

We analysed graphs for vertices and the number of edges (or "branches"), noting the functional spaces for their activities, and then grouped and re-labelled them based on activity types. The process is to achieve graphs' labelling parity for all houses. For instance, a room labelled "tataring", the traditional kitchen, functions more than just as a service area but also as a living and communal space: the hearth (Waterson, 2010). In grouping and re-labelling, we valued "tataring" in the same bracket as living-dining-kitchen functional spaces anywhere else in the justified graphs. We reiterate this process until we clarify consistency leading to pattern creation.

3.2. ORIGINAL GRAPH

3.2.1. Overall Graph Pattern

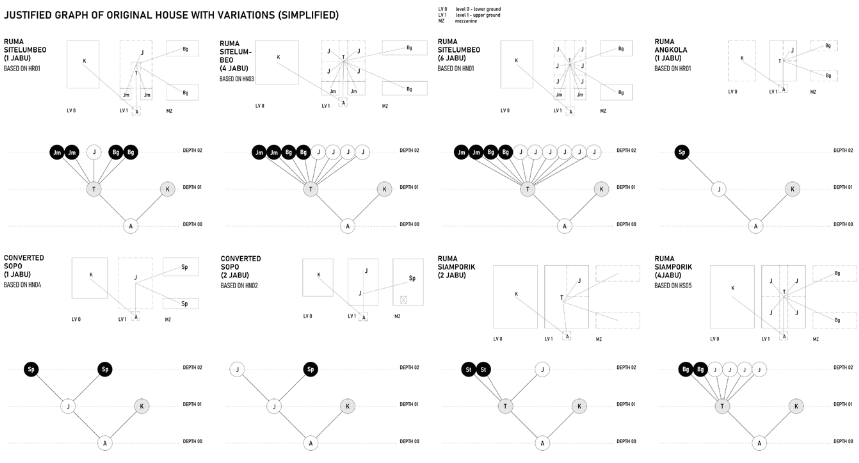


Figure 2 Originals Justified Graphs

Based on the analysis of the graphs, the basic configuration of a traditional Batak Toba houses can be explained as a network branching from an anchor into two primary vertices: the service space, "kandang", and the living spaces, passing through the "tataring" as the first node. From "tataring", it branches to multiple vertices of two groups of functional spaces: (1) living units, or the "jabu", and (2) storage spaces. A "kandang" originally is a buffalo shed or pigsty to keep their livestock, doubling as a pit latrine before the introduction of the water closet system (Boer, 1920; Modigliani, 1892). When aligning with other functional spaces, we identify these as service or storage spaces.

The basic pattern also represents elevational division, with "kandang" on the ground level, the living spaces on the upper ground level, and roof storage spaces doubling as a stage for ceremonies on the mezzanine level. Note that this division refers to three levels of worlds in Batak Toba's cosmology: the underworld, the middle world or the world of the living, and the upper world: the heavens. (Waterson, 2010)

3.2.2. Depth Consistency

In the original graphs, the average depth *(L)* is 2 for all the variations of available traditional houses (Table 1). Variations are in the number of functional spaces (or vertices, in the graph) at a second level (depth = 2). They reflect different amounts of "jabu" as living units, "jambur", their additional storage rooms in "sitelumbeo" house type, and "bonggar-bonggar", a mezzanine level common in all house types. The number of "jabu" in a house varies from 2 to 4 or 6 (Table 1). "Jambur" are usually two, flanking the entrance and "bonggar" also two, each in front and back of a house. In a converted "sopo", separating the living unit as "jabu" is not usual, with the main floor serving as one big living unit.

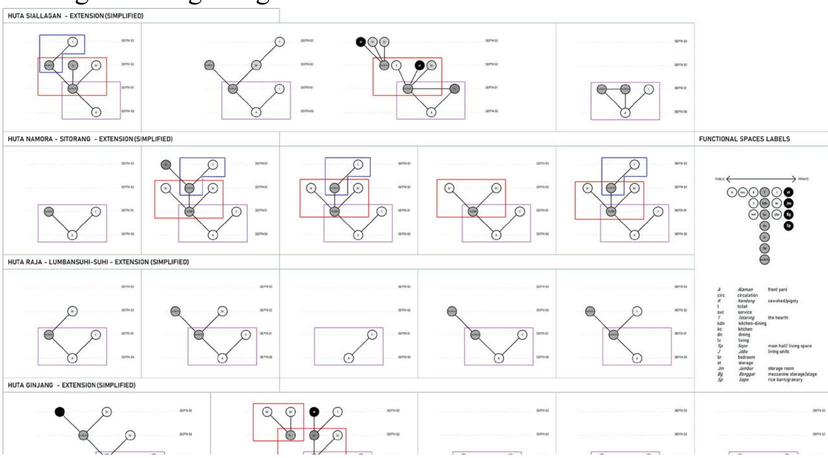


Figure 3 Extensions Justified Graph

Extensions with depth *(L)* = 1 and *(L)* = 2 dominate (amounting to 61%), with 5 and 6 houses from our sample possessing those configurational depths, respectively. The house with *(L)* = 3 and *(L)* = 4 is next, with the most depth measured in the configurations of our samples. Lastly, we found one house with *(L)* = 0 or no addition to the house.

Table 1 Space Configuration Analysis of Original Traditional Houses

	Analysis	*Sitelumbeo (1)*	*Sitelumbeo (4)*	*Sitelumbeo (6)*	*Angkola (1)*	*Conv. Sopo (2)*	*Conv. Sopo (1)*	*Siamporik (2)*	*Siamporik (4)*

K	Nodes (functional space)	8	11	13	4	4	5	6	9
L	Depth, distances "Alaman."	2	2	2	2	2	2	2	2
D0	nodes in depth = 0	1	1	1	1	1	1	1	1
D1	nodes in depth = 1	2	2	2	2	2	2	2	2
D2	nodes in depth = 2	5	8	10	1	1	2	3	6
TD	Total Depth	12	18	22	4	4	6	8	14
MD	Mean Depth	1.71	1.80	1.83	1.33	1.33	1.50	1.60	1.75
RA	Relative Asymmetry (of carrier)	0.24	0.18	0.15	0.33	0.33	0.33	0.30	0.21
i	Integration (of carrier)	4.20	5.63	6.60	3.00	3.00	3.00	3.33	4.67
		*Values in brackets ( ) denote the number of *jabu*							

3.2.3. Functional Spaces as Parallels between Originals and Extension

Typical functional rooms in the extensions are the kitchen, living room, dining room, bedrooms, and toilets. In many instances, living, dining, and kitchen may occupy the same room or zone. The shared nature of living, dining and kitchen spaces parallels the hearth or "tataring" role in the originals. "Tataring" space is also shared as the main circulation zone. Its placement suggests centrality to other functional spaces (Figure 2). Integrity (i) values of "tataring" reflect its characteristic accessibility (Table 1).

Service spaces like toilets, bathrooms, kitchens, or cowsheds are traditionally separated into different levels in the originals. In extension buildings, the separation is by allocating a service space to a different shed or an extension. A horizontal instead of vertical division. Even when they are visibly attached or located inside an extension, most service spaces have a separate entry from outside. Separating service spaces is typical throughout Austronesian houses and characteristically influences rule-based patterns (Figure 2 and Figure 3).

From our modelled graph, extensions configured from "alaman" as $(L)0$ are connected to either a living room, a dining room, a kitchen, or a service space in depth $(L)1$. Next, in the same depth level, various numbers of vertices function as bedrooms are found. We found the variation to this pattern in adding extra living space, usually at extensions with two-floor levels. These variations can be added at depth $(L)2$. Other functional spaces added from this depth level are a toilet or walk-in closet as storage for personal items. These are rare, and storage of this type is never used to keep harvests and produce goods.

4. Rules-Based Configuration Generation

4.1. RULES DEFINITION

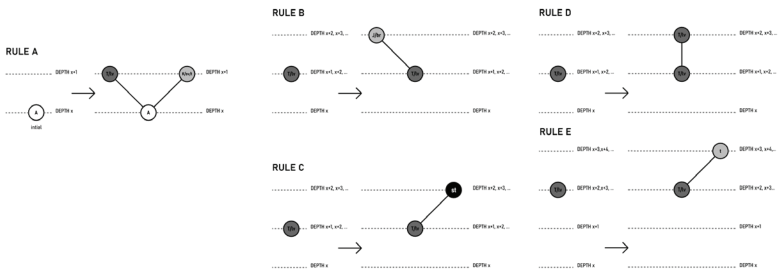


Figure 4 Rules based on patterns

Rules are presented in the form of justified graphs. Vertices in justified graphs represent functional spaces with edges connecting each other in a pattern that makes up a space network, following a characteristic of the pattern we found on Batak Toba houses and their extensions.

In justified graphs, graph rules of the Batak Toba are always drawn and progressed within its depth level $(L)x$. A rule starts from one functional space in a depth level as an anchor and continues by adding more functional spaces in the next depth level. It starts from $(L)0$, or $(L)x$ moving one step in depth sequence to $(L)x + 1, (L)x + 2, \dots (L)x + n$ consequently until all desirable depth acquired and possible functional spaces are exhausted.

First, the configuration's depth level must be set to apply the rule. The number of depths in Batak Toba houses follows the characteristic average depth levels of two, and it is multiple. Additions or complexity by variation within the rules is achieved by adding more functional spaces or repeatedly applying a rule according to their depth level. Figure 4 exhibits the rules graphs.

- Rule A branches the spaces from the anchor "alaman" into living space, service space, or zones (Figure 4). Patterns from originals show this beginning rule separating, vertically, the service spaces at ground level and living spaces at upper ground level. While in the extension buildings, the separation of service spaces is horizontal, in either two different sheds or a room within the extension with a separate service entrance. The rule is drawn departing from "alaman" in depth $(L)x$ into functional spaces in depth $(L)x + 1$.
- Rule B branches the communal living spaces, double as a circulation hub, into living units or bedrooms. The rule changes the depth level from $(L)x + 1$ to $(L)x + 2$. Extension bedrooms are configured surrounding the communal living space (Figure

2) with various amounts, mimicking multiple “jabu” (from two to six “jabu”) encircling “tataring”. The rules reflect continued alignment with the Batak Toba kinship social structure (Sargeant & Saleh, 1973).

- Rule C links a communal living space with storage areas. The number of these stores also varies according to the house size. In the originals, storage areas are located at the mezzanine level, called “bonggar-bonggar”. These are commonly used for harvest storage, just like their granary or “sopo”. During a ceremony, a front “bonggar-bonggar” seats musicians playing for the processions (see the section in Figure 1). It acts as a stage for the gathering public at “alaman”. Another form of storage is “jambur”, enclaves at the front, each at the right- and left-hand side of the entrance stair.
- Rule D with depth $(L)x + 1$, adds an extra living room from another living room in the previous depth level. This rule emerges from patterns found in extensions with multiple floor levels. The extra living room may be located at ground or upper ground level, aligned with “kandang” or with the main living hall originals. The addition of living rooms or application of Rule D allows for other rules that anchor on living spaces to be applied (Rule B and Rule C).
- Rule E in depth $(L)x + 2$ add an extra service space, e.g., an extra toilet or bathroom (Fig. 4) to a communal living space. This rule, applied at a depth level x, suggests that it only applies to an existing living space but not to the initial. Its application will add complexity to the configuration. It is a rare find among our sample houses as the process would inflict higher construction costs and adequate technical knowledge of waterworks and plumbing systems, a rarity in the area.

4.2. RULES APPLICATION AND ITERATIONS

Application of the rules follows the justified graph format in Space Syntax. We begin by assigning depth $(L)x$, with an anchor, a yard, to apply the first rule and continue to the next depth level. If we begin with a single, unattached configuration, then an anchor is at $(L)0$. The rules applied in the following sequence (Figure 5):

(1) Apply Rule A to start the configuration, resulting in an anchor at level $(L)x$ and two branches: “T/lv” (“tataring” or living room) and “K/sv” (“kandang” or services) at level $(L)x + 1$. This rule initiates the configuration with branch “K/sv”, a terminal vertex.

(2) Apply Rule B two times to add individual living spaces or bedrooms by adding vertices’ “J/br” (“Jabu” or bedrooms) to justified graph in depth $(L)x$+1 and so on. The rule can be re-applied each time one finds anchor “T/lv” in the appropriate depth level.

(3) Apply Rule C to anchor “T/lv” to add storage spaces by adding vertices’ “st” (storage) to justified graph in depth $(L)x$+1 and so on. The rule can be re-applied each time anchor “T/lv” exists in the appropriate depth level.

(4) Apply Rule D to anchor “T/lv” to add another communal living space by adding vertices T/lv (“tataring” or living room) to justified graph in depth $(L)x$+1 and so on. The rule can be re-applied each time one finds anchor “T/lv” in the appropriate depth.

(5) Apply Rule E to add a bathroom or toilet to an anchor of communal living spaces “T/lv”. The rule only applies to “T/lv” in-depth level $(L)x$+2, resulting in a

branch vertex in $(L)x$+3.

Rules can be applied and re-applied, adhering to the basic rule of depth level in graphs until the desired configuration is reached. The last iterations show how Rule B and C are applied at different depth levels. The resulting justified graphs should keep that shallow characteristic of graph set with grafted graph set can be added in the next depth level. Figures 5 (a) to (g) demonstrate the application of rules and their possible

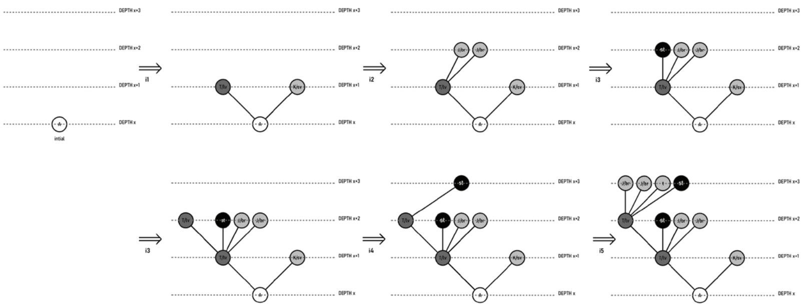


Figure 5 a-g Rule Iterations

iterations.

5. Conclusion

Revealing patterns by abstraction using graph and space syntax analytics assists our understanding of the basic social structure reflected in spatial configurations. By looking at not only the original but also extensions, one can identify the consistency or disruptions of an otherwise repeated nature of traditions. One outcome of this observation and experimentation is that rules can be established to formalise patterns for future spatial configuration regeneration, with the potential to be replicated in other communities or ethnic groups.

Configurational continuity does not translate into forms, constructions, or material. The study shows that the social structure represented in space configuration holds more extended longevity, even in different forms. With extensions more loosely reinterpreted; economy, material availability, and practicality are more of a driving force.

Worth to be discussed that the implementation of space syntax and rules should be based on sufficient social and anthropological data. The hybrid of data: geometry, social, and anthropology will assert meaning and understanding to the otherwise graphical representation and mathematical model.

Finally, the rule assists us in observing the trajectory of house-dwelling patterns, especially for a traditional community's ongoing adaptation to their way of living. One may regard the context of habitation in constantly changing culture and domination of worldwide trends or even tools for socially appropriate housing, more tailored to the dweller, to avoid generic solutions.

Acknowledgement

This study is supported by The Indonesia Endowment Fund for Education (LPDP), Ministry of Finance, The Republic of Indonesia.

Glossary

Alaman	The front yard of a house, forming part of a street or village square
Angkola	A house type with only two roof frames, marked with straight instead of angled walls
Bonggar-bonggar	A mezzanine floor is used as storage and a stage during ceremonies in a Sitelumbeo type.
Huta	A Batak Toba village.
Kandang	A service space below the main floor of Batak Toba house. Commonly used as buffalo shed or pigsty and pit latrine
Jabu	A single living space belonging to a family unit in a ruma.
Jambur	Two small spaces or rooms in front of a Batak Toba house are used as storage space.
Ruma	A traditional Batak Toba house.
Siamporik	A Batak Toba house type. A direct entrance from the front characterises it.
Sitelumbeo	Batak Toba house type is marked by an entrance of a stair from beneath, with no doors visible from the front elevation, and has complete elements. They are also referred to as "ruma bona".
Sopo	Rice barns.
Tataring	The hearth, traditional kitchen and communal space in a Batak Toba house

References

Boer, D. W. N. de. (1920). Het Toba-Bataksche Huis. G. Kolff. Retrieved from: https://catalog.hathitrust.org/Record/100175127

Cagdas, G. (1996). A Shape Grammar: The Language of Traditional Turkish Houses. Environment and Planning B: Planning and Design, 23, 443–464. https://doi.org/10.1068/b230443

Chiou, S.-C., & Krishnamurti, R. (1995). The Grammar of Taiwanese Traditional Vernacular Dwellings. Environment and Planning B: Planning and Design, 22(6), 689–720. https://doi.org/10.1068/b220689

Domenig, G. (2004). Consequences of functional change Granaries, granary dwellings, and houses of the Batak Toba. In Indonesian houses (pp. 61–97). Singapore University Press.

Hillier, Bill. (1984). The Social Logic of Space. Cambridge University Press.

Li, A. I.-K. (2001). A Shape Grammar for Teaching the Architectural Style of the Yingzao Fashi. Retrieved from: https://dspace.mit.edu/handle/1721.1/8631

Modigliani, E. (1892). Fra i Batacchi Indipendenti: Viaggio. Società Geografica Italiana. https://catalog.hathitrust.org/Record/008389634

Sargeant, G., & Saleh, R. (1973). Traditional Buildings of Indonesia Volume I: Batak Toba (Vol. 1). Regional Housing Center; Perpustakaan Kementerian Pekerjaan Umum dan Perumahan Rakyat.

Situmorang, S. (2009). Toba Na Sae: Sejarah Lembaga Sosial Politik Abad XIII-XX. Komunitas Bambu.

Waterson, R. (2010). Living House: An Anthropology of Architecture in South-East Asia. Tuttle Publishing. http://ebookcentral.proquest.com/lib/usyd/detail.action?docID=931457

EL RETABLO DIGITAL (THE DIGITAL ALTARPIECE)

Restoration of Mortuary Rituals with Deep Learning

GIOVANNA PILLACA[1] and DANIEL ESCOBAR[2]
[1]OLA.Research [2]OLA.Research.
[1]giovanna.pillaca@urp.edu.pe, 0000-0002-5980-5853
[2]descobar01490@gmail.com

Abstract. The Digital Altarpiece is an immersive virtual temple that seeks to communicate a traditional funeral ritual from Ayacucho, Peru, and revive it in a contemporary environment. Due to the confinements in times of pandemic, these Peruvian Andean rituals of social integration were in danger of disappearing. The ancestral knowledge in humanity has a cultural and spiritual value. Through this project, the theory of dehumanization due to the excess of new technologies and the preservation of cultural memory using digital tools and machine learning are questioned. For this, digital memories were collected, photographs of physical spaces openly uploaded to the web by their authors. This data was extracted to learn features using GANs or Generative Adversarial Networks. The intention of using photographs from social networks was to reinterpret and discover a new local and collective perspective. Although artificial intelligences have the power to excite us with new figure proposals, we humans are experiencing the sensations of these pieces, and we can direct our technical explorations with new technologies that change so quickly, to cultural integration.

Keywords. Digital Heritage, Machine Learning, Architecture, Social Networks, Digital Artefact.

1. Introduction

Computational tools are evolving the ways we collect, store and inhabit memories. The digital altarpiece is a research into what it means to preserve memories of rituals and places in an immersive 3D virtual environment. By studying traditional craftsmen in Ayacucho, Peru making physical altarpieces, their analog methods are reinterpreted through digital tools. The proposal leverages publicly available images from social media to extract features using machine learning. These features are transformed to 3D architectural elements that are assembled into a digital temple deriving its organization from mortuary rituals. It concludes in a project preserving cultural tradition as a form of digital heritage.

Varinlioglu and Balaban exploit the use of computational tools to reconstruct social networks of the past using deep learning to simulate movements of caravans and

HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2023, Volume 1, 757-766.

automate heritage site prediction. Heritage has gained its place in artificial intelligence studies as a dataset of humankind that can be copied and learned from (Varinlioglu et al., 2021).

Significant digitization projects such as the EU funded “Time Machine Project" is creating a big dataset of European history and culture (TMP 2019). It represents the creation of a European memory in the digital. According to Addison digital heritage as an emergent discipline is made up of three key processes of 3D: documentation, representation, and dissemination including immersive networked worlds to augmented reality (Addison 2000). The development of AI can expand Addison’s definition for automating certain tasks and uncovering new paradigms by running simulations (Varinlioglu et al., 2021).

The following section describes recent case studies of digital heritage in architecture and background of traditional regional rituals in Ayacucho, Peru that inspired the creation of a traditional altarpiece into a digital architectural temple.

2. Background

The creation of temples serves as a way to establish permanence in the passing of time and cultural traditions. Understanding regional iconography as observed through the lens of visitors and locals can provide a sense of culture. Defining what makes up culture is particular to an area that only through experience can be fully understood. Thus, it is possible through digital methods to create and preserve a tradition that is emotionally and culturally relevant to humans in an immersive virtual environment.

2.1. DIGITAL HERITAGE IN ARCHITECTURE

Appropriating Addison’s definition of digital heritage, recent works can be categorized into documentation and analysis, representation and dissemination, and applied artificial intelligence.

2.1.1. *Documentation and Analysis*

The use of photogrammetry has proven useful for reconstruction of heritage sites through crowdsourced photographs. The Project Mosul seeks to digitally reconstruct the lost heritage, whether through war, conflict, natural disaster, or other means, and preserve its memory through digital preservation schemes (Vincent et al. 2015). This highlights a framework for managing crowdsourced reconstructions which can extend to other various purposes.

2.1.2. Representation and Dissemination

Using a game engine, Unity, Ubisoft restored Notre-Dame de Paris cathedral, which then became a digital building in the game Assassin's Creed (Gombault, 2020). Taking a similar approach, the digital Teos Project is an interdisciplinary research project investigating and digitally animating the architectural heritage of Teos. The project includes digital fabrication of historic buildings, a virtual reality tour, augmented reality, and smart phone game application (Varinlioglu et al. 2021).

2.1.3. Applied AI

Barcelo (2008) argues that computers can perform analysis acting as automated archeologists. An attempt to predict missing portions of an ancient Greek transcript uses Pythia, which is a deep neural network that makes 20 different suggestions for missing words. The best suggestion is meant to be chosen by subjective domain knowledge (Assael et al., 2019). Similarly, a proposal for the reconstruction of 2D wall paintings from fragments uses a genetic algorithm, which allows a larger and more consistent reconstruction than a local assembly (Sizikova and Funkhouser, 2017). These methods focus on puzzle solving and reconstructing missing artifacts by employing deep learning algorithms for cultural archaeology and preservation.

2.2. TRADITIONAL MORTUARY RITUALS

Rituals consist of choreographed actions and symbols that are encoded as cultural knowledge. When a person dies the rites are organized by their community to mark an important event and memory. In the family it serves to accompany the mourners, comfort them and offer condolences. Each regional culture interprets this differently.

2.2.1. The Ritual in Huamanga, Ayacucho, Peru

The traditional Andean mortuary commemorates our life. As of 2019 this ritual is a hybrid between the indigenous culture and the Catholic religion indoctrination. The mortuary ritual is a community meeting space where the deceased is the protagonist, evidencing networks of family connections that pay tribute through songs, prayers, flowers, food, drinks, candles, etc. In a procession, the local streets and architecture memorable for the deceased is experienced, symbolizing the farewell to the earthly world.

Figure 1. Ayacucho Location on the left. B. Traditional Retablo in the center. 'Ancestor and Future' digital artwork on the right.

2.3. THE ALTARPIECES OF AYACUCHO

The traditional altarpieces have their origin in colonial times when Spanish priests in the process of evangelization toured all the towns of the Peruvian highlands. They carried articulated boxes with images of Catholic saints so that they would be recognized by the neighbors. These were called "Cajas de San Marcos" and centuries later they were taken as a reference by artisans from Ayacucho to make the altarpieces. During the 1940s, the creation of these works of art began to design scenes with themes related to local motifs (bullfights, cockfights, festivals and traditional dances), with rural scenes and agricultural work. (Ulfe 2004).

The handmade altarpieces become the iconographic representation of socially relevant events to narrate a cultural memory. Artisans use photographic memory to create sculptural pieces inside a box whose doors, when opened, present drawings of flowers and architectural elements such as doorways, openings, levels, and vegetation. See Figure 1.

2.4. RESPONSE TO A RELEVANT SOCIOCULTURAL EVENT

During 2020, when the highest number of deaths was recorded worldwide, rituals that summoned families were prohibited. The meeting and farewell space was broken. This makes us think of new ways to say goodbye to our dead. The altarpieces and their traditional construction processes by the artisans inspire the proposal of a digital altarpiece. Becoming a virtual temple where you can connect with other family members and maintain ancestral mortuary rituals. And for this, the Andean worldview must be understood.

In the Andean culture the term Upamarca is used to refer to the passage to another world. Hannan Pacha, like the world above, the sky; Ukupacha to refer to the world below; and Kai Pacha to refer to the world of the center of the earth as narrated by Garcilaso de la Vega. However, we do not know what these places are. And, without getting into philosophical debates, there is the possibility of free will to decide where to go.

2.5. COLLABORATING WITH AI

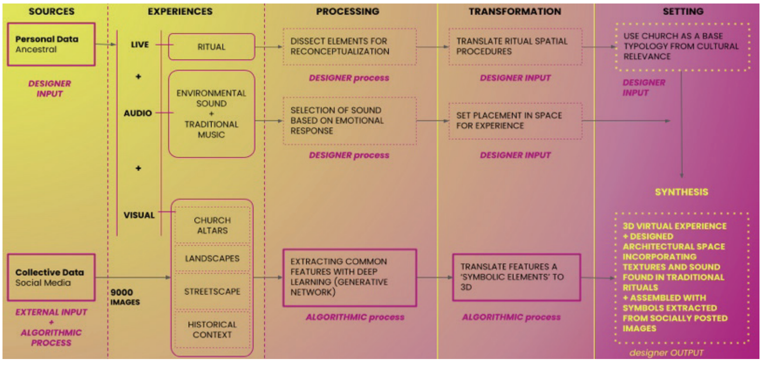


Figure 2. Workflow of 'The Digital Altarpiece'

The background research shows the current approaches using digital computational tools for digital cultural heritage. AI approaches require access to datasets and computation resources for training of models. Crowdsource data is often preferred when no available dataset. Most methods require the supervision of a domain expert. This makes the case for collaborative methods of human subjective experience alongside machine intelligence. The project aims to transport the visitor to a virtual reality, like a digital hybrid architectural restoration, inspired by the Ayacucho altarpiece to activate memories and reflect on transmutations of its people. The methods consist of 4 stages: Documentation and Analysis, Application of Artificial

Intelligence, Collaborative Synthesis; and Representation and Dissemination. See Figure 2.

2.6. DOCUMENTATION

2.6.1. Types of Data

Three types of datasets are established for building a regional identity of place.

2.6.2. Social Media Dataset

Using crowdsourced images creates an unbiased recollection of memory which serves to understand what people find important enough to share online. A total of 9,000 images are mined from Flickr and Instagram using their API. A geographical radius is established at 10 miles from the city center of Ayacucho to filter out images. The data range for the images is set from 2012 to 2021. The following relevant hashtags are used to search: #AYACUCHO #RETABLOS #IGLESIASAYACUCHO #LANDSCAPEAYACUCHO.

2.6.3. Historical Dataset

Contextual images from the 1960s produced by the photographer Baldomero Alejos, who portrayed families from different social strata of Ayacucho society in a time before the civil war, are also used. These 150 photographs convey a reality of gender roles and racism. This dataset aims at integrating local history by making it accessible to the people of the Andes.

2.6.4. Personal Dataset

In 2019, the author participated in their grandmother's funeral rite. During the experience the importance of revaluing Andean culture was made evident. Elements such as sounds, visuals, and architectural space are understood as necessary for an immersive experience. In the representation step, these elements are composed to generate an immersive environment.

2.7. APPLIED AI

Stylegan2 and CycleGan are used to extract and remap visual features. These are generative deep learning algorithms. They learn a representational distribution of a ground truth dataset through unsupervised learning.

Since our ground truth data has high variance, this approach serves to find commonality in visual features by relating them in a searchable latent space.

2.7.1. Feature Extraction

StyleGAN2 generates realistic images of faces that resemble the original dataset (Karras, 2020). The 9,000 crowdsourced images are used to learn a representation of regional site features. Using visual inspection during the training process the generative quality is assessed. When quality is determined as coherent, the latent space is investigated by selecting a set of feature vectors representing points in the distribution. A set of paired images selected by the author are feedforward into the StyleGAN to create an embedding that is a representation of the images mapped to a learned latent space. An interpolation is made between the two embeddings and outputs a resulting set of frames showing how features from two images are related. This process is repeated several times to find commonalities between images. With this method a natural landscape and architectural building dataset is created to represent the extracted features of the site. See Figure 4.

CycleGAN is able to learn from an unpaired set of images and map features to input images (Zhu, 2020). This algorithm is used to map a set of unrelated features from two generated datasets from the previous process. By forcing the CycleGAN to learn a mapping between them, the predicted results display the distillation of architectural unique elements and placing natural landscapes in new relationships. This provides a visual discriminative output of learned features that are then used to process 2D architectural elements into 3D shapes in a later step. See Figure 3.

Domain expertise is used as a way to guide the processes to the intended result while allowing for the discovery of new elements that might be encountered during the outcome.

Figure 3. First image iteration from data set extracted online.

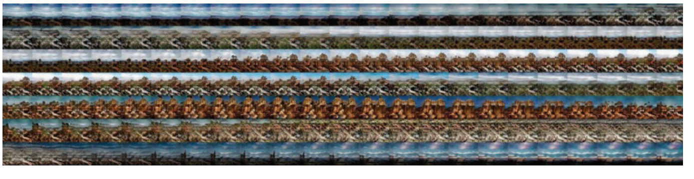


Figure 4. Images from the StyleGAN process and new images from the CycleGAN process

2.8. COLLABORATIVE SYNTHESIS

In the following method, domain expertise is used to form a synthesis of the previous processing steps with AI. The traditional Altarpiece of Ayacucho has the particularity of almost invariable elements such as the cover, the openings, levels, vegetation, details and nuances. These are considered as categories that will serve as the basis for classifying the generated images from the previous process.

2.8.1. Data Analysis

Similarly, the generated images are categorized into: covers, openings, levels, vegetation, details and nuances. The covers of the temples, usually made of gray stones, are bathed in gold. The openings that contain sculptures of saints are built as open openings towards the sky. The levels of stairs and environments are evident as isolated and focused elements. The vegetation appears as a pattern that invades levels, openings and portals. Interior and exterior details are evident with changes in material and contrast of figure and background. The material hues or textures change from stone to gold, from gray to reds. See Figure 5.

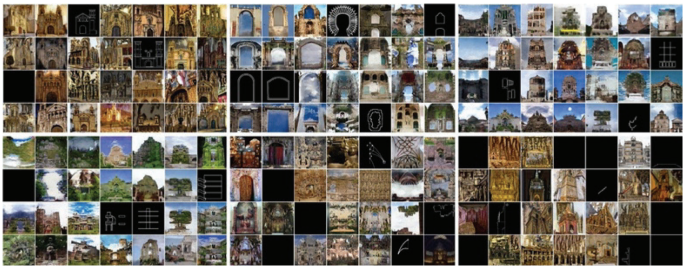


Figure 5. Images curated by designers to create 3D pieces.

2.8.2. Data Transformation

The next collaboration in human and AI, begins with selecting images generated by the GAN's models within each category. These images have to present clarity in the profiles of visual features to delineate vectors on the images. Then, a Sobel filter is applied to images to extract edges and lines. These edges can be vectorized using Adobe Illustrator. To go from 2D to 3D elements, the vectorized edges are transformed using basic operations: extrusions, booleans, revolutions, scaling, rotations in Rhino3D automated with grasshopper. This results in a catalog of 3D architectural elements. See Figure 6.

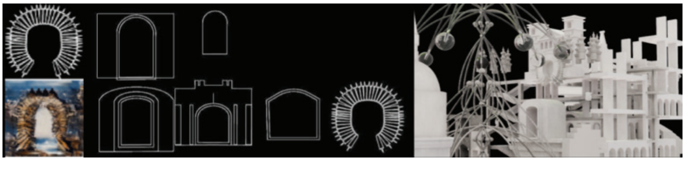


Figure 6. Images to vectors extracted from selection of patrons in GAN's. To create 3D Geometries for architectural spaces.

2.8.3. Composition of 3D (Digital artisanal craftsmanship)

The 3D elements are organized like a colonial temple. Circulation is developed as an interpretation of the Andean Worldview of three levels: the world above, the middle, and below. The 3D results are architectural artifacts that with an aggregation logic are used as unique pieces for the composition of the project.

The digitally architecture uses the typology of a church to represent its importance in social customs during death, wakefulness and burial rituals.

Texture integrates material patterns observed in the images extracted from the resulting GANs, where particularly vegetation, gold and stone take center stage and mix with each other.

2.9. REPRESENTATION AND DISSEMINATION

2.9.1. Augmented Reality, Virtual Environment and Web

This research classifies and re-conceptualized categories of the Ayacucho Altarpiece with Deep Learning processes. For virtual reconstruction, beautifully rich spaces are unfolded in the matrix of spatial results compiled into a narrative scene to be experienced in Augmented Reality, Virtual Reality, and web.

The web experience environment and Virtual Reality is developed in Unity. The element´s materials is based on the categorization of images. The final file is optimized and compressed to be uploaded to the web, where it is possible to walk through the project as first and third person. Virtual Reality is configured for Oculus. This experience is first person based. See Figure 7.

Figure 7. Section cut with Hannan Pacha (the world above), Ukupacha (purgatory) and Kay Pacha (the world in the center of the earth). VR and Video Game Environments images.

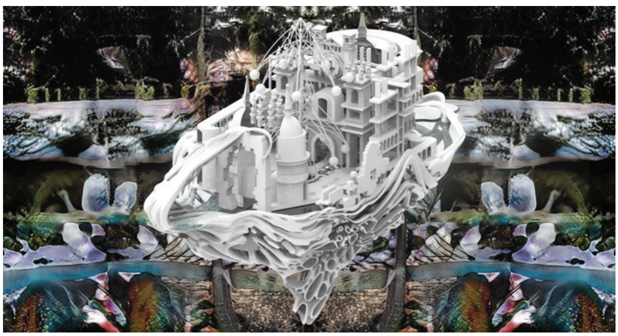


Figure 8. Digital Altarpiece, created to AR experience.

Augmented Reality (AR) is done with Sketchfab. For this type of visualization, the geometry is optimized and texturized as a single monochrome material to speed up the opening of any type of digital device. See Figure 8.

To activate the sensory-spatial memory in the AR and web experience, bright moving elements and surround sound are included. The sound is a strategy to guide the visitor on the tour and invite the user to have a certain degree of privacy in virtuality. The sounds are worked with real recordings of local bands in the procession and are inserted in two areas of the virtual space, one close to the entrance and another at the highest point.

2.9.2. The 3D printed artifact

The transition from the digital altarpiece to the physical medium materialized in 168 hours of 3D printing and assembly. The elements are printed in 3 colors of gold, white and natural PLA material to differentiate metallic, tectonic and organic elements. See Figure 9.

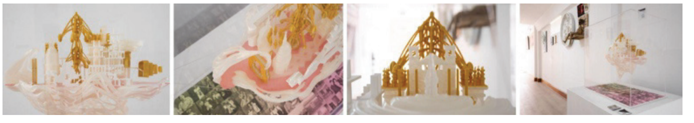


Figure 9. Images of 3D Printing Digital Altarpiece

3. Results and discussion

The Digital Altarpiece proposes to expand personal memory to a collective memory, reduce the limitations for face-to-face attendance and promote the use of technologies to preserve cultural identity.

The confluence of personal and collective memory to recreate a single identity approach is a key aspect. Experiential and sensory data allowed us to rethink how virtuality allows generating participatory work. Subsequently, one of the main achievements is linking the Ayacucho culture in an immersive experience for technological users and connecting the Ayacucho population through experiences with familiar elements and language.

4. Future works

This work focuses on unique users. However, the idea of multiple users in a metaverse broadens the horizon of the project. In addition, more advanced AI methods would speed up the creation of custom digital altarpieces with higher fidelity in rebuilding from real to digital.

5. Conclusion

Tangible artifacts and traditions serve to preserve cultural heritage. The Digital Altarpiece contributes by reinterpreting how we produce, store and disseminate rituals, adapting to a contemporary context. Using technologies that emulate traditional analogue knowledge and methods of artisanal and artistic creation. The project is a proposal for the hybridization of cultural/historical preservation with technology and the construction of the world through a design proposal. The project makes a contribution to the field of computational design and cultural appreciation.

Acknowledgements

I would like to thank the ancestors, the Andean knowledge, rituals and traditions which still live nowadays. Thanks to the grandmothers Felicitas Rivera and Melchorita

Garibay who is the principal inspiration and motivation to do the project. And thanks to all the women who work in AI creatively.

References

Addison, A. C. (2000). Emerging Trends in Virtual Heritage. IEEE Multimedia, 7(2), 22-25. https://doi.org/10.1109/93.848421

Assael, Y., Sommerschield, T., & Prag, J. (2019). Restoring Ancient Text Using Deep Learning: A Case Study on Greek Epigraphy. Empirical Methods in Natural Language Processing (EMLP), https://doi.org/10.48550/arXiv.1910.06262

Barcelo, J. A. (2008). Computational Intelligence in Archeology. Information Science Reference. DOI: 10.4018/978-1-59904-489-7

Gombault, A. (2020). Notre-Dame Is Burning: Learning from the Crisis of a Superstar Religious Monument. International Journal of Arts Management, 22(2), 83-94

Karras, T., Laine, S., Aittala, M., Hellsten, J., Lehtinen, J., & Aila, T. (2020). Analyzing and improving the image quality of stylegan. arXiv.org. Retrieved December 12, 2022, from https://doi.org/10.48550/arXiv.1912.04958

TheTime Machine Project (2019). Retrieved from https://www.timemachine.eu.

Ulfe, M. E. (2004) El Arte de los Retablos Ayacuchanos: Religiosidad, Historia y Práctica Cultural Emergente. In F. Armas Asin, In Angelus Novus. Prácticas evangelizadoras, representaciones artísticas y construcción del catolicismo en América, siglos XVII-XX. Lima:Fondo Editorial de la PUCP.

Sizikova, E., & Funkhouser, T. (2017). Wall Painting Reconstruction Using a Genetic Algorithm. Journal on Computing and Cultural Heritage, 11(1), 1-17. https//doi.org/10.1145/3084547

Varinlioglu, G., Balaban, O. (2021) Artificial Intelligence in Architectural Heritage Research: Simulating Networks of Caravanserais through Machine Learning. In I. As and P. Basu, In The Routledge Companion to Artificial Intelligence in Architecture. (pp 207-223). Routledge. https://doi.org/10.4324/9780367824259

Vincent, M. L., Gutierrez, M. F., Coughenour, C., Manuel, V., Bendicho, L. M., Remondino, F., & Fritsch, D. (2015). Crowd-sourcing the 3D Digital Reconstructions of Lost Cultural Heritage. In 2015 Digital Heritage (Vol. 1, pp. 171-172). Piscataway, NJ: IEEE. https//doi.org/10.1111/j.0031-868X.2004.00278.x

Zhu, J.-Y., Park, T., Isola, P., & Efros, A. A. (2020). Unpaired image-to-image translation using cycle-consistent adversarial networks. arXiv.org, https://doi.org/10.48550/arXiv.1703.10593.

INDEX OF AUTHORS

In the 21st century, global immersion in technology has connected the world like never before, and technological impact is reaching even some of the most remote and impoverished areas of the world. As a consequence, the advancement and adoption of new technologies in all aspects of our society have profoundly changed how we design our environment to be healthy, liveable, and equitable.

Computational design, simulation, analysis, fabrication, and management allow us to evaluate and forecast the performance of habitats with a deeper understanding of contexts.

Technology for social good is an agenda to articulate values, behaviours, and attitudes that focus on social impact – placing people at the focal point. In the context of a rapidly changing climate, it has become evident that we are approaching a moment of no return; a moment in which our fate hangs in the balance. We as a species, and the societies we construct, must take immediate action and work collectively towards a brighter and more sustainable future.

CAADRIA2023 (Human Centric) invited contributions addressing the conference theme by discussing the influence of technology on society, economy, environment, and governance pertaining to habitats where people are at the heart of the construction industry. We invited conversations and debate around key questions:

How do different computational design approaches contribute to the design, development, and policy-making of complex societies?

To what extent do automation, machine intelligence, and control of complex systems impact societal development?

How do the practical and real-world applications of computational research and technologies advance broader social well-being?

ISBN: 978-988-78917-9-6

9 789887 891796

www.caadria2023.org

ISSN: 2710-4257 (PRINT); 2710-4265 (ONLINE)