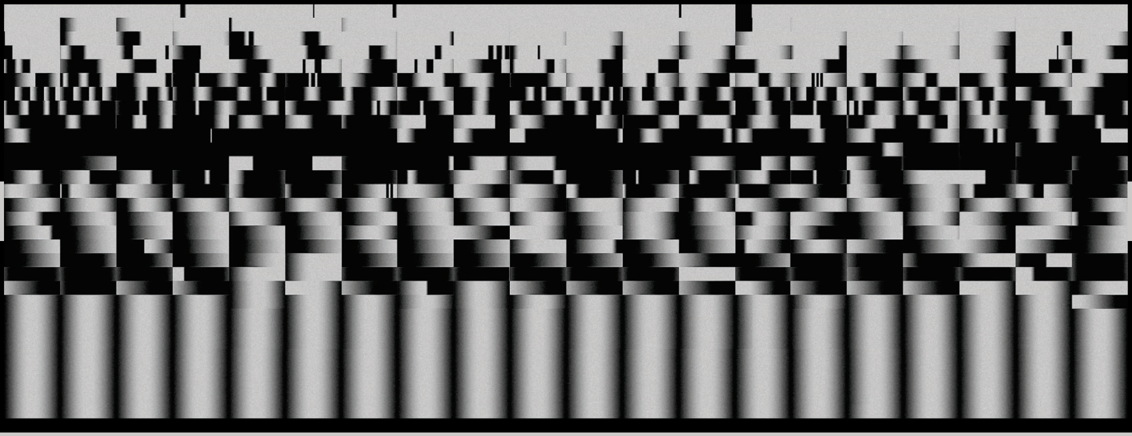


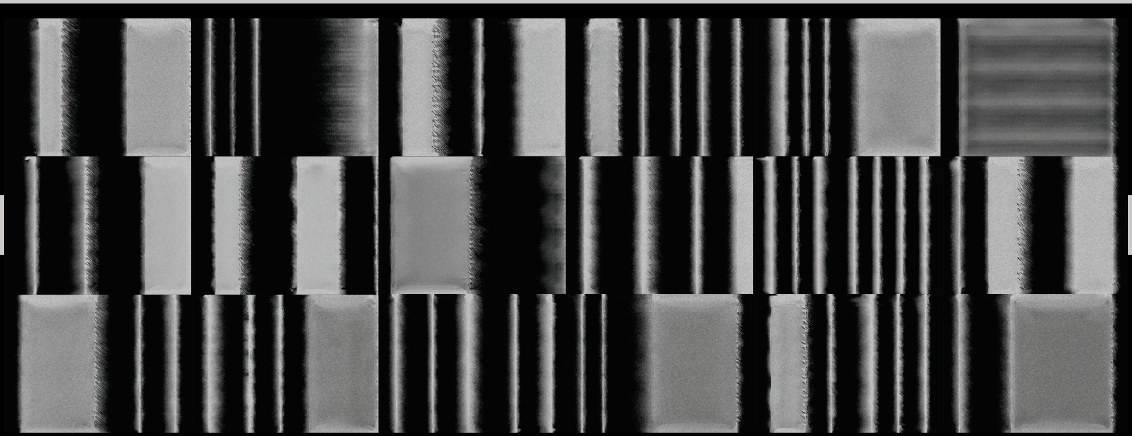
↗ INPUT INDEX: #000 — #600 COLLECTED 05/10/2023 — 26/02/2024

CAADRIA '24

20 — 26 APRIL 2024 VOLUME 1 SINGAPORE, SG

ACCELERATED DESIGN

↘ OUTPUT INDEX: #000 — #018 SAMPLER: EULER A



ACCELERATED DESIGN

Proceedings of the 29th International Conference on Computer-Aided Architectural Design Research in Asia (CAADRIA 2024)

Volume 1

Edited by

Nicole Gardner
University of New South Wales, Australia

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

Likai Wang
Xi'an Jiaotong - Liverpool University, China

Hirano Toshiki
The University of Tokyo, Japan

Sumbul Ahmad Khan
Singapore University of Technology and Design, Singapore

Accelerated Design
29th International Conference on Computer-Aided Architectural Design Research in Asia (CAADRIA 2024)

20-26 April 2024

Hosted By:

Singapore University of Technology and Design,
SINGAPORE

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

ISBN: 978-988-78918-1-9
ISSN: 2710-4257 (print)
2710-4265 (online)

Printed in Singapore, SINGAPORE

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 three volumes, contain the research papers that were accepted for presentation at the 29th International CAADRIA Conference, organised by Singapore University of Technology and Design, Singapore. 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 203 reviewers from 33 different countries working across the globe. It marks another year of growth 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 2023 resulted in a record high stage 1 submission of 541 abstracts from 37 countries, of which 397 were invited for further development as full paper submissions. For the stage 2 double-blind review, at least 2 international reviewers were assigned to each of the submitted full papers. In February of 2024, following the reviewers' recommendations and the Paper Selection Committee's evaluation, 159 papers were accepted for inclusion in the publication and presentation at the conference.

We thank all authors for their research contributions and congratulate them on their publication achievement. We extend our sincere thanks to the International Review Committee members who volunteered their valuable time and expertise. We also thank the hosts at the Singapore University of Technology and Design for organising the 2024 CAADRIA Conference and developing a thought-provoking, creative, and engaging theme and schedule of events. Special thanks are offered to Gabriel Wurzer for his patient, generous, and expert contribution to the CAADRIA proceedings production. Finally, we thank the CAADRIA organisation for providing the opportunity and honour to serve on the Paper Selection Committee for the CAADRIA 2024 conference.

PAPER SELECTION COMMITTEE

Nicole Gardner	*University of New South Wales (Chair)*
Christiane M. Herr	*Southern University of Science and Technology*
Likai Wang	*Xi'an Jiaotong - Liverpool University*
Hirano Toshiki	*The University of Tokyo*
Sumbul Ahmad Khan	*Singapore University of Technology and Design*

Theme: ***‘Accelerated Design’***

Can and should design be accelerated? To what extent and to what end? Or perhaps, design must decelerate instead? Accelerated Design is an urgent call for a critical reflection of and creative action by architecture during this challenging time of accelerating climate crisis, unrestricted data surveillances, generative AI copyright infringements, global geopolitical conflicts, hyperconcentration of digital power, post-pandemic mental health deterioration, and widespread disinformation attacks.

CAADRIA2024 seeks contributions in addressing the conference theme by discussing and debating the role of design and designers in the midst of accelerated changes brought about by and on technology, economy, environment, and governance, to construct new ways of thinking, teaching, researching and practising architecture in the age of artificial intelligence and climate change.

The 29th 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 towards an inclusive future for humans and non-humans. Contributions focusing on the Asia / Pacific context are particularly encouraged.

CAADRIA 2024 invited 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,

Immanuel Koh	*Singapore University of Technology and Design (Chair)*
Jianxi Luo	*Singapore University of Technology and Design*
Jeffrey Chan	*Singapore University of Technology and Design*
Sumbul Ahmad Khan	*Singapore University of Technology and Design*
Elissa Hartanto	*Singapore University of Technology and Design*

{https://caadria2024.org}

About CAADRIA

CAADRIA - The *Association for Computer-Aided Architectural Design Research in Asia* promotes teaching, research and innovative practices in Computer-Aided Architectural Design (CAAD) and beyond, spanning across the larger Austral-Asian and Pacific regions with a global membership across six continents.

CAADRIA was founded in 1996 with the following key 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.

Among other initiatives, CAADRIA organises an annual conference. The first of which was held in 1996 in Hong Kong, where the association is registered to date. Since then, 28 conferences have been held in India, Australia, China, Hong Kong, Japan, Korea, Malaysia, New Zealand, Singapore, Taiwan, and Thailand. The annual CAADRIA conferences provide an opportunity to meet each other and learn about the latest research in the field. CAADRIA is a growing community with many motivated volunteers who generously give their time and energy to support the organisation in various roles. One of CAADRIA's missions is to engage with emerging researchers (especially PhD students), supporting and expanding such initiatives as our Postgraduate Consortium and the World CAAD PhD workshop.

CAADRIA fosters long-standing bonds with its sister organisations: CAAD Future Foundation (CAAD Futures), Education and Research in Computer Aided Architectural Design in Europe (eCAADe), Sociedade Ibero-Americana de Gr áfica Digital (SIGraDi), Association for Computer-Aided Design in Architecture (ACADIA), and Arab Society for Computer-Aided Architectural Design (ASCAAD).

The 29th conference, in 2024, is hosted by the Singapore University of Technology and Design, Singapore. CAADRIA 2024 is held as a face-to-face conference, continuing the Association’s mission to bring together researchers, practitioners, and schools of the Asia-Pacific region. CAADRIA 2024 is themed ‘Accelerated Design’ and aims to bring together academics, researchers, and practitioners to contribute to the fields of computational design methods, instruments, and processes towards an inclusive future for humans and non-humans.

All CAADRIA conference proceedings are indexed in CumInCAD - a Cumulative Index of Publications About Computer-Aided Architectural Design. The proceedings are available both online and in research libraries around the world. CAADRIA is one of the four founding organizations of the *International Journal of Architectural Computing* (IJAC) and supervises one issue each year. IJAC is published by Sage Publications in both paper and electronic versions.

Anastasia Globa
President, CAADRIA

CAADRIA Officers

President: Anastasia Globa
University of Sydney, Australia

Secretary: Urvi Sheth
Craft Quest, 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

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

Administrative Officer: Marc Aurel Schnabel
Xi'an Jiaotong - Liverpool University, China

Outreach Officer: Ding Wen 'Nic' Bao
RMIT University, Australia

Organisation Officer: Likai Wang
Xi'an Jiaotong - Liverpool University, China

Communication Officer: Chitraj Bissoonauth
Xi'an Jiaotong - Liverpool University, China

CAADRIA 2024 Conference Committees

Hosting Institution:
Singapore University of Technology and Design
SINGAPORE

Organising Committee:

Immanuel Koh	*Conference Chair*
Jianxi Luo	*Conference Coordinator*
Jeffrey Chan	*Conference Coordinator*
Sumbul Ahmad Khan	*Conference Coordinator*
Elissa Hartanto	*Conference Coordinator*
Stylianos Dritsas	*Workshop Coordinator*
Jason Lim Teck Chye	*Exhibition Coordinator*

Elissa Hartanto	*Tour Coordinator*
Suryakala Mahankali	*IT services*
Angeline Li	*Communication team*
Evelin Tay	*Communication team*
Artificial-Architecture	*Exhibition Design*
Practice Theory x Artificial-Architecture	*Graphics Design* *Website Design*

Paper Selection Committee:

Nicole Gardner — *Chair*
University of New South Wales, Australia
Christiane M. Herr
Southern University of Science and Technology, China
Likai Wang
Xi'an Jiaotong - Liverpool University, China
Hirano Toshiki
The University of Tokyo, Japan
Sumbul Ahmad Khan
Singapore University of Technology and Design, Singapore

Proceedings Production

Gabriel Wurzer
TU Wien, Austria

Post Graduate Students Consortium:

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

Young CAADRIA Award Committee:

Ding Wen 'Nic' Bao — *Chair*
Anastasia Globa
Christiane M. Herr
Immanuel Koh
Sky Tiantian Lo
Jeroen van Ameijde

International Reviewing Committee

Sheng-Fen Chien
National Cheng Kung University

Thian-Siong Choo
Singapore Polytechnic

I-Ting Chuang
The University of Auckland

Kristof Crolla
University of Hong Kong

Sida Dai
Clemson University

Avishek Das
Aalborg Universitet

Mukta Deshpande
MIT ADT University

Danilo Di Mascio
The University of Huddersfield

Ben Doherty
BVN

Jonathan Dortheimer
Ariel University

Theodoros Dounas
University of Antwerp

Garvin Goepel
The University of Hong Kong

Nik Eftekhar Olivo
Digital Building Technologies (DBT / ITA / ETH)

Halil Erhan
Simon Fraser University

Sarvin Eshaghi
Texas Tech University

Wu Fangning
The Chinese University of Hong Kong

Nora Fankhauser
Gensler / UCL

Shayani Fernando
ETH Zürich, WSL

Darren Fong
University of South Australia

Haruyuki Fujii
Tokyo Institute of Technology

Tomohiro Fukuda
Osaka University

Song Gao
Independent researcher

Fernando García Amen
Universidad de la República

Nicole Gardner
University of New South Wales

Nadja Gaudillière-Jami
TU Darmstadt

John Gero
The University of North Carolina at Charlotte

Ali Ghazvinian
Texas Tech University

Anastasia Globa
The University of Sydney

Ning Gu
University of South Australia

Zhuoxing Gu
Tongji University

Derya Güleç Özer
Istanbul Technical University

Qi Guo
Harbin Institute of Technology

Zhe Guo
Hefei University of Technology

M. Hank Haeusler
University of New South Wales

Gilles Halin
Université de Lorraine

Zhen Han
National University of Singapore

Xinyuan Hao
Nanjing University

Farahbod Heidari
Tarbiat Modares University

Christiane M. Herr
Southern University of Science and Technology

Pablo C Herrera
Universidad Peruana de Ciencias Aplicadas

Toshiki Hirano
The University of Tokyo

Anca Horvath
Aalborg University

Kensuke Hotta
Keio University

June-Hao Hou
National Yang Ming Chiao Tung University

Wei Hu
Tongji University

Chenyu Huang
Tongji University

Chien-hua Huang
China Academy of Art

William Sheng-Yang Huang
University College London

Xiaoran Huang
North China University of Technology

Markus Hudert
Aarhus University

Erin Hunt
Texas Tech University

Magdy Ibrahim Mahmoud
Illinois Institute of Technology

Yasushi Ikeda
The University of Tokyo

Taysheng Jeng
National Cheng Kung University

Guohua Ji
Nanjing University

Muxin Jia
New Jersey Institute of Technology

Adarsh Joy Joseph
CEPT University

Sam Joyce
Singapore University of Technology and Design

Abdallah Kamhawi
Taubman College, University of Michigan

Anna Karagianni
Technical University of Crete

Sumbul Khan
Singapore University of Technology and Design

Chin Koi Khoo
Deakin University

Chavanont Khosakitchalert
Chulalongkorn University

Daiki Kido
Kajima Corporation

Joachim Kieferle
HSRM

Takuya Kikuchi
Kajima Corporation

Nayeon Kim
The Catholic University of Korea

Dongyun Kim
Harvard University Graduate School of Design

Geoff Kimm
Swinburne University

Immanuel Koh
Singapore University of Technology and Design

Satakhun Kosavinta
King Mongkut's Institute of Technology Ladkrabang

Ivana Kuzmanovska
University of New South Wales

Ih-Cheng Lai
Tamkang University

Christoph Langenhan
Technical University of Munich, Karlsruhe Institute of Technology

Maria Larsson
The University of Tokyo

Antonio Leitao
Instituto Superior Tecnico

Surapong Lertsithichai
MQDC Idyllias

Andrew Li
Kyoto Institute of Technology

Bin Li
South China University of Technology, Politecnico di Torino

Chenxiao Li
The University of Hong Kong

Qingpeng Li
Tianjin University

Yunqin Li
Nanchang University

Yuqian Li
Tsinghua University

Fernando Lima
Belmont University

Chieh-Jen Lin
Tainan University of Technology

Yin-shan Lin
University of Hawaii at Manoa

Yuezhong Liu
Nanyang Technology University

Jie Liu
BUPT

Pengkun Liu
Carnegie Mellon University

Yiming Liu
Southern University of Science and Technology

Quan LIU
Nanjing University

Sky Lo
The Hong Kong Polytechnic University

Paul Loh
Bond University

Davide Lombardi
Xi'an Jiaotong-Liverpool University

Werner Lonsing
Independent researcher

Shuai Lu
Tsinghua Shenzhen International Graduate School

Siliang Lu
Bosch

Yao Lu
University of Pennsylvania

Chuanwen Luo
North China University of Technology

Mohammed Makki
University of Technology Sydney

Bob Martens
TU Wien

Iasef Md Rian
Xi'an Jiaotong-Liverpool University

Virginia Melnyk
Virginia Tech

Lin (Leo) Meng
University of New South Wales

Saurabh Mhatre
Harvard Graduate School of Design

Joy Mondal
WEsearch Lab

Zayad Motlib
University of New South Wales

Walaiporn Nakapan
Chulalongkorn University

Taro Narahara
New Jersey Institute of Technology

Worawan Natephra
Mahasarakham University

Peter Ortner
Singapore University of Technology and Design

Mine Özkar
Istanbul Technical University

Stefana Parascho
EPFL

Hyoung-June Park
University of Hawaii at Manoa

Inês Pereira
INESC-ID/IST, University of Lisbon

Panagiota Pouliou
Aalborg University Copenhagen

Deyan Quan
Xi'an Jiaotong-Liverpool University

Ahmad Rafi
Multimedia University

Dinesh Raghavan E S
Vellore Institute of Technology

Balaji Rajasekaran
dMac Group

Cristina Ramos Jaime
University of New South Wales

Rabee Reffat
University of Nottingham Ningbo

Dagmar Reinhardt
The University of Sydney

Jinmo Rhee
University of Calgary

Nicolas Rogeau
Ecole Polytechnique Fédérale de Lausanne

Rohit Priyadarshi Sanatani
Massachusetts Institute of Technology

Marc Aurel Schnabel
Xi'an Jiaotong-Liverpool University

Gerhard Schubert
Technical University of Munich

Kyle Schumann
University of Virginia

Yinyi Shi
South China University of Technology

Daniela Silva
ISCTE-IUL

Yang Song
University of Liverpool

Jingwen Song
The University of Hong Kong

Lawson Spencer
Cornell University

Milena Stavric
Graz University of Technology

R. Spencer Steenblik
Indiana University

Djordje Stojanovic
University of Melbourne

Rudi Stouffs
National University of Singapore

Marcin Strzala
Warsaw University of Technology

Ioanna Symeonidou
University of Thessaly

Ying Yi Tan
Singapore University of Technology and Design

Peng Tang
Southeast University

Michela Turrin
TU Delft

Sepehr Vaez Afshar
Texas Tech University

Jiaqi Wang
South China University of Technology

Likai Wang
Xi'an Jiaotong-Liverpool University

Shih-Yuan Wang
National Yang Ming Chiao Tung University

Sining Wang
Soochow University

Xiao Wang
Southeast University

Glen Wash
Xi'an Jiaotong-Liverpool University

Claudia Westermann
Xi'an Jiaotong-Liverpool University

Albert Wiltsche
Graz University of Technology

Kaicong Wu
The University of Hong Kong

Jeroen van Ameijde
The Chinese University of Hong Kong

Chao Yan
Tongji University

Jiawei Yao
Tongji University

Ho Man Yau
Robert Gordon University

Boyuan Yu
The Chinese University of Hong Kong

Kuai Daniel Yu
University of New South Wales

Rongrong Yu
University of South Australia

Zhang Yue Ying
Southeast University

Darcy Zelenko
Monash University

Shaoting Zeng
Beijing University of Technology

Jiahao Zhang
Osaka University

Jiaxin Zhang
Nanchang University

Wei Zhang
Huazhong Agricultural University

Yuqing Zhang
South China University and Technology

Mingming Zhao
The Chinese University of Hong Kong

Jiangyang Zhao
University of Liverpool

Hao Zheng
City University of Hong Kong

Xinjie Zhou
Tongji University

Guanqi Zhu
The University of Queensland

Zhelun Zhu
Xi'an Jiaotong-Liverpool University

Xinwei Zhuang
University of California

Guest of Honour (Conference Event)
Chong Tow Chong
President, Singapore University of Technology and Design (SUTD)

Professor Chong Tow Chong was appointed President of the Singapore University of Technology and Design (SUTD) in April 2018. He had been SUTD's founding Provost since 2010, where he played an instrumental role in steering the strategic development and operationalisation of SUTD. As President, he provides leadership and guidance in the next phase of the University's growth and development, as well as ensuring continuity in the pursuit of SUTD's vision, mission and strategic goals. Prior to joining SUTD, Prof Chong had been the Executive Director of A*STAR's Science and Engineering Research Council and Data Storage Institute for 15 years. He also had a 30-year academic career with the National University of Singapore as Professor of Electrical and Computer Engineering.

Keynote 1

Akihisa Hirata

Architect, Professor at Kyoto University
Akihisa Hirata Architecture Office

Born in Osaka, Japan in 1971. Graduated from the Department of Architecture in the Faculty of Engineering at Kyoto University in 1994. Received a master's degree from the Kyoto University Graduate School of Engineering in 1997. After working at Toyo Ito & Associates, Architects, he established "akihisa hirata architecture office" in 2005. Currently, He is Professor at Kyoto University. His important works include showroom "Masuya" (2006), apartment "Alp" (2010), "Bloomberg Pavilion" (2011), "Art Museum & Library, OTA" (2017) and "Center of Yatsushiro Folk Performing Arts" (2021). He was awarded 19th Japan Institute of Architects Newcomer Prize (2008), Golden Lion at the Venice Biennale with Toyo Ito, Naoya Hatakeyama and other 2 architects (2012), Colored Concrete Works Award at Berlin (2015) and The Murano Togo Prize (2018) etc. His publications include "Tangling" (LIXIL, 2011), "Discovering New"(TOTO, 2018)etc. He has lectured at Bauhaus Dessau (Germany), Harvard University (U.S.A), University of British Colombia (Canada) and currently lectured at Architecture Foundation (U.K.), etc. He has done personal exhibitions in Tokyo, U.K. and Belgium, etc. And he sent his works to variety of Art fair such as Art Basel, Frieze Art Fair and Milano salone.

"Tangling" and "Resonance"

I have been wishing to create architecture that resonates with the world of life. I would like to see human activities including architecture as an integral part of the world of life, not as something special. It means that keeping away from 'human.' It is at such boundaries that define what it is to be human - at the water's edge of the human - that a new architecture for the age of the environment and life will be appeared.

"Tangling" and "Resonance" are two complementary clues to explore the water's edge of the human. "Tangling" corresponds to proximity, while "Resonance" corresponds to remoteness.

In this lecture, I will introduce some geometric challenges to create the "Tangling" where various things can be intertwined – "Karamarishiro." Additionally, I will talk about architecture that is intertwined with the datistical = 'resonance', like a swarm of thoughts from different people, and creating "Resonances of Resonances," to resonate "Resonance" of the single horizon each other.

Keynote 2

Jinjoon Lee FRSA

Professor of Graduate School of Culture Technology in KAIST (Korea Advanced Institute of Science and Technology)

Founding director of KAIST Art & Technology Center (KATEC)

Distinguished Professor of New Contents Academy (NCA), Korea Creative Content Agency (KOCCA)

FRSA, Life Fellow of Royal Society of Arts, UK

MRSS, Member of Royal Society of Sculptors, UK

Dr. Jinjoon Lee FRSA is a celebrated new media artist, renowned for his groundbreaking research and artistic creations that leverage cutting-edge technologies to examine East-Asia utopian vision and liminoid experiences. His unique artistic journey was initiated in 2007, with his solo exhibition Art Theatre – Role Play held at the Arko Museum of Korea Art Council. Since then, his body of work has graced over 50 global exhibitions and recently he successfully held the solo exhibition Audible Garden (2023) in Korea Cultural Centre UK, London, UK. At present, Dr. Lee shares his wealth of knowledge as an assistant professor at KAIST and as a distinguished professor at New Contents Academy (NCA), Korea Creative Content Agency (KOCCA). After acquiring a B.B.A. in Business School from Seoul National University in 2001, he proceeded to secure a B.F.A and an M.F.A. from same institution. He quests for knowledge led him to Royal College of Art (RCA) for an M.A., and finally, the Ruskin School of Art at the University of Oxford, where He obtained a DPhil. As a Fellow of the historic 270-year-old Royal Society of Arts (FRSA) and a full member of the Royal Society of Sculptors (MRSS), he delves into the study of data-driven art and design, sound art, and the intriguing concept of XR performance for future opera through the use of virtual reality (VR) and artificial intelligence (AI). Lee has invited as a guest artist by the ZKM Center for Art and Media at Germany in 2023 and by Vermont Studio Center (VSC) with full-fellowship at USA in 2024.

Total Experience in Liminal Spaces with VR, AR, and MR Inspired by East Asian Garden Philosophy

This keynote speech delves into the fusion of East Asian garden philosophy with Virtual Reality (VR), Augmented Reality (AR), and Mixed Reality (MR) technologies to craft transformative liminal spaces. The talk will commence by outlining the core tenets of East Asian garden philosophy, spotlighting the themes of natural harmony and the artistic representation of landscapes. It will then progress to discussing the notion of liminoid experiences, which are environments that facilitate transformative encounters beyond the commonplace. The primary focus will be on leveraging VR, AR, and MR technologies to actualize these philosophical concepts, thereby engendering a Total Experience that engrosses users within an amalgamation of virtual and tangible realms. This presentation will underscore the significance of an integrated and synesthetic approach that synergizes technology, design, psychological insights, and cultural legacy. The conclusion will speculate on the future role of these technologies in altering our perception and understanding of space and reality. The objective of this keynote is to inspire a profound appreciation of how enduring aesthetic principles can augment contemporary technological experiences.

Keynote 3

Philip F. Yuan

Professor and Associate Dean, College of Architecture and Urban Planning at Tongji University

Philip F. Yuan is a professor and associate dean of the College of Architecture and Urban Planning at Tongji University, Honorary Fellow of American Institute of Architecture (Hon. FAIA). He is also the co-founder of DigitalFUTURES Association, a global educational initiative with a particular emphasis on the latest computational design and fabrication technologies, Editor-in-Chief of Architectural Intelligence journal, and founding partner of Archi-Union Architects & Fab-Union Technology. Yuan has served as Thomas Jefferson professor at University of Virginia (2019), the visiting professor at Massachusetts Institute of Technology (2019), and Royal Melbourne Institute of Technology (2021). He has also served as council member of UIA Professional Practice Commission (PPC). Yuan was attributed with UIA 2023 The Auguste Perret Prize for Technology in Architecture. His work has been recognized with notable awards, including 2022 AIA Open International | Architecture Honor Award, 2022 Dezeen Award Best Civic Building, 2020 ACADIA Innovative Academic Program Award of Excellence and etc. Yuan has participated in Venice biennale, Chicago biennale, Milan triennial, Tallin biennale, etc. His works have been collected by MOMA New York City, M+ Hong Kong and Centre National d'art et de Culture Georges Pompidou.

Crafting Robotics for Architectural Intelligence

This lecture navigates the intersectional study of AI for architectural design and robotic fabrication, a union encounting the future of architecture in the post-humanistic era. Amidst global challenges, it urges a re-evaluation of the relationship between globalization and locality, advocating for the integration of globalized technologies into localized practices. The production paradigm and knowledge landscape of architecture are being reshaped by crafting robotics for architectural intelligence. As the field of design and practice transits into a new era of flourishing development of intelligent and autonomous technologies, it becomes imperative for architectural academia to redefine the symbiotic relationships between humans and technology. We'll delve into the 'symbiotic' intelligence, aiming to rethinking on the global warming and new paradigm shift under the influence of future technologies. Based on research and practice, the discussion will illustrate how this symbiotic relationship can drive innovative, sustainable solutions, transform the socio-cultural, industrial economics, and environmental resiliency. We are aiming to inspire architects to blur the lines among architects, artifacts, and social processes, contributing to a new understanding of architecture in the post-humanistic era.

Sister Special Roundtable

Accelerated Research

Sherif Abdelmohsen
President | ASCAAD
Professor and Chair, Department of Architecture, American University in Cairo, Egypt

Sherif Abelmohsen is Professor of Digital Media and Design Computing, Chair of the Department of Architecture, Director of the Robotic Fabrication Laboratory, and co-founder of the Educational Virtual Environments Lab at the American University in Cairo, Egypt. He is also President of ASCAAD, and Co-Founder and Managing Partner at UDAAR.

Anastasia Globa
President | CAADRIA
BAE Program Director The University of Sydney

Anastasia is a researcher, academic and designer working in the field of architecture, with strong research interests in algorithmic design, interactive systems, and simulations. Anastasia is based at The University of Sydney. She closely collaborates with the Computer Aided Architectural Design in Asia community, being elected CAADRIA President in 2022.

Christiane M. Herr
Vice-Chair | CAADFutures Foundations
Professor, School of Design
Southern University of Science and Technology, Shenzhen, China
Past President, CAADRIA

Christiane M. Herr is an architectural researcher and educator focusing on cross-disciplinary innovation in engineered ecosystems on buildings as well as digitally supported design. She is currently Professor at the School of Design, Southern University of Science and Technology. Christiane is Past President of CAADRIA as well as board member and vice chair of the CAADFutures Foundation.

Rudi Stouffs
Vice-President (Emeritus) | eCAADe
Dean's Chair Associate Professor

Rudi Stouffs is Dean's Chair Associate Professor in Architecture and Assistant Dean (Research) in the College of Design and Engineering, NUS. He is Vice-

President (Emeritus) of the association for Education and research in Computer Aided Architectural Design in Europe (eCAADe). He has held previous appointments at Carnegie Mellon University, ETH Zurich, and TU Delft.

Paula Gómez Zamora
  Vice-President | SIGraDi
  Senior Research Engineer with the Georgia Tech Research Institute, USA.

Dr. Paula Gómez, a senior research engineer at GTRI, earned her Ph.D. in Computational Design from Georgia Tech. She leads several projects in systems modeling, emphasizing the built environment, energy, sustainability, and human well-being. Recognition of her work includes international awards, service on scientific committees, and IJAC editorials. She currently serves as Vice-President of SIGraDi.

CAADRIA Fringe City Events

AI-NITE Guest of Honour

Chee Su Eing

President, Design Business Chamber Singapore (DBCS)

Chee Su Eing, a pioneering figure in Singapore's design landscape, notably serves as the first female President of Design Business Chamber Singapore (DBCS) in its 36-year history. As Founder and Director of D'Perception Pte Ltd and Managing Director of D'Perception Ritz Pte Ltd, she spearheads full-suite interior design ventures across Singapore and Asia, acclaimed for their excellence in residential, social, and hospitality interior architecture. With over two decades of expertise, she champions innovative design solutions and has served as a juror for various prestigious design awards. Su Eing is deeply committed to leveraging design for societal and business impact, exemplified through her leadership at DBCS. Under her guidance, DBCS actively fosters partnerships and collaborations with major design players, positioning itself at the forefront of discussions on cutting-edge topics such as AI and sustainability. Through these initiatives, Su Eing and DBCS drive forward the dialogue on how design can shape a more sustainable and technologically advanced future.

AI-NITE Keynote

Immanuel Koh

Assistant Professor, Singapore University of Technology and Design (SUTD)
Director, Artificial-Architecture

Immanuel Koh is the founder and director of Artificial-Architecture. He is an Assistant Professor in Architecture & Sustainable Design (ASD) and Design & Artificial Intelligence (DAI) at the Singapore University of Technology & Design (SUTD) where he leads research and teaching at the intersection of AI and Architecture. Trained at the Architectural Association (AA) in London and holds a PhD from the School of Computer Sciences and Institute of Architecture at the École polytechnique fédérale de Lausanne (EPFL), he is a pioneer in AI x Architecture who operates in an original and trans-disciplinary way. He is the Principal Investigator for several funded AI research projects with support from, among others, the National Research Foundation (NRF), Ministry of Defence (MINDEF), National Arts Council (NAC), AI Singapore (AISG), DesignSingapore Council (DSG), Urban Redevelopment Authority (URA), DSO National Laboratories, Temasek Laboratories, Hokkien Foundation, Sunray Woodcraft Construction, and National Supercomputing Centre (NSCC). Immanuel is an expert consultant at the ZJU-BOSIDENG Joint Research Centre on AI Design and currently conducts research for high-profile architecture practices such as Zaha Hadid Architects (London) and MVRDV (Rotterdam) in developing custom state-of-the-art deep learning models. He has also published and exhibited widely, ranging from AAAI, ICCV, CVPR, NeurIPS and AD to Singapore’s Arts House, Venice Architecture Biennale, and V&A Museum; and taught internationally such as at UCL, RCA, Bauhaus Dessau, Harvard GSD, and Strelka. Immanuel co-founded Neural Architecture Group, directs DesignerlyAI, and is the author of the book ‘Artificial & Architectural Intelligence in Design’ published in 2020.

AI-NITE Roundtable

Accelerated Practice

Marios Tsiliakos
Partner | Applied R+D, Foster+Partners
Lecturer | Bartlett School of Architecture, UCL

Marios is a Design Systems Analyst and Partner at the Applied R+D team at Foster+Partners, expanding the boundaries of applied computation and problem solving for AEC industry. Specializing in complex geometry, performance driven design, and interoperability, he leads the development of F+P’s in-house interoperability tool.

Hui Min Chan
Director and the Head of Green Well Tech at DP Architects

Green-Well-Tech (GWT) is DP Architects' strategic initiative to drive growth and excellence throughout its business thrusts and unify its actions and capabilities toward innovative and sustainable solutions. As part of her role, she also oversees DPA's Smart Sustainability Unit, which extends the core business into tech-related business lines.

Vignesh Kaushik
Principal & Regional Design Technology Director at Gensler

Vignesh leverages his passion for people empowerment, process development and technological innovation to lead the digital transformation of the Gensler practice, which covers multiple typologies, complexities and scales across different sectors in the APME and GC regions. He has taught advanced computational design in many institutions in Southeast Asia.

Eva Castro
Professor of Practice at SUTD ASD
Co-founder & Director GroundLab / Plasma studio / formAxioms

Eva Castro currently is the coordinator of core studio 2 and co-leads the advanced option studio on speculative futures and oceanic inhabitations at SUTD. Castro is co-founder of form_axioms lab, a territorial agency for academic research purposes operating from within Singapore. As a practitioner, Castro is cofounder of Plasma Studio and GroundLab.

Pia Fricker
Professor of Practice in Computational Methodologies in Landscape Architecture and Urbanism, Vice Head of Department, Aalto University Finland, School of Arts, Design and Architecture.

Prof. Dr. Pia Fricker, is a leading expert in computational methodologies for landscape architecture and urbanism. Her pioneering work utilizes diverse AI-driven computational methods to tackle pressing challenges in urban design amidst the accelerating climate crisis and technological advancements. With a focus extending beyond the human-centered paradigm, Fricker's internationally showcased research redefines trajectories for architecture in the era of artificial intelligence and environmental transformation.

AI-NITE Exhibition

Neural MONOBLOC Black: Artificial Intelligence and its Aesthetic Subversion

By Immanuel Koh | Artificial-Architecture

The white stackable plastic Monobloc chair is the world's most widely/cheaply/quickly produced and disposed chair. Most people would have at some point in their lives sat on a Monobloc chair without even noticing it. The Monobloc chair is also the most common chair imagery on the internet, thus automatically finding its way into any datasets used to train today's most powerful foundation AI models such as ChatGPT and Stable Diffusion. The Monobloc is a blanket term and therefore it is one and many chairs at the same time. It doesn't really have a name(s) nor an acknowledged designer(s). However, its physical and digital ubiquity makes it the perfect baseline artefact to discuss design with anyone, from anywhere, and at any time. In fact, the Monobloc is closer to a design concept than a designed object.

The exhibition "Neural MONOBLOC Black" presents 3 acts of aesthetic subversion through the Monobloc – from the original 'Everyday Normal Monobloc' to the transgressive 'Artist Bricolage Monobloc', and finally to the

AI-hallucinated ‘Neural Monobloc Black’. The ‘Everyday Normal Monobloc’, with its endless machinic production but homogenous design variations, is generally hated and despised by designers, but has continued to exert its usefulness, messiness, and even place-making-ness in public spaces everywhere. This is Act 1. The ‘Artist Bricolage Monobloc’, with its D-I-Y/collage-based formal operations and layered narratives, is a form of design critique by hacking everyday normal monoblocs and showing them as one-off art pieces in esteemed museums. This is Act 2.

The ‘Neural Monobloc Black’ is the Final Act. It is the focus of this exhibition which features eight physical artefacts that are generated and fabricated directly in 3D through a custom fine-tuned text-to-3D AI model developed at Artificial-Architecture. The ‘Neural Monobloc Black’ is the hallucinated result of an AI model’s attempt in deconstructing and reconstructing a hypothetical 3D monobloc from its flatland training set of 2D monobloc imagery. Its appearance of uncanniness suggests an underlying Freudian ‘repetition’ and ‘doubling’. Its appearance of wrongness might be explained by a quote from Mark Fisher -- “The weird thing is not wrong, after all: it is our conceptions that must be inadequate.” The blackbox of artificial intelligence and the charred-black chairs are instrumentalised here to question our all-too-human conception and perception of what design is and can be.

Artificial-Architecture (A-A) is based at the Singapore University of Technology & Design (SUTD) across the academic programmes of Architecture & Sustainable Design (ASD) and Design & Artificial Intelligence (DAI). It is a speculative and experimental design research group that not only explores the technics of the artificial, but also its epistemological, cultural, political, and aesthetic implications through the expanded lens of the architectural, and vice versa. Our team includes architects, designers, urban planners, engineers, programmers, mathematicians, theorists, and computer scientists.

Exhibition Project Team:
Ashley Chen, Benedict Tan, Christopher Ooi, Elissa Hartanto, Lynus Lim, and Tan Zhi Sheng.

TABLE OF CONTENTS

Artificial Intelligence & Machine Learning in Design **7**

A Method for Real-Time Prediction of Indoor Natural Ventilation in Residential Buildings 9
Zewei Shi, Xiaoxin Wang, Jinyu Wang, Yu Wang, Yixin Jian, Chenyu Huang, Jiawei Yao

A Pattern Coloring Method Using Rank-Based Interactive Evolutionary Algorithm 19
Elif Gamze Dedeler, Michael S. Bittermann

A Rapid Prediction Model for View-Based Glare Performance With Multimodal Generative Adversarial Networks 29
Xiaoqian Li, Zhen Han, Gang Liu, Rudi Stouffs

Accelerating Future Scenario Development For Concept Design With Text-Based GenAI (ChatGPT) 39
Linus Tan, Thom Luke

Adapting the Software Design Pattern Model for AI-Enabled Design Computing 49
Geoff Kimm, Marcus White, Mark Burry

AI-Enhanced Performative Building Design Optimization and Exploration: A Design Framework Combining Computational Design Optimization and Generative AI 59
Chuwen Zhong, Yi'an Shi, Lok Hang Cheung, Likai Wang

ARCHICLIP: Enhanced Contrastive Language–Image Pre-training Model With Architectural Prior Knowledge 69
Shengtao Xia, Yiming Cheng, Runjia Tian

Architectural Generative Model Evaluation Methods: Image Quality Assessment Metrics and Expert-Based Approach 79
Ying Lin, Fei Ye

Can Generative AI Models Count? Finetuning Stable Diffusion for Architecture Image Generation with Designated Floor Numbers Using a Small Dataset 89
Weishun Xu, Mingming Li, Xuyou Yang

ChatDesign: Bootstrapping Generative Floor Plan Design With Pre-trained Large Language Models 99
Jinmin Li, Yilu Luo, Shuai Lu, Jingyun Zhang, Jun Wang, Rizen Guo, ShaoMing Wang

Deep Spatial Memory: Quantifying Architectural Spatial Experiences through Agent-driven Simulations and Deep Learning 109
Shuhan Miao, Wenzhe Peng, Daniel Tsai, Takehiko Nagakura

DSNL in Architecture – A Deep Learning Approach to Deciphering Architectural Sketches and Facilitating Human-AI Interaction 119
Wei Hu

From Topology to Spatial Information: A Computational Approach for Generating Residential Floorplans 129
Mohamed El Mesawy, Nawal Zaher, Ahmed El Antably

GRAPH2PIX: A Generative Model for Converting Room Adjacency Relationships into Layout Images 139
Zhen Han, Xiaoqian Li, Ye Yuan, Rudi Stouffs

Latent Petroglyphs: Pattern Extraction From Prehistoric Rock Art Through Generative Workflows for a Design Project in Greece 149
Smaro Katsangelou, Parker Wilmoth, Dimitris A. Pados, Emmanouil Vermisso

Learning and Generating Spatial Concepts of Modernist Architecture via Graph Machine Learning 159
Erik Bauscher, Anni Dai, Diellza Elshani, Thomas Wortmann

Drag2build: Interactive Point-Based Manipulation of 3D Architectural Point Clouds Generated From a Single Image 169
Jun Yin, Pengjian Xu, Wen Gao, Pengyu Zeng, Shuai Lu

Pseudo-cross-modal Translation: Bridging Architectural Plan and Perspective through a pix2pix Network 179
Yujun Mao, Wenzhe Peng, Takehiko Nagakura

The Intersection of Technology and Architecture: Smartphone Photography in Carbon Analysis 189
Iuan-Kai Fang, Shen-Guan Shih

Building Information Modelling 199

BatchPlan: A Large Scale Solution for Floor Plan Extraction 201
Burak Yildiz, Javier Cuartero, Fatemeh Mostafavi, Seyran Khademi

BIM-Enabled Regulatory Design Rule Checking for Building Circulation 211
Cheryl Lee, Khystelle Yeo, Stylianos Dritsas

Bridging Bim and AI: A Graph-Bim Encoding Approach for Detailed 3D Layout Generation Using Variational Graph Autoencoder 221
Jiadong Liang, Ximing Zhong, Immanuel Koh

Extracting Spatial Structure From Building Information Models 231
Stylianos Dritsas, Khystelle Yeo, Cheryl Lee

More Words from Facility Operation Managers: Text-based Building Information for Inclusive Accessibility 241
Ai Kasai, Tsukasa Ishizawa

Predicting Pedestrian Trajectories in Architectural Spaces: A Graph Neural Network Approach 251
Runyu Yang, Weili Wang, Peng Gui

Review on the Use of Conversational AI NPC Avatars in Teaching and Learning BIM: A Preliminary Observation of Its Introduction in a Built Environment Related Course in Singapore 261
Juniarto Hadiatmadja

UrbanSight: Application of Digital Twin to Augmented Reality 271
Tzu-Hsin Hsieh, Tay-Sheng Jeng

Generative, Algorithmic, & Evolutionary Design 281

A Hybrid Multi-objective Model for Multi-story Warehouse Design: A Case Study in Shenzhen 283
Yi Wen, Wei Ye, Gang Yu

Architectural Topological Form-finding Integrating Solid and Fluid Structural Performances 293
Xin Yan, Kun Di Shu, Ding Wen Bao

Clicking is All You Need: Implementing Wave Function Collapse in Early-Stage Design for Manufacturing and Assembly Projects 303
Owen Zhiyuan Lu, Leo Lin Meng, Cristina Ramos Jaime, M. Hank Haeusler

Exploring Rule-Based Discrete Timber Design With Mortise-Tenon Joints 313
Qianbin Xu, Frederico Fialho Teixeira, Maryam Shafiei

Extracting Actionable Information from the Site Context Using A Phenotype-Based Strategy 323
Ban Liang Ling, Bige Tunçer

Formal Variation Exploration for Performance-Based Building Massing Design Optimization 333
Tong Shao, Likai Wang, Guohua Ji

From Sketch to Design: A Cross-scale Workflow for Procedural Generative Urban Design 343
Jin Gao, Sayjel Vijay Patel

Re-imagining The Urban Development of Western Sydney: The Case Study of Oran Park 353
Monika Krncevic, Niousha Arjaghi, Mohammed Makki, Mathers Jordan

Rule-Based Generation of Interwoven Assemblies in Architectural Design: A Computational Approach Integrating Attribute Grammar 363
Kevin Harsono, Shen-Guan Shih, Ye Yint Aung, Felicia Wagiri, Tsung-Wei Cheng

SIMForms: A Web-Based Generative Application Fusing Forms, Metrics, and Visuals for Early-Stage Design 373
Baizhou Zhang, Yichen Mo, Biao Li, Yanyu Wang, Chao Zhang, Ji Shi

Turning Empirical Knowledge into Digital Assets: A Workflow Assisting Intern Architects in High-Rise Conceptual Design 383
Chengyu Sun, Yilin Li, Yueli Wang

Performance-based Design & Analytics **393**

A Knowledge Graph Model for Performance-Based Generative Design and Its Applications in Accelerated Design 395
Zhaoji Wu, Zhe Wang, Jack C.P. Cheng, Helen H.L. Kwok

A Parametric Method in Room Acoustics Simulation With Performance Verification: Real-Time Ray Tracing Techniques in Parametric Modeling 405
Bo-Young Lin, Wei Lin

A Phygital Form-Finding Method for Body-Scale Soundscape Installation With Flexible Materials 415
Sijia Gu, Chao Yan, Tianyi Gao, Yufei Wu, Ruishu Wang, Guoliang Lian, Rixin Wang, Yufei Peng, Fangshuo Mo, Philip F. Yuan

An Approach to Improve Individual Thermal Comfort Based on Mobile Measurement of Biological Reaction to Proximate Thermal Environment 425
Shun Kimura, Yasushi Ikeda

Bridging Exterior and Interior Climate: Interdisciplinary Design of a Dual-functional Adaptive Kinetic Façade Prototype 435
Moon-Young Jeong, Reiner Rössle, Michael Böhm, Maria Matheou, Oliver Sawodny

Coastal Infrastructure Design: Researching Sea-Waves and Textured Surfaces Interaction Using Physical and Virtual Wave Flumes 445
Roy Kozlovsky, Yasha Grobman, Hanna Levy

D-Predict: Integrating Generative Design and Surrogate Modelling with Design Analytics 455
Esmaeil Mottaghi, Ahmed M. Abuzuraiq, Halil Erhan

Design Factor-Oriented Life-Cycle Energy Optimisation for Early-Stage Building Design 465
Yang Yang, Likai Wang, Marco Cimillo

Design Optimization Through CFD-ABMS Integration for Controlling Virus Spread 475
Chao-Nien Chen, Pei-Hsien Hsu

Explainable AI and Multi-objective Optimization For Energy Retrofits in Existing Residential Neighborhoods 485
Jiahong Ye, Yanting Shen, Chenyu Huang, Jinyu Wang, Rong Qu, Jiawei Yao

Glare Prediction in Check-in Halls of Airport Terminals Using Integrated Algorithms and Transfer Learning Strategy: A Case Study of Guangzhou Baiyun International Airport 495
Yinyi Shi, Zhichao Ma, Cunyuan Wang, Xiong Chen

Integrating Genetic Algorithms and RBF Neural Networks in the Early Design Stage of Gymnasium for Multi-Objective Optimization Framework 505
Zhaoxiang Fan, Shuoning Tang, Mengxuan Liu

Making a Case for Design Analytics: Complementing Designers' Toolbox for Data-Informed Creative Decision-Making 515
Halil Erhan, Ahmed M. Abuzuraiq

Mixed-Use Eco-Computational Design in Hot Climates: A Two-Step Cross-Climate Parametric Workflow 525
Eilam Sklar, Shachar Lorberbom, Jonathan Natanian

Multispecies Building Envelopes: Adopting Plant Habitat Suitability Modelling for Ecological Design Decision-Making 535
Surayyn Uthaya Selvan, Laura Windorfer, Mariasole Calbi, Soultana T. Saroglou, Enrica Roccotiello, Shany Barath, Yasha Grobman

Uncovering the Circular Potential: Estimating Material Flows for Building Systems Components Reuse in the Swiss Built Environment 545
Shuyan Xiong, Edwin Zea Escamilla, Guillaume Habert

Artificial Intelligence & Machine Learning in Design

A METHOD FOR REAL-TIME PREDICTION OF INDOOR NATURAL VENTILATION IN RESIDENTIAL BUILDINGS

ZEWEI SHI[1], XIAOXIN WANG[2], JINYU WANG[3], YU WANG[4], YIXIN JIAN[5], CHENYU HUANG[6], JIAWEI YAO[7]
[1,3,5,6,7]*College of Architecture and Urban Planning, Tongji University.*
[2,4]*Lingsi Technology (Shanghai) Co., Ltd.*
[1]*szw_shi2009@163.com, 0009-0007-3644-6513*
[2]*1428433871@qq.com*
[3]*johnwang1998@163.com*
[4]*wyu0211@foxmail.com*
[5]*zjwzjianyixin@126.com*
[6]*cyhuang@tongji.edu.cn, 0000-0002-6360-638X*
[7]*jiawei.yao@tongji.edu.cn, 0000-0001-7321-3128*

Abstract. Against the backdrop of energy crises and climate change, performance-oriented architectural design is increasingly gaining attention. Early-stage assessment of natural ventilation performance is crucial for optimizing designs to enhance indoor environmental comfort and reduce building energy consumption. However, traditional numerical simulations are time-consuming, and existing data-driven surrogate models primarily focus on predicting partial indicators in indoor airflow or single-space airflow. Predicting the spatial distribution of airflow is more advantageous for addressing global issues in building layout design. This paper introduces a surrogate model based on Generative Adversarial Networks. We constructed a dataset of floor plans, with 80% of the data generated using parameterized methods and 20% sourced from real-world examples. We developed a 3D encoding method for the floor plans to facilitate machine understanding of spatial depth and structure. Finally, we conducted airflow simulations on the dataset, with the simulated results used to train the Pix2pix model. The results demonstrate that the Pix2pix model can predict indoor airflow distribution with high accuracy, requiring only 0.8 seconds. In the test set, the average values for MAPE, SSIM, and R^2 are 2.6113%, 0.9798, and 0.9114, respectively. Our research can improve architectural design, enhance energy efficiency, and enhance residents' comfort, thereby contributing to the creation of healthier indoor environments.

Keywords. Generative Residential Buildings, Natural Indoor Ventilation, Early Design Stage, Real-time Prediction, Generative Adversarial Networks (GAN)

1. Introduction

Due to the increasingly severe pollution and climate change, the challenges faced by urban and indoor environments in providing good air quality and thermal comfort are becoming more pressing. As people spend a significant portion of their lives indoors, the indoor air circulation environment directly affects daily life, particularly in terms of health, comfort, and productivity. Therefore, creating a healthy and comfortable indoor environment, especially in residential settings, is of paramount importance.

To enhance the thermal comfort of residents, a substantial amount of energy is used for the air conditioning systems of residential buildings. However, adopting passive design methods such as natural ventilation can regulate indoor air quality and comfort, while also reducing the energy consumption of air conditioning equipment (Yik and Lun, 2010), thereby promoting sustainable development. During the early design stages, buildings have greater potential for performance optimization, emphasizing the need to predict and assess indoor airflow during these phases (Wen and Hiyama, 2018).

2. Related Work

Scholars both domestically and abroad have conducted extensive research on simulating wind environments. Over the years, Computational Fluid Dynamics (CFD) has been widely and consistently applied to simulate airflow in building environments. However, the parameter inputs of traditional CFD simulation software are complex, and simulations require a considerable amount of time, resulting in low simulation efficiency. This significantly limits the development of performance-based design in architecture.

With the continuous progress of machine learning, data-driven surrogate models have been able to significantly improve prediction speed while maintaining predictive accuracy (Nguyen Van and De Troyer, 2018) (Meddage et al., 2022). As the volume of data increases, artificial neural networks (ANN) have become a research focus due to their high prediction accuracy and low computational costs. Studies have utilized neural network models as surrogate models to predict low-dimensional distributions of pollutant concentrations (Cao and Ren, 2018), forecast indoor airflow patterns and temperature distributions (Zhou et al., 2021). Furthermore, research has employed ANN for rapid energy consumption prediction in early-stage complex architectural form design (Li et al., 2019), optimizing residential building HVAC systems through ANN-based model predictive control (Afram et al., 2017), and improving indoor airflow prediction accuracy by sequentially connecting two independent ANN models, avoiding the need for large training datasets (Kim et al., 2023).

However, current predictions of indoor airflow data mostly focus on key coefficients such as average wind speed, wind pressure coefficients, and air age. The use of indoor space is characterized by non-uniformity, meaning that people tend to stay in specific areas with higher requirements for environmental quality. Therefore, predicting detailed global airflow within the entire space is crucial for addressing local issues in indoor design, such as spatial distribution, optimization of door and window shapes, sizes, and layouts, ensuring comfort and environmental quality indoors.

The emergence of Generative Adversarial Networks (GANs) has propelled the rapid development of image prediction and has been continuously applied in the field of predicting architectural environmental performance. Some studies have used pix2pix to learn wind distributions around buildings (Mokhtar et al., 2020) and predict wind pressure images (Hu et al., 2020). Additionally, pix2pix models have been used to real-time predict outdoor wind environments, comfort, and solar radiation, accelerating performance-driven urban design (Huang et al., 2022). Furthermore, combining pix2pix-based outdoor wind field and solar radiation prediction models with urban design and optimization systems has been explored (Duering et al., 2020). Apart from urban environments, GAN models have also been applied to indoor environmental predictions, such as using CNN and GAN as surrogate models for planar sunlight simulation, predicting static, annual sunlight metrics, and spatial illuminance distributions (He et al., 2021). Additionally, for indoor airflow environment predictions, a new boundary condition CGAN model has been created, generating two-dimensional airflow distribution images based on continuous input parameters (Faulkner et al., 2023).

Based on the above research, GANs have high application value and potential in predicting indoor airflow fields. However, current research on the prediction of global indoor airflow mostly relies on fixed boundary spaces or single spaces, leaving certain gaps in predicting the global wind environment in diverse indoor spaces.

3. Methods

This paper proposes a rapid prediction method for indoor natural ventilation based on a CGAN model. Initially, the model constructs samples based on parameterized generative models and real residential models. Subsequently, these samples are imported into the Butterfly plugin for batch simulation of indoor airflow environments. Next, a dimensionality reduction and encoding process is applied to the three-dimensional models. The model encoding results, combined with simulation results, constitute the training dataset for the Pix2pix model. Finally, the Pix2pix model is utilized to construct a predictive model for indoor environmental performance. The technical path of the study is illustrated in Figure 1.

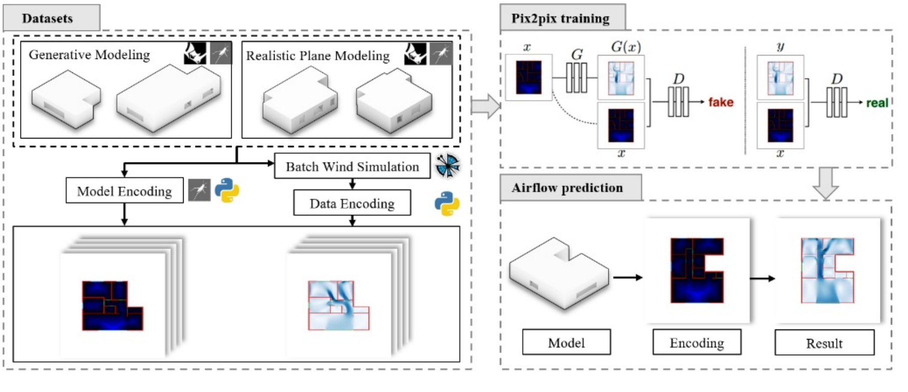


Figure 1. Workflow

3.1. DATA COLLECTION

This study established two distinct datasets, with 80% derived from generative parametric modelling and 20% based on actual residential floor plans. These datasets were employed for model training, and the results were subsequently compared.

3.1.1. Parametrically Generated Residences

The generative residential model is based on Grasshopper shape grammar, utilizing the residential outline boundaries, the edges of exterior walls, and the entrance door position as inputs to generate residential floor plans parametrically (PlanFinder, 2023). Secondary morphological control indicators, including floor height, door width and height, window sill height, and window width and height, were established based on relevant research to generate three-dimensional models. The logic of model generation is illustrated in Figure 2. To comply with residential architectural design standards and energy efficiency specifications, constraints were added, such as the window-to-wall ratio (WWR). The parameter values are provided in Table 1. The collected generative residential dataset comprises a total of 319 samples.

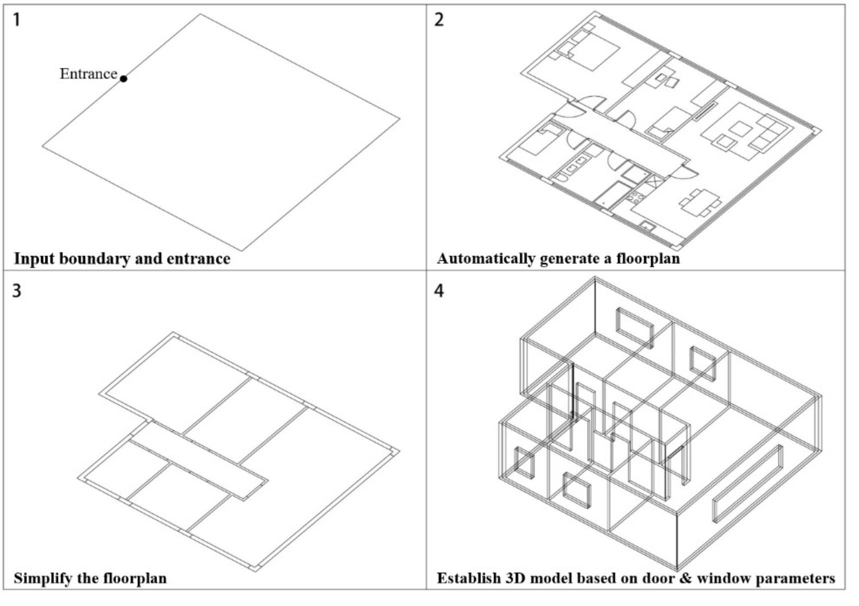


Figure 2. Parametrically generated residences

3.1.2. Real Residential Dataset

In order to increase the rationality of the floorplan of the training samples, a real data set is added, which is derived from real residential floor plans. Three-dimensional models were constructed using Grasshopper, with consistent values set for window sill height (0.9m) and window height (1.5m). To expand the dataset and mitigate overfitting, cases were transformed into new floor plans for data augmentation. By changing the window width to enhance real residential dataset, the window width was reduced to 80% and 70% of the original, respectively. The actual dataset comprised a total of 84 samples.

Table 1： Values for Residential Floor Height, Door Width and Height, Window Sill Height, Window Width, and Window Height Parameters

Parameters	Values (Range)
Floor Height	3.0m
Door Width	0.9m
Door Height	2.0m
Window Sill Height	0.8m~1.2m
Window Width	0.5m~2m
Window Height	0.9m~1.8m
WWR	Northward ⩽ 0.25; East-west ⩽ 0.30; Southward ⩽ 0.35.

3.2. IMAGE ENCODING

The input for pix2pix is defined as a three-dimensional tensor of size 256*256*3, thus requiring the encoding of geometric data from the residential models into a three-dimensional tensor of the same size, as illustrated in Figure 3(a).

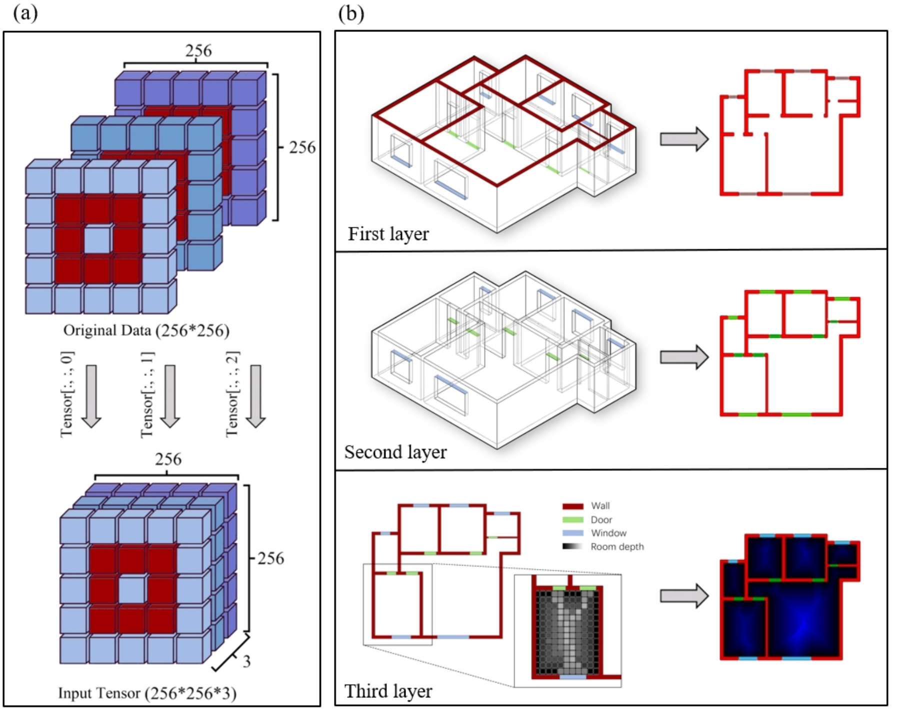


Figure 3. Model encoding method

Considering the significant impact of room walls and openings on indoor wind speed, the distance from the indoor space to the walls is considered in the image encoding. To maximize the geometric information contained in the encoding, the following encoding design was carried out: stacking planar geometric information, height information, and the distance from the space to the walls in the depth direction. The information covered by the first layer includes planar geometry information, height from the ground to the lower edge of door openings, height from the ground to the lower edge of window openings, and floor height. The second layer contains information on the height from the ground to the upper edge of door openings and window openings. The third layer provides information on the distance from the interior space to solid walls. A monitoring surface of 20m*20m is set at height of 1.5m above the ground, divided into 256*256 evenly distributed monitoring points. The Manhattan distance from the monitoring points inside the room to the nearest wall is calculated. As shown in Figure 3(b), this encoding method clearly displays the planar geometry and height information of the residential walls and doors/windows, indicating the impact of room shape, door/window opening positions, and sizes on the indoor airflow field.

3.3. INDOOR AIRFLOW SIMULATION

Butterfly is an interactive interface based on the OpenFOAM CFD Code, an open-source CFD library written in C++ and run in a Linux environment. It is used to visually invoke OpenFOAM commands. The parameter settings during the CFD simulation process in this study, including the configuration of the computational domain, boundary conditions, and mesh size, follow existing guidelines and best practices (Tominaga et al.,2008). The boundaries of the computational domain are set in the inflow and outflow directions, with the sides and top positioned 1.2H away from the wind tunnel boundaries, where H represents the maximum building height. The south-facing side is designated as the inflow direction, while the north-facing side is set as the outflow direction. Inlet boundary conditions are defined as velocity inlet boundaries, with a wind speed of 1.8 m/s. All surfaces are defined as no-slip walls. Subsequently, a hexahedral unstructured mesh is generated with grid counts of 200, 200, and 50 in the x, y, and z directions, respectively. A monitoring surface is set at a height of 1.5m above the ground, with 65,536 uniformly distributed monitoring points on a 20m*20m monitoring plane.

The simulation is based on the three-dimensional steady Reynolds-averaged Navier-Stokes (RANS) equations, utilizing the standard k-ε turbulence model. The stopping criterion is set as residuals less than 10^{-3}. Finally, the accuracy of the GAN airflow prediction model is validated by comparison with a wind tunnel database. Sensitivity tests, including the evaluation of model parameters such as color schemes, are conducted to assess potential influences on the model.

3.4. PIX2PIX TRAINING

In this study, a surrogate model based on pix2pix is established to predict the indoor airflow. The Pix2pix algorithm consists of two components: a generator and a discriminator. The generator adopts a U-net architecture, comprising an encoder and a decoder. Each block in the encoder is composed of convolution-batch normalization-Leaky ReLU, while each block in the decoder is comprised of transpose convolution-batch normalization-dropout (applied to the first three blocks)-ReLU. Skip connections exist between the encoder and decoder. The discriminator employs PatchGAN to determine the "authenticity" of independent patches in an image. Each block in the discriminator is convolution-batch normalization-Leaky ReLU. The losses in Pix2pix include CGAN loss and L1 loss. L1 loss represents the mean absolute error between the generated image and the target image.

The pix2pix network was trained using the TensorFlow deep learning framework, with architectural geometry-encoded plans as input and predicted wind field images as output. The dataset was partitioned, allocating 70% for training, 15% for validation, and 15% for testing. Considering the trade-off between image accuracy and computational cost, the iteration count was ultimately set at 200. The training process involved a total of 282 instances from the training set. Model training parameters adhered to the default settings of the Pix2pix algorithm. The experimental setup utilized an AMD R7 5800H, a 64-bit Windows 10 operating system, and an NVIDIA RTX 3070 Ti.

4. Result

The generators and discriminators of the Pix2pix model converge after 200 epochs. To quantitatively assess the accuracy of the model's predictions, MAPE, SSIM, and R^2 are selected as evaluation metrics. MAPE, the Mean Absolute Percentage Error, measures the average percentage error between predicted and actual values. A lower MAPE indicates better predictive performance. SSIM, the Structural Similarity Index, is a widely used objective image quality assessment metric based on the assumption of the highly adaptive structure information of the human visual system. SSIM values are generally not greater than 1, with a value of 1 indicating identical images. R^2 is an indicator measuring the degree of fit of the model to the observed data, representing the ratio of the variance explained by the model to the total variance.

The test set is input into the trained model, and MAPE, SSIM, and R^2 are calculated. Figure 5 displays real images, predicted images, error images, and corresponding evaluation metrics for a subset of the test set. Visually, the predicted indoor airflow distribution in the model's output aligns well with the actual images. The average values of MAPE, SSIM, and R^2 for all test results are 2.6113%, 0.9798, and 0.9114, respectively. The experimental results indicate a high degree of agreement between the model's predictions and the calculations from simulation software. Therefore, this model can provide intuitive feedback on indoor airflow distribution predictions for design professionals in the early stages of residential building design, assisting designers in performance-based design while ensuring accuracy.

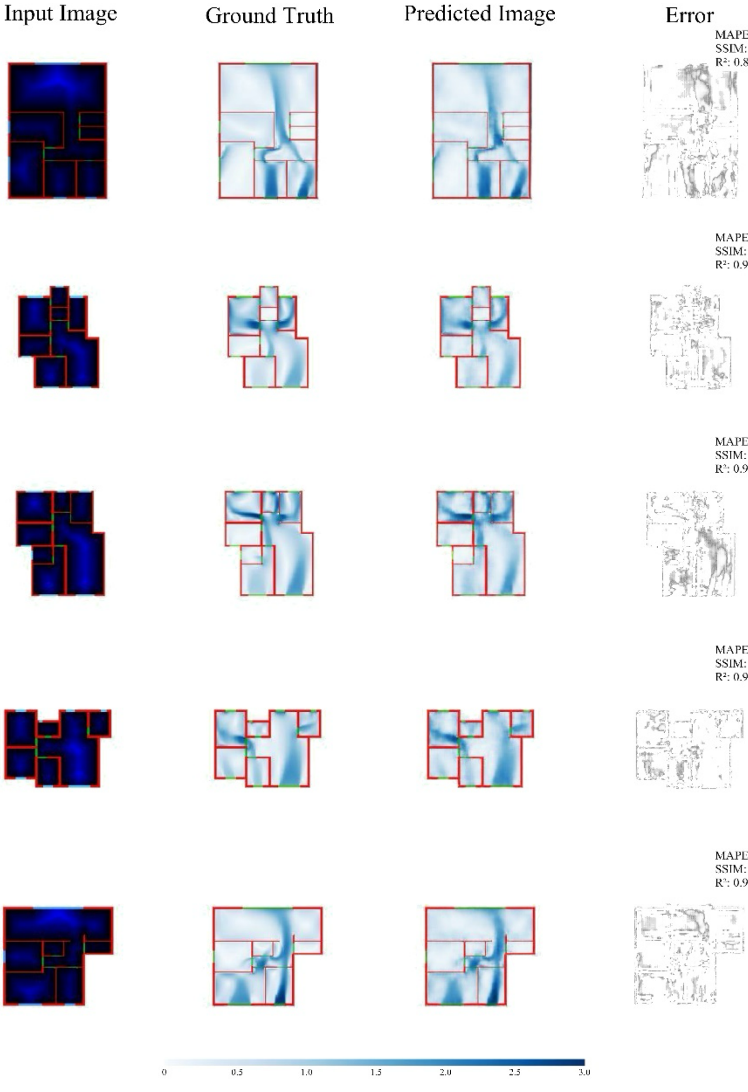


Figure 5. partial results

5. Discussion and conclusion

Against the backdrop of climate change, the assessment of indoor environmental performance becomes increasingly crucial. This study focuses on the design of naturally ventilated residences, aiming to simplify and expedite the indoor airflow simulation process. The goal is to enhance the efficiency of early-stage energy-efficient and comfortable residential design, contributing to building energy savings, emissions reduction, and sustainable development.

This research leverages GAN to achieve rapid prediction of indoor airflow under natural ventilation conditions. The results indicate an average prediction time of 0.8 seconds per sample, compared to 1200 seconds using traditional simulation software for the same samples. The model significantly reduces the time required for predictions compared to numerical simulations. The average MAPE, SSIM, and R^2 for the test set are 2.6113%, 0.9798, and 0.9114, respectively. This predictive model meets the needs of designers to instantly visualize the natural ventilation performance of designs in the early stages of residential design. This model can provide rapid indoor airflow predictions for typical residential building forms without complex input parameters.

However, the training samples in this study are limited, and it is possible to achieve predictions of indoor airflow in more diverse residential floor plans by increasing the volume of real residential dataset. Currently, the research is limited to residential building types, and the scope can be further expanded to include other types of buildings. In addition, this study uses fixed wind directions and speeds for training. In the future, more data can be simulated with multiple wind directions, and training can be conducted using wind speed ratios to enhance the generalizability to different regions and climates.

Acknowledgements

This research was financially supported by the National Natural Science Foundation of China under Grant NO. 52278041 and the Fundamental Research Funds for the Central Universities.

References

Yik, F. W. H., & Yu Fat Lun. (2010). Energy saving by utilizing natural ventilation in public housing in Hong Kong. *Indoor and Built Environment*, 19(1), 73–87. https://doi.org/10.1177/1420326x09358021

Wen, L., & Hiyama, K. (2018). Target air change rate and natural ventilation potential maps for assisting with natural ventilation design during early design stage in China. *Sustainability*, 10(5), 1448. https://doi.org/10.3390/su10051448

Wen, L., & Hiyama, K. (2018). Target air change rate and natural ventilation potential maps for assisting with natural ventilation design during early design stage in China. *Sustainability*, 10(5), 1448. https://doi.org/10.3390/su10051448

Meddage, D. P. P., Ekanayake, I. U., Weerasuriya, A. U., Lewangamage, C. S., Tse, K. T., Miyanawala, T. P., & Ramanayaka, C. D. E. (2022). Explainable machine learning (XML) to predict external wind pressure of a low-rise building in urban-like settings. *Journal of Wind Engineering and Industrial Aerodynamics*, 226, 105027. https://doi.org/10.1016/j.jweia.2022.105027

Cao, S.-J., & Ren, C. (2018). Ventilation control strategy using low-dimensional linear ventilation models and artificial neural network. *Building and Environment*, 144, 316–333. https://doi.org/10.1016/j.buildenv.2018.08.032

Zhou, Q., & Ooka, R. (2021). Influence of data preprocessing on neural network performance for reproducing CFD simulations of non-isothermal indoor airflow distribution. *Energy and Buildings*, 230, 110525. https://doi.org/10.1016/j.enbuild.2020.110525

Li, Z., Dai, J., Chen, H., & Lin, B. (2019). An ann-based fast building energy consumption prediction method for complex architectural form at the early design stage. *Building Simulation*, 12(4), 665–681. https://doi.org/10.1007/s12273-019-0538-0

Kim, M., & Park, H.-J. (2023). Application of artificial neural networks using sequential prediction approach in indoor airflow prediction. *Journal of Building Engineering*, 69, 106319. https://doi.org/10.1016/j.jobe.2023.106319

Mokhtar, Sarah and Sojka, Aleksandra and Davila, & Carlos Cerezo. (2020). Conditional Generative Adversarial Networks for Pedestrian Wind Flow Approximation. In *Proceedings of the 11th Annual Symposium on Simulation for Architecture and Urban Design* (Vol. 58, pp. 1–8). San Diego, USA; Society for Computer Simulation International.

Hu, G., Liu, L., Tao, D., Song, J., Tse, K. T., & Kwok, K. C. S. (2020). Deep learning-based investigation of wind pressures on tall building under interference effects. *Journal of Wind Engineering and Industrial Aerodynamics*, 201, 104138. https://doi.org/10.1016/j.jweia.2020.104138

Huang, C., Zhang, G., Yao, J., Wang, X., Calautit, J. K., Zhao, C., An, N., & Peng, X. (2022). Accelerated Environmental Performance-driven urban design with generative Adversarial Network. *Building and Environment*, 224, 109575. https://doi.org/10.1016/j.buildenv.2022.109575

Duering, S., Chronic, A. & Koenig, R. (2020). Optimizing Urban Systems: Integrated optimization of spatial configurations. In *Proceedings of the 11th Annual Symposium on Simulation for Architecture and Urban Design* (Vol. 74, pp. 1–7). San Diego, USA; Society for Computer Simulation International.

He, Q., Li, Z., Gao, W., Chen, H., Wu, X., Cheng, X., & Lin, B. (2021). Predictive models for daylight performance of general floorplans based on CNN and Gan: A proof-of-concept study. *Building and Environment*, 206, 108346. https://doi.org/10.1016/j.buildenv.2021.108346

Faulkner, C. A., Jankowski, D. S., Castellini, J. E., Zuo, W., Epple, P., Sohn, M. D., Kasgari, A. T., & Saad, W. (2023). Fast prediction of indoor airflow distribution inspired by synthetic image generation Artificial Intelligence. *Building Simulation*, 16(7), 1219–1238. https://doi.org/10.1007/s12273-023-0989-1

Planfinder. PlanFinder. (n.d.). https://www.planfinder.xyz/

Tominaga, Y., Mochida, A., Yoshie, R., Kataoka, H., Nozu, T., Yoshikawa, M., & Shirasawa, T. (2008). AIJ guidelines for practical applications of CFD to pedestrian wind environment around buildings. *Journal of Wind Engineering and Industrial Aerodynamics*, 96(10–11), 1749–1761. https://doi.org/10.1016/j.jweia.2008.02.058

A PATTERN COLORING METHOD USING RANK-BASED INTERACTIVE EVOLUTIONARY ALGORITHM

ELİF GAMZE DEDELER[1] and MICHAEL S. BITTERMANN[2]
[1,2]Istanbul Technical University.
[1]dedeler19@itu.edu.tr, 0009-0006-5345-0146
[2]bittermann@itu.edu.tr, 0000-0003-3640-0754

Abstract. Interactive evolutionary algorithm (IEA) is a form of evolutionary computation designed to utilize information provided through human subjective assessments. This study proposes a design method based on an interactive evolutionary algorithm using a non-dominated sorting method that is well-known in the context of multi-objective evolutionary algorithm. The method developed is applied to a coloured geometric pattern as a graphic design element. The aim of this experiment is to investigate and evaluate a pattern coloring method using IEA that have the potential to augment creativity by providing robust support for exploration process of designers. The questions to the participants were designed to capture participants' perspectives on various aspects of the experiment, including satisfaction with the IEA, confidence in the exploration process, inspiration drawn from the generated designs in terms of visual pattern perception and color combinations. The results from the questionnaire showed that IEA can contribute the designers creative exploration process, become influential in visual perception of the patterns and supports initial design phase.

Keywords. Evolutionary algorithm, pattern, interactive design, pattern colouring, creative exploration

1. Introduction

Evolutionary algorithms, acknowledged as a stochastic search methodology, have been extensively explored and applied within the realm of engineering design. Notably, their benefits have transcended engineering domain and found relevance in diverse fields such as design. Whether optimizing engineering parameters or enhancing various design aspects, in both, purpose of evolutionary algorithms is to maximize the fulfilment of one or multiple objectives. IEAs, a subset of interactive evolutionary computation, aim to integrate human subjective evaluations into the optimization process. Unlike conventional evolutionary algorithms, interactive evolutionary algorithms (IEAs) emerge as a distinctive approach designed to bridge the gap between human and computer. This integration allows for the inclusion of tacit knowledge of human; including their psychology, emotions, preferences, and intuition.

In the realm of digital art and design, the quest for aesthetically pleasing pattern coloring methods are to be an area of exploration. The intersection of computational intelligence and artistic expression this paper presents a novel approach to pattern coloring, by using of a rank-based interactive evolutionary algorithm. The objective of the study is to empower individuals to explore diverse coloured pattern designs that they may not have been considered through conventional methods. In this approach, IEA is used to augment the aesthetic appeal of the pattern based on the user's aesthetic judgment and preferences.

2. Background

In the subsection 2.1, we gathered background information about patterns, color and visual perception. In the subsection 2.2, we described genetic algorithms, how designs are generated by using genetic algorithm, a specific type of of genetic algorithm, known as non-dominated sorting Genetic Algorithm (NSGA-II) which is used in this work. We also reviewed a number of interactive genetic algorithms studies focusing on the implementation.

2.1. PATTERN

2.1.1. Patterns in Graphic Design

Pattern implementation in graphic design has viewed as a pervasive and dynamic area of exploration. Early works by design theorists such as William Morris emphasized the significance of patterns in creating visual interest and conveying specific aesthetic intentions. In contemporary graphic design, patterns are employed across various mediums, from print to digital, to enhance visual appeal, communicate brand identity, and evoke emotional responses (Lupton & Phillips, 2015).

2.1.2. Role of Colors in Patterns

Colors play a pivotal role in the effectiveness and aesthetic impact of patterns. Research by Albers (1975) and Itten (1961) laid the groundwork for understanding color interactions within patterns, emphasizing the significance of color harmony and contrast. The selection and arrangement of colors contribute to the visual hierarchy within a pattern, guiding the viewer's focus and conveying specific meanings (Landa, 2019). Studies by Palmer and Schloss (2010) have investigated the subjective nature of color perception within patterns, suggesting that individual differences in color preferences can influence the overall visual experience.

2.1.3. Color Perception in pattern

The role of color in pattern perception is grounded in literature. Research by Itten (1961) and Albers (1975) laid the foundation for comprehending color interactions and their effects on visual perception.

Moreover, Adams (2012) and Smith et al. (2018) have explored cultural influences on color interpretation within patterns, revealing the ways in which societal norms and individual experiences contribute to the subjective understanding of color in different contexts.

2.2. GENETIC ALGORITHM

When a search problem involves objectives that are characterized by non-linearity, uncertainty or discreteness, point-by-point search lacks in robustness, so that parallel search strategies apply (Goldberg, 1989) Treating a problem by means of a GA conventionally requires expressing the design criteria in the form of one or multiple objective functions (Deb, 1995; Coello, Veldhuizen, and Lamont 2003). In the genetic algorithm framework, a design solution is represented as a chromosome, comprising individual genes. The collection of solutions constitutes the optimization's set of possible solutions, referred to as a population. Each gene within the chromosome denotes a value corresponding to a specific design variable. In the context of genetic algorithms, usually values represent in binary form, i.e., as a string of bits.

2.2.1. NSGA-II Algorithm

This work uses the non-dominated sorting genetic algorithm II (NSGA-II), a widely used genetic algorithm known for its effectiveness across diverse applications. NSGA-II, specifically designed when the objective functions are multiple. The essential mechanism for such algorithms to cope with the multitute of solution directions that arise due to multitute of objective functions, is a comparison among population members known as Pareto ranking. In the case of Pareto ranking the fitness values obtained from multiple objective functions and are used to establish a higher order fitness criteria, known as the degree of non-dominance. Non-dominance refers to the relative superiority of a solution compared to the rest of known possible solutions with respect to the simultaneous superiority with regards to all objective functions. Explicitly a solution A is said to dominate a solution B, when A is equally fit as B with respect to the objective functios, while at least in one of the objective functions A is superior to B. Then a solution C is said to have superior Pareto rank compare to solution D, when the number of solutions that dominate C is lower compared to the number of solutions that dominate D. With this definition of a higher order “fitness,” namely fitness in a multiobjective sense, it is clear that the “fitness,” of an individual solution is not independent from the other solutions.

2.2.2. Interactive Evolutionary Algorithm

A design problem presents challenges for computational design, especially while expressing aesthetical criteria. Since aesthetics cannot be translated into mathematical expressions, in the absence of a objective function, a genetic algorithm is to be implemented in interactive manner, by a human. In the realm of Interactive Genetic Algorithms (IGAs) applied to design, various studies have explored diverse domains. Leelathakul and Rimcharoen (2020) focused on ornamental motifs, employing a population size of 9 individuals who selected 1 out of 3 options nine times. Their assessment involved statistical analysis of shape features and questionnaire responses. Dou et al. introduced an IGA the context of car dashboard design. They utilized six real numbers in order to get input from designers judgement in their approach of customization of product. Brintrup et al. proposes an algorithm within the realm of furniture design to determine the ergonomics of a chair. They took into account both qualitative and quantitative objectives. Yoon and Kim, focused on generation of shape of buildings in a video game. They used 12 integers from the interval between 0 and 5 to format designer's judgement. Hernandez et al. proposed an interactive algorithm in order create solution for facility layout. They use 9 integer numbers from the interval of 1 and 5 as a human input. Gong et al. used interactive genetic algorithm for women's fashion elements with a multi-population approach. Their grading scale is in between 1 to 20 in order to format human preference. Kim and Cho, also studied IGA in the field of fashion design. The diverse dress styles, was displayed on a screen, and users assigned fitness values to each. Except for the Brintrup et al. and Quiroz et al. the interactivity is not to let the grading of chromosomes be accomplished by a human in place of a fitness function but to refine the application of the objective functions.

3. Methodology

This section presents our methodology for generating designs using IEA. The interactivity of proposed IEA and selection scheme developed in this work are explained.

3.1. INTERACTIVITY

Interactive ranking fitness function decided according to human preferences and judgements. In this approach, chromosomes within the population are ranked according to user preferences. The most favored chromosome receives the highest rank (Rank 1), the second-best obtains Rank 2, and the remaining non-preferred chromosomes are assigned Rank 3. This method establishes non-dominancy among chromosomes through human preferences. The algorithm is laid out in the Figure 1.

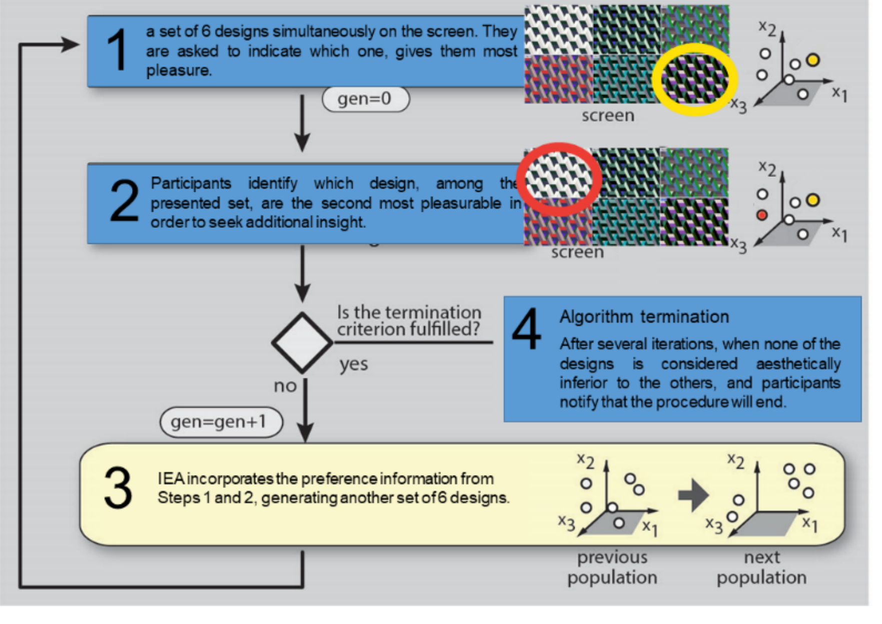


Figure 1. Interactivity Flow Chart of the proposed interactive evolutionary algorithm.

3.2. SELECTION SCHEMA

Aesthetic judgement is a measurement of a pleasure due to humans' perception. Given the absence of abstraction in this evaluative process, it is challenging on determining exact scores on an abstract scale but rather on identifying the design that offers a more aesthetically pleasing experience, in other words, which one is superior to the others.

Taking this perspective, the method of ranking based tournament selection, as it is implemented in NSGA-II, is deemed appropriate in this work since it allows comparative assessments.

3.3. BINARY TOURNAMENT SELECTION

Within genetic algorithms, the selection process can take different forms, among them two major ones are fitness proportionate selection and tournament selection. A key distinction lies in tournament selection, where the selection pressure is consistently maintained throughout the entire search process (Deb, 1995).

3.3.1. Tournament Formation

In each generation, 6 binary tournaments occurs, where two contenders compete to be the winner. Binary tournaments, a popular form of tournament selection (used in algorithms such as NSGA-II), involve determining a single winner per tournament. The winner contender is then copied into the mating pool for subsequent genetic operations, this process referred to as selection in genetic algorithm terminology.

The two contenders for winning the tournament are determined as follows: The first contender in the first tournament is pre-determined to be chromosome number one; in the same way, the first contender in the second tournament is chromosome number two;, the first contender in the ninth tournament is the chromosome number nine. This deterministic element is done to ensure that every chromosome at least is selected at one time into a tournament, so that it has at least some chance to be selected. The second contender, is randomly chosen from the remaining chromosomes to avoid repetition within the same tournament.

In a binary tournament featuring chromosomes A and B, the ranking hierarchy is established: Rank 1 is the highest, followed by Rank 2, and Rank 3 as the lowest. The tournament outcome is determined as follows: if Rank(A) < Rank(B), A wins; if Rank(A) = Rank(B), the probability of A winning is equal to the probability of A losing (both are 0.5); if Rank(A) > Rank(B), A loses.

It is important to note that in the case of a chromosome holding Rank 1 and being the sole occupant of this top rank, it is guaranteed to win at least one tournament. However, if there are multiple Rank 1 solutions, it is possible that a particular Rank 1 solution may not participate in next generation, depending on its opponents, if it contends against another.

4. Experiment

In this experiment, we utilize a geometric pattern composed of multiple geometric elements. A pattern with geometric elements are selected due to its regular arrangement of geometric shapes. Each geometric element within the pattern is assigned a color parameterized in the HSL color model using Grasshopper. Each shape has following paremeters: hue, saturation, luminance and alpha value between 0 and 1. At the first phase of the experiment, participants experiment colour combinations on the pattern by changing slider values for each shape. A total number of 6 geometric shapes with 4 different slider values, makes 24 number of parameters. At the second phase, the experiment with IEA was implemented as the steps followed which are mentioned in the section 3.1. Through the experiment, the distribution index (n_c) used by the SBX (Simulated Binary Crossover) operator is kept constant. The n_c distribution index significantly influences convergence speed; smaller values of n_c result in slower convergence as offspring solutions move away from their parents. In our experiment, we set $n_c = 2$ to foster a variety of design solutions that are distinct from their parents. Mutation probability set as 0.1. The mutation operation introduces a parameter called the mutation index which is n_m (Deb, 2001). When n_m is set to a lower value, the diversity of mutated solutions increases, and they deviate more from their parent solutions. This is the rationale behind setting n_m to 10, aiming to enhance the diversity among mutated solutions and reduce their resemblance to the parent solutions. These values are determined by a number of test runs using trial and error, not by systematic tuning. In Figure 2, one of the experiment screen is shown.

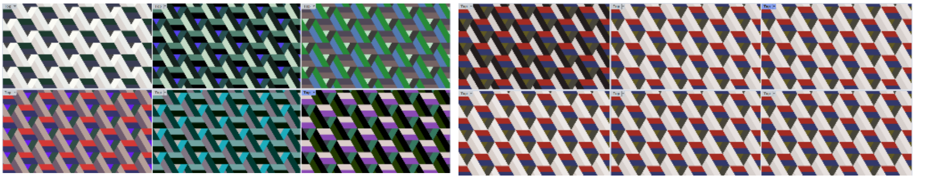


Figure 2. First generation (left), last generation (right) of one of the experiments

5. Results

5.1. OVERVIEW OF THE SURVEY

A total of 27 participants took part in the survey, representing diverse professional backgrounds, including graphic designers, artists, architects, interior architects, and industrial designers. The participant breakdown included 3 undergraduate students, 8 graduate students, and 16 professionals. Among the professionals, 10 had 1-5 years of experience, 4 had 5-10 years of experience, and 2 had more than 10 years of experience in their respective fields. They are asked the following questions:

- How satisfied are you with the generated colored patterns?
- Do you find the number of design solutions provided sufficient? Would you prefer fewer or more choices? If so, why?

- Do you believe that IEA can take your aesthetic judgement? (Did the colored patterns become more aesthetically pleasing as the search progressed? and/or Do the patterns become similar to each other when the search goes on?)
- When comparing the process of coloring patterns using the conventional method (selecting each parameter individually) with the interactive algorithm, how do the two techniques differ in terms of your confidence in throughly exploring possible solutions?
- Did the colored pattern generated solutions inspire you? If so, how?
- Do you think that IEA could contribute your explorative creative process?
- When exposed to different coloring combinations on the pattern, do you perceive them differently? How does your perception affected?
- Do you have any additional comments or suggestions regarding your experience with the IEA algorithm in pattern coloring and the overall experiment?

5.2. RESULTS OF THE SURVEY

5.2.1. Satisfaction

Participants were asked about their satisfaction with the generated coloured patterns. Out of the total 27 participants, 2 were very satistfied (7.4%), 14 participants reported being satisfied (51.9%), 9 indicated a neutral stance (33.3%), and the remaining 2 participants (7.4%) reported dissatisfaction with the experience.

In response to the question about the sufficiency of design solutions provided, the participants' preferences varied. Out of the total 27 participants, one participant (3.7%) expressed a preference for fewer design choices, citing a desire to concentrate more on the pattern. Conversely, 12 participants (44.4%) expressed a preference for more design solutions, expressing a desire for increased diversity in the design set. The majority, comprising 14 participants (51.9%), found the number of design solutions provided to be satisfactory.

In response to the question regarding the capability of IEA to take aesthetic judgment, a significant majority, comprising 22 participants (81.5%), affirmed that they believed the IEA could take their aesthetic judgement. In contrast, 3 participants (11.1%) expressed a neutral stance on the matter, while 2 participants (7.4%) disagreed with the notion that the IEA could effectively take their aesthetic judgement. Participants who held a neutral stance on this question provided feedback indicating that, through the search, although the designs become more similar to each other, the designs did not entirely align with their individual aesthetic preferences. Furthermore, the participants that agreed with this question noted that as an overall, the designs were progressively becoming more aesthetically pleasing, however, they expressed a desire for more pleasing options.

5.2.2. Exploration

In comparison of the conventional parametric method (selecting each parameter individually) with the IEA, the participants provided insights in their confidence regarding exploring possible solutions, over the 27 participants, 24 of them (88.89%) agreed on that. The conventional parametric method, which involved selecting each parameter individually using sliders, deemed confusing by the respondents. Despite control, participants found the process tiresome, particularly due to high number of parameters involved. In contrast, the IEA method was appreciated by the participants since the generations produced by the IEA were explorative. 3 of the participants (11.11%) commented IEA is not exploratory due to high speed of convergence.

For the question whether participants found the IEA process inspirational or not, the majority of them, 25 out of 27 respondents (92.59%) agreed on that the IEA method inspire them while 2 of the respondents (7.41%) stay neutral. None of the participants disagreed with it. Some of them reported that interacting with generated colour options on the triggered imaginative thinking and influenced their own creative process.

The view on whether the IEA could contribute to the explorative creative process varied among participants. Out of 27 participants, 9 of them (33.3%) responded agree, 12 of them (44.4%) responded partially agree and 4 of them (14.81%) had neutral stance. Although the majority partially or not partially agreed on the contribution of IEA, 2 of the participants (7.41%) expressed a preference for selecting the color pallette before the design process and, as a result, did not feel that the IEA significantly contributed to their creative design process. On the other hand, another participant highlighted seeing different options as a contribution to design. The ability to explore various design alternatives through the IEA was considered a positive influence on their explorative creative process. Another view from the participants was noted that while the IEA was inspirational, it couldn't fully lead the design process; its impact was more pronounced during the initial conceptual phase.

In the study with 27 respondents, we investigated the impact of various coloring combinations applied on pattern as visual perception. 26 of the participants (96.29%) reported participants did, in fact, perceive patterns differently when exposed to different coloring combinations. (20 of them responded " Significantly enhances perception and 6 of them responded "enhances perception") They commented that variations in color seemed to have an effect on their overall perception so that it become inspirational in the design process.

5.2.3. Suggestions

The feedback from participants regarding their experience with the IEA algorithm in pattern coloring and the overall experiment highlighted a desire for increased options. In that manner some of the participants suggested that conducting the experiment for multiple times could be beneficial due to the potential wideness of the exploration space. Additionally, a few participants expressed the idea that it would be helpful if they could remove designs they did not like. Another participant provided feedback suggesting the option to limit certain parameters based on their design criteria.

6. Discussion

6.1. RANDOMIZED SOLUTIONS

Regarding slider-based design, the starting point within the decision variable space is arbitrary, with a single trajectory. In case of interactive evolutionary algorithm, the starting points are also arbitrary, and more specifically they are random. Furthermore, there are numerous starting points, ensuring that randomness covers the entire search domain. The search process within the space, due to the parallel evaluation of designs, implies the simultaneous pursuit of multiple trajectories. In each generation, in other words, at every turn of a trajectory in decision variable space, the designer's evaluation information comes into the process. Among each turning point, those that are relatively close to a desirable design solutions are distinguished from those that are relatively distant. As the search is a multi-trajectory search, the probability of reaching a desirable region should be somewhat higher than in case of the single trajectory search, which parametric design is. Participants commented that the presence of multiple sliders complicates the design process. Consequently, they believe that the IEA method has the potential to be influential, considering it a beneficial tool for conceptual exploration.

6.2. INITIAL POPULATION

The randomly generated initial population may not be satisfactory for the participants. However, due to the experimental evaluation process, participants are still required to choose designs based on their aesthetic judgments among the alternatives. This could potentially misdirect the algorithm in a wrong trajectory. Consequently, some participants commented that, despite the algorithm could take their preference information, the results did not entirely reflect their aesthetic preferences or the designs could be better. Selecting worst design or conduct the experiment with a different the initial population could be helpful to address this issue.

6.3. POPULATION SIZE

Determining the population size is a crucial aspect of interactive evolutionary computational processes. The population size should not be too large to maintain user's focus on the task, yet not too small, as a low population hinders the user's ability to explore diverse alternatives for solutions. The limited number of design solutions creates a lack of diversity, therefore the randomized solutions might direct the process as it is biased. In order to prevent from the disadvantage of low population size, the experiment could be conducted multiple times.

7. Conclusion

The intricacy of design arises from the simultaneous evaluation of multiple parameters and the incorporation of tacit knowledge, which are encompassing human preferences and judgments. To address this challenge, the interactive evolutionary approach emerges as a unique solution. This is primarily because the interactive evolution incorporates the judgment of preference directly into the computational process, eliminating the need for an explicit explanation of the preference.

We had applied a rank-based interactive evolutionary algorithm (IEA) to generate colour combinations for patterns. Exploring each color applied to the pattern is a challenging task because visual perception is influenced. In the examplary design study, it addresses the challenge of considering multiple color parameters for each geometric element of the pattern. This task is inherently complex due to the intricate relationships between colors, geometric shapes and the number of parameters. In this regard, the IEA stands out as an effective process for design exploration for color selection and visual appeal in the final pattern. When compared to conventional parametric design methods, experimental results showed that this approach facilitates creative exploration and inpire the initial phases of design due to IEA's ability to start with randomized design solutions and due to fitness according to human preferences.

References

Albers, J. (1975). Interaction of Color. *Yale University Press.*

Brintrup, A. M., Ramsden, J., Takagi, H., & Tiwari, A. (2008). Ergonomic chair design by fusing qualitative and quantitative criteria using interactive genetic algorithms. *IEEE Transactions on Evolutionary Computation*, 12(3), 343-354.

Deb, K. (2001). Multi-objective optimization using evolutionary algorithms. *John Wiley & Sons.*

Deb, K., & Beyer, H. G. (2001). Self-adaptive genetic algorithms with simulated binary crossover. *Evolutionary computation*, 9(2), 197-221. https://doi.org/10.1162/106365601750190406

Dou, R., Zong, C., & Nan, G. (2016). Multi-stage interactive genetic algorithm for collaborative product customization. *Knowledge-Based Systems*, 92, 43-54. https://doi.org/10.1016/j.knosys.2015.10.013

Goldberg, D. E. (1989). Genetic Algorithms. *Addison-Wesley*: Massachusetts.

Gong, D.-W., Hao, G.-S., Zhou, Y., & Sun, X.-Y. (2007). Interactive genetic algorithms with multi-population adaptive hierarchy and their application in fashion design. *Applied Mathematics and Computation*, 185(2), 1098-1108.

Hernandez, L. G., Morera, L. S., & Azofra, A. A. (2011). An Interactive Genetic Algorithm for the Unequal Area Facility Layout Problem. *Soft Computing Models in Industrial and Environmental Applications, 6th International Conference SOCO 2011* (pp. 253-262). Advances in Intelligent and Soft Computing (AINSC), 87.

Itten, J. (1961). The Art of Color: The Subjective Experience and Objective Rationale of Color. *John Wiley & Sons.*

Kim, H. S., & Cho, S. B. (2000). Application of interactive genetic algorithm to fashion design. *Engineering Applications of Artificial Intelligence*, 13(6), 635-644.

Landa, R. (2019). Color, Space, and Style: All the Details Interior Designers Need to Know but Can Never Find. *Rockport Publishers.*

Leelathakul, N., & Rimcharoen, S. (2020). Generating Kranok patterns with an interactive evolutionary algorithm. *Applied Soft Computing*, 89, 106121.

Lupton, E., & Phillips, J. C. (2015). Graphic Design: The New Basics.

Quiroz, J. C., Louis, S. J., Shankar, A., & Dascalu, S. M. (2007). Interactive Genetic Algorithms for User Interface Design. *2007 IEEE Congress on Evolutionary Computation* (pp.1366-1373). IEEE. DOI: 10.1109/CEC.2007.4424630

Yoon, D.-M., & Kim, K.-J. (2013). Interactive Evolution of 3D Models based on Direct Manipulation for Video Games. *Procedia Computer Science*, 14 (1), 1-16. https://doi.org/10.1016/J.PROCS.2013.10.036

A RAPID PREDICTION MODEL FOR VIEW-BASED GLARE PERFORMANCE WITH MULTIMODAL GENERATIVE ADVERSARIAL NETWORKS

XIAOQIAN LI[1], ZHEN HAN[2], GANG LIU[3] and RUDI STOUFFS[4]
[1,2,3] *School of Architecture, Tianjin University.*
[1,2,4] *Department of Architecture, National University of Singapore.*
[1]*lixiaoqian_95@tju.edu.cn, 0009-0009-0160-6817*
[2]*zhen.han@u.nus.edu, 0000-0002-9547-9063*
[3]*liug@tju.edu.cn, 0000-0002-7864-7846*
[4]*stouffs@nus.edu.sg, 0000-0002-4200-5833*

Abstract. Machine learning-based glare prediction has greatly improved the efficiency of performance feedback. However, its limited generalizability and the absence of intuitive predictive indicators have constrained its practical application. In response, this study proposes a prediction model for luminance distribution images based on the multimodal learning approach. This model focuses on objects within the field of view, integrating spatial and material features through images. It also employs semantic feature mapping and multimodal data integration to flexibly represent building information, removing limitations on model validity imposed by changes in design scenarios. Additionally, the study proposes a multimodal Generative Adversarial Network tailored for the multimodal inputs. This network is equipped with unique feature fusion and reinforcement blocks, along with advanced up-sampling techniques, to efficiently distill and extract pertinent information from the inputs. The model's efficacy is verified by cases focusing on residential building luminance distribution, with a 97% improvement in computational speed compared to simulation methods. Offering both speed and accuracy, this model provides designers with a rapid, flexible, and intuitive supporting approach for daylight performance optimization design, particularly beneficial in the early design stage.

Keywords. Glare Prediction, Prediction Model, Multimodal Model, Generative Adversarial Networks

1. Introduction

Daylight is crucial for indoor environments; it offsets electric lighting costs and promotes the health and well-being of occupants (Al Horr et al., 2016). However, uncontrolled penetration of daylight into buildings could lead to undesirable luminous environments, impairing vision or causing visual discomfort. Therefore, glare autonomy is a vital part of improving the quality of building indoor environments.

However, during the early stages of building design, assessing glare presents a unique computational challenge because of complex and time-consuming rendering and calculations. This makes it hard for designers to receive real-time performance feedback during the process of selecting or improving design schemes, limiting its application in practical design (Jones et al., 2018). Three trends in accelerating glare evaluation research are 1) Idealizing the light transmission process or screening out focus areas for simplified calculations (Giovannini et al., 2020); 2) Accelerating rendering by tracing multiple primary rays in parallel on a Graphics Processing Unit (Jones, 2019); 3) Implementing rapid feedback with predictive models based on Machine Learning algorithms (Ayoub, 2019). Predictive models are increasingly favored for their cost-effectiveness and rapid processing capabilities.

Some relevant papers are reviewed below (Table 1). Predictions mainly focus on numerical evaluation metrics, such as Daylight Glare Probability (DGP), vertical eye illuminance (E_v), or annual metrics Annual Sunlight Exposure (ASE). However, there is less research on intuitive image-based indicators, such as luminance distribution maps, even though intuitive visual formats are more readily accepted and understood, matching better with the requirements of the design process. Moreover, the generalizability of prediction models is a key focus for users, as it determines models' adaptability to various design scenarios. However, it is a significant limitation in existing prediction models (Ngarambe et al., 2022). Crucial reasons for this are, besides the network structure's inherent attributes, the selection of design elements inputted and their feature representation methods. These are crucial as they dictate whether solutions with significant variations can be effectively translated through a uniform rule. In reviewed research, inputs typically consist of geometry, materials, and environment features, which are converted into numerical variables for the network's inputs (Pierson et al., 2018). However, this approach's adaptability is typically suboptimal, for instance, failing to accommodate diverse spatial forms or material configurations. One potential reason is the reliance on single-modal data, which is inadequate for representing all features comprehensively. In contrast, multimodal data, integrating the advantages of different data types, has proven highly effective in other research areas. This approach, however, is seldom employed in the field of building performance prediction, suggesting an untapped potential for multimodal data to enhance the generalizability of representing building information.

Therefore, to develop a sufficiently generalized and intuitive method for glare prediction, this paper proposes a multimodal feature representation method combined with semantic information mapping and its corresponding network structure. This method employs images to depict features of sub-aspects within the field of view, such as the type, material, and location of surfaces, and combines these with numerical data on environmental conditions and viewpoints. This combination allows for a versatile portrayal of the spatial features, adaptable to different design scenarios. Furthermore, based on Generative Adversarial Networks (GANs), the proposed approach includes operations such as multimodal feature fusion, key feature reinforcement, and refined up-sampling to satisfy the requirements of a network with multimodal inputs. This model enables the rapid generation of image-based glare evaluation from any viewpoint, effectively adjusting to any changes in building details, which provides efficient and effective support for designers.

Table 1. Machine learning-based prediction model for indoor glare prediction

Ref.	Inputs			Output			Algorithm
	Variables			Perfor-mance metrics	Data type		
	geome-try	mate-rial	wea ther		numer-ical	image	
Xie et al. 2023	√	×	×	DGP	√	×	GBRT
Radziszewski et al., 2018	√	×	√	DGP	√	×	ANN
Xie et al. 2021	×	×	√	DGP	√	×	KNN, RF
Liu et al.2020	×	×	√	Iuminance	×	√	DNN
Luo et al. 2022	√	×	√	Ev	√	×	ANN
Lin et al.2021	√	×	√	ASE	√	×	ANN
Nourkojouri et al. 2021	√	√	×	ASE	√	×	ANN
Ayoub et al.2019	√	×	×	ASE	√	×	ANN
Mostafavi et al. 2022	√	×	×	Iuminance	×	√	GAN

2. Methods

2.1. SEMANTIC INFORMATION MAPPING

The essence of semantic information mapping is to utilize pixel brightness and spatial attributes to detail interior surfaces, including their type, material, shape, and location. This mapping is grounded in the imagery captured from a single viewpoint by a perspective camera. The approach narrows the learning focus to the field of view, rather than the entire space, and the conversion to images impervious to spatial form alterations. Subsequently, the view field content is segmented into subregions. Each is tailored to different semantic objects based on surface types. These segments are merged with their respective material properties. This fusion forms the final semantic information matrix, a cohesive representation of spatial and material attributes. Additionally, capturing imagery with a fisheye camera typically requires rendering a 3D model. Due to the fisheye camera's broader range compared to a perspective camera, this study introduces a technique to grid the hemisphere in the viewing direction to reduce computational load, i.e., merging multiple perspective views to simplify fisheye capture. The specific calculation steps are as follows:

Step 1: Construct the 3D building model, identifying and color-coding various surface attributes.

Step 2: Determine the target viewpoint and its direction. Construct a hemisphere centered at the viewpoint and perpendicular to the line of sight. Grid the hemisphere based on precision requirements. Create sub-viewing directions from the viewpoint to the center of each grid, set the camera's focal length, and capture interior images along these viewing directions. The details are illustrated in Figure 1.

Step 3: Based on captured images, construct semantic matrixes that categorize the types of interior surfaces.

Step 4: Identify the material properties of each surface type and map them in situ to the semantic matrix, creating the image input for this model. It is important to note that transparent materials, such as glass, may cause an overlap of multiple types of surfaces, for instance, seeing another wall through a window. In such cases, their material properties are a blend of multiple materials, as detailed in Table 2.

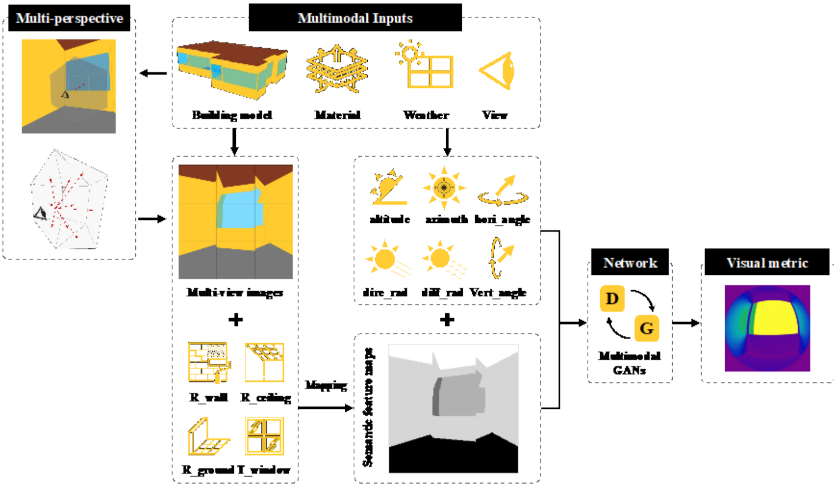


Figure 1. The framework of the prediction model for view-based glare performance.

Table 2. Definitions of surface material properties.

Surface type	Material index
Window+Wall	window_t * wall_r
Window+Window+Wall	window_t * window_t * wall_r
Window+Window+Ceiling	window_t * window_t * celing_r
Window+Window+Floor	window_t * window_t * floor_r

2.2. MULTIMODAL INPUT VARIABLES

In addition to the spatial and material features, environmental conditions and viewpoint information are also key factors influencing glare. The environmental variables are chosen based on existing research. Solar altitude and azimuth angles determine the daylight level, while sky conditions are abstracted to direct normal radiation and diffuse horizontal radiation, reflecting the external daylight transmission. Additionally, the horizontal and vertical angles of the viewpoint are considered to illustrate the relationship between the view direction and the angle of incoming daylight. These six variables, quantified numerically, combine with the semantic matrix to form the multimodal input for the prediction model, as shown in Figure 1.

2.3. MULTIMODAL GENERATIVE ADVERSARIAL NETS

In this study, a novel multimodal GAN (mGAN) is proposed for multimodal inputs. This framework is constructed based on the Pix2Pix model (Isola, 2017), which consists of a Generator and a Discriminator, receiving multimodal inputs composed of images and vectors. The network structure is illustrated in Figures 2 to 4.

To accommodate multimodal inputs, the mGAN incorporates a vector-based feature encoding and a feature fusion mechanism. Vector features are sequentially input into the network and transitioned from a low-dimensional latent space to a higher spatial dimension through a series of transposed convolution blocks, aligning their dimensions with image features. Simultaneously, image features are exacted with a series of convolution blocks, progressively increasing their channel count, integrating more refined information, and preserving spatial structure. Finally, the Hadamard product is used to fuse the extracted vector and image features.

Beyond the standard convolution and transposed convolution blocks in Pix2Pix, the Generator's up-sampling integrates PixelShuffle blocks with transposed convolution blocks. PixelShuffle blocks upscale by rearranging channel data instead of inserting blank pixels, transforming feature maps from lower resolution with more channels to higher resolution with fewer channels. This approach maintains uniform pixel distribution and spatial consistency during up-sampling, thus avoiding the checkerboard effect often associated with transposed convolution only. This blend of upscaling techniques enables the model to flexibly manage feature representation and information flow across various layers, balancing learning capacity, image quality, and computational efficiency.

Additionally, to solve the considerable dimensional discrepancy between image and vector inputs, the Generator of the mGAN, besides its existing skip connections, has added reinforcement connections. This involves exacting vector-based inputs with various depths and incorporating these multi-level deep features as part of the resources into the corresponding up-sampling location. This strategy is designed to prevent the loss of vector information due to excessive dimensional differences. During the feature extraction process, inception blocks from GoogleNet are used, utilizing convolution kernels of different scales for multi-scale feature refinement, thereby enriching detail.

The standard pix2pix loss function is used in the mGAN. The total loss function is a weighted sum of the adversarial loss and the L1 loss to improve the generalizability and robustness of the network.

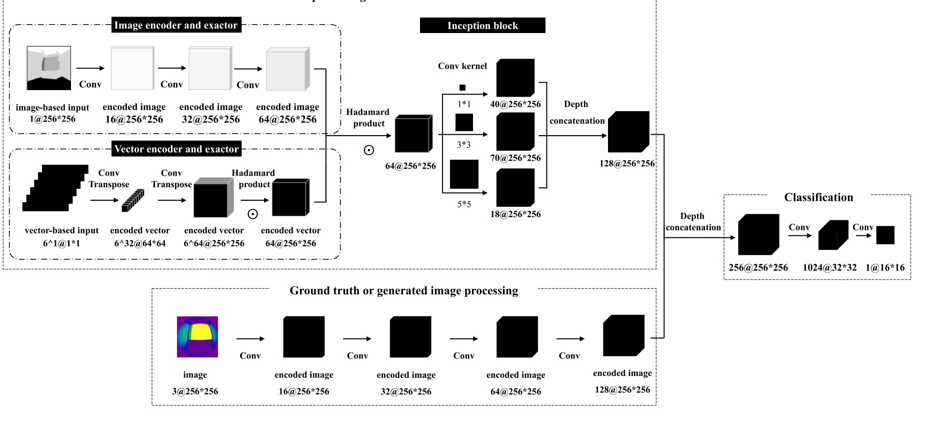


Figure 2. The construction of the Discriminator.

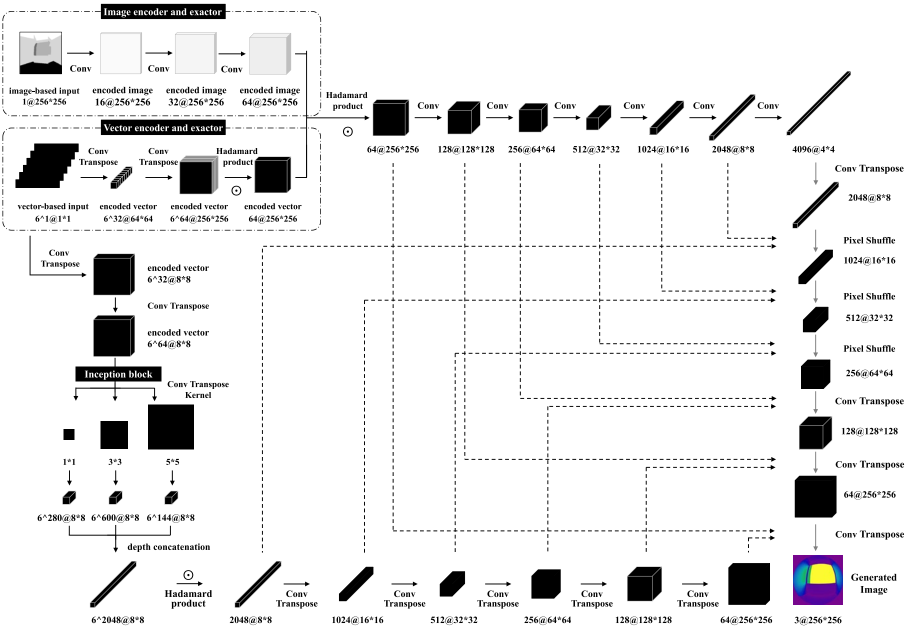


Figure 3. The construction of the Generator.

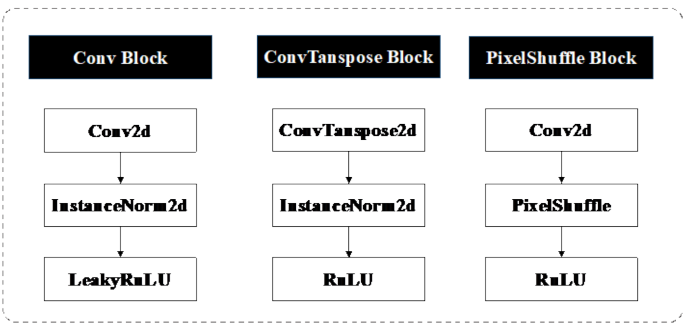


Figure 4. The construction of key blocks in the mGAN.

3. Evaluation of the prediction model

To test the performance of the proposed glare prediction model, it has been applied to predict the luminance distribution within residential buildings.

The data sources for this study comprise a building form scheme dataset, a material dataset, and a weather dataset. The scheme dataset employed in this study is RPLAN, an Asian residential building floor plan dataset assembled by Wu et al. in 2019. 5000 building models were generated with Grasshopper (single-wall models for daylight simulation). Material attribute parameters are guided by the Chinese Daylighting Design Code, involving the reflectance of the walls, ceiling reflectance, floor reflectance, and the transmittance of windows. The weather data is sourced from twelve cities in China, each belonging to one of five different light climate zones, and covers the time range from 8:00 AM to 6:00 PM. Data from these multiple sources are randomly selected and integrated to develop a simulation model, which is utilized for generating the learning target with Ladybug.tools. To standardize the visualization mapping relationship, the display lower boundary for cloud maps is set to 0 lux, with an upper boundary of 2000 lux.

The dataset comprises 4,800 samples, allocated in an 8:2 ratio for training and testing purposes. To improve training efficiency, the resolution of the input images has been downscaled to 256x256 pixels. Additionally, numerical features have been normalized to a range between 0 and 1, ensuring a balanced influence on the overall data distribution. The detailed hyperparameter settings are shown in Table 4.

The performance of the prediction model is evaluated in terms of both data accuracy and quality of the images. The accuracy metrics include MAE (Mean Absolute Error) and MAPE (Mean Absolute Percentage Error), while image quality is assessed through SSIM (Structural Similarity Index Measure) and Image Difference Analysis (IDA).

Table 3. Ranges of material, weather and view variables.

Variables of material	Range
Reflection rate of wall (wall_r)	[0.3, 0.7]
Reflection rate of ceiling (celing_r)	[0.5, 0.9]
Reflection rate of floor (floor_r)	[0.1, 0.5]
Transmittances of windows(window_t)	[0.5, 0.8]
Solar azimuth angle(az_sun) Solar altitude angle (al_sun) Direct normal radiation (dir_rad) Diffuse horizontal radiation (dif_rad)	Data from Beijing, Shanghai, Chengdu, Guangzhou, Harbin, Hohhot, Kunming, Lhasa, Shenyang, Urumqi, Xian, Xining
View azimuth angle(az_sun)	[0, 360]
View altitude angle (al_sun)	[0, 360]

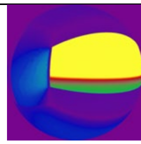
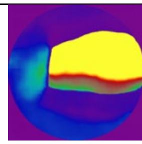
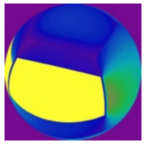
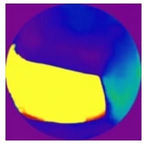
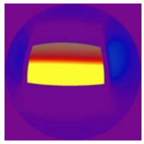
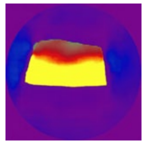
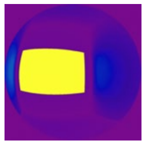
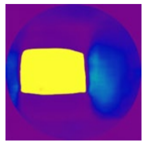
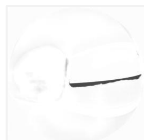
Table 4. Hyperparameters of training.

Parameters	Learning Rate_D	Learning Rate_G	Batch Size	n_epoch	Optimizer
Settings	Initial value: 0.0005 Drop period: 5 Drop rate: 0.3	Initial value: 0.0015 Drop period: 5 Drop rate: 0.3	8	200	Adam

4. Results

The Discriminator converges after around 120 epochs, while the Generator converges after around 140 epochs. Table 5 provides an overview of various cases from this test set, involving the inputs, ground truth, predicted images, and IDA maps. Additionally, Figure 6 illustrates MAE, MAPE, and SSIM for the test set. The training set shows an average MAE of 3.654, an average MAPE of 0.124, and an average SSIM of 0.919. In contrast, the test set has an average MAE of 18.233, an average MAPE of 0.466, and an average SSIM of 0.814.

Table 5. Example results of the test set.

Imgae inputs	Vector inputs		Ground truth	Prediction	IDA
	al_sun:57.231 az_sun:110.561 dir_rad:792	dif_rad:30 az_view:285.109 al_view:89.790			
	al_sun:33.153 az_sun:197.812 dir_rad:753	dif_rad:152 az_view:7.571 al_view:104.526			
	al_sun:55.426 az_sun:110.503 dir_rad:621	dif_rad:200 az_view:20.579 al_view:87.230			
	al_sun:74.489 az_sun:241.523 dir_rad:487	dif_rad:242 az_view:227.846 al_view:99.916			

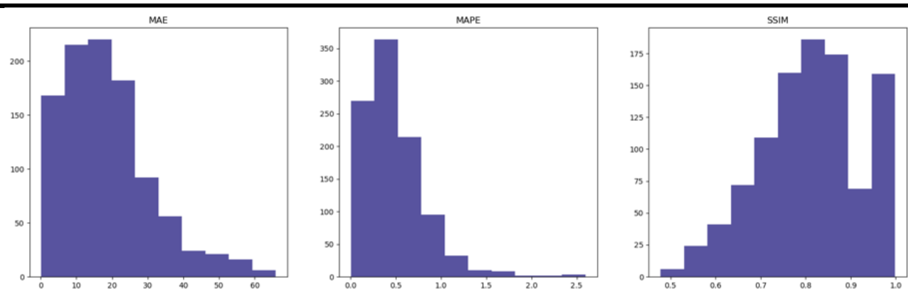


Figure 5. Accuracy indicators of the test set.

5. Discussion

The efficacy of the proposed prediction model has been verified through the case study, and its strengths and limitations deserve further discussion.

1) Time gap with traditional simulation: The proposed model's calculation time involves the time required for capturing the view content, constructing the semantic feature matrix, and generating the predictive image. Constructing the semantic feature matrix is the most time-consuming part. For example, in a 9-viewpoint splice case, the average processing time is about 7.6s, while the other steps take less than 1s. By comparison, simulations for the same dataset range from 276 to 432s, averaging 354s. This model achieves a 97% improvement in speed compared to traditional simulations.

2) Generalizability of the proposed prediction model: This model focuses on

objects within the field of view rather than the entire building. The internal surfaces' types, shapes, and locations are represented as images, thus remaining unaffected by structural changes. Material information, integrated into the images through semantic feature mapping, allows the model to adapt various material setups. The environmental setting is distilled into essential weather parameters, releasing the model from climatic constraints. Overall, this model enhances generalizability in spatial, material, and environmental domains, catering to various design scenarios.

3) Partial sample with poor performance: A few predicted results with a large error. There are two potential reasons. The first is overfitting within mGANs, suggested by the superior performance of the training set over the test set. Despite efforts to enhance model generalizability — like increasing sample size, adding regularization, and using robust layers — overfitting remains a challenge. The second reason might be the partial loss of vector-based feature information during processing. In some low-performance samples, while the surface blocks appear similar, great differences in the luminance distribution of external windows are observed. This could be due to the model's limited ability to learn weather features, attributed to a dimensional gap between vector and image features. Addressing these two issues is crucial for future work in this research.

6. Conclusion

In this paper, a novel rapid predictive model for glare is proposed, designed to quickly visualize glare conditions without limitation from design scenario variations. In this model, building spaces are represented as standardized viewpoint snapshots. Semantic feature mapping is employed to integrate spatial characteristics and material properties. Images containing comprehensive information, along with vectors that reflect environmental conditions, constitute the multimodal input of this model. To fulfill them, mGANs have been developed, focusing on multimodal features. The network integrates feature fusion, feature reinforcement blocks, and advanced upscaling methods, enabling it to effectively eliminate redundant information from multimodal inputs, deeply extract features, and generate predictive images. The model's efficacy is verified through case studies on luminance distribution prediction in residential buildings. Compared to simulation methods, it achieves a 97% improvement in computational speed. Overall, the model provides designers with a faster, more flexible, and intuitive glare visualization method, useful in scenarios such as requiring quick evaluation for the performance of numerous schemes or rapid intuitive understanding of glare conditions. Currently, this model exists as a conceptual prototype. Future research will further refine the overall process and develop related associated tools based on popular design platforms to truly support design practice.

Acknowledgements

The first author benefited from a China Scholarship Council grant (202206250073).

References

Al Horr, Y., Arif, M., Kaushik, A., Mazroei, A., Katafygiotou, M., & Elsarrag, E. (2016). Occupant productivity and office indoor environment quality: A review of the literature. Building and environment, 105, 369-389. https://doi.org/10.1016/j.buildenv.2016.06.001

Ayoub, M. (2019). A multivariate regression to predict daylighting and energy consumption of residential buildings within hybrid settlements in hot-desert climates. Indoor and built environment, 28(6), 848-866. https://doi.org/10.1177/1420326X18798164

Giovannini, L., Favoino, F., Verso, V. R. M. L., Serra, V., & Pellegrino, A. (2020). GLANCE (GLare ANnual Classes Evaluation): An approach for a simplified spatial glare evaluation. Building and Environment, 186, 107375. https://doi.org/10.1016/j.buildenv.2020.107375

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).

Jones, N. L. (2019, September). Fast climate-based glare analysis and spatial mapping. In Proceedings of building simulation 2019: 16th Conference of IBPSA.

Jones, N. L., & Reinhart, C. F. (2019). Effects of real-time simulation feedback on design for visual comfort. Journal of Building Performance Simulation, 12(3), 343-361. https://doi.org/10.1080/19401493.2018.1449889

Liu, Y., Colburn, A., & Inanici, M. (2020). Deep neural network approach for annual luminance simulations. Journal of Building Performance Simulation, 13(5), 532-554. https://doi.org/10.1080/19401493.2020.1803404

Luo, Z., Sun, C., Dong, Q., & Qi, X. (2022). Key control variables affecting interior visual comfort for automated louver control in open-plan office--a study using machine learning. Building and Environment, 207, 108565 https://doi.org/10.1016/j.buildenv.2021.108565

Lin, C. H., & Tsay, Y. S. (2021). A metamodel based on intermediary features for daylight performance prediction of facade design. Building and Environment, 206, 108371. https://doi.org/10.1016/j.buildenv.2021.108371

Mostafavi, F., Tahsildoost, M., Zomorodian, Z. S., & Shahrestani, S. S. (2022). An interactive assessment framework for residential space layouts using pix2pix predictive model at the early-stage building design. Smart and Sustainable Built Environment. https://doi.org/10.1108/SASBE-07-2022-0152

Ngarambe, J., Adilkhanova, I., Uwiragiye, B., & Yun, G. Y. (2022). A review on the current usage of machine learning tools for daylighting design and control. Building and Environment, 109507. https://doi.org/10.1016/j.buildenv.2022.109507

Nourkojouri, H., Shafavi, N. S., Tahsildoost, M., & Zomorodian, Z. S. (2021). Development of a machine-learning framework for overall daylight and visual comfort assessment in early design stages. Journal of Daylighting, 8(2), 270-283. https://doi.org/10.15627/jd.2021.21

Pierson, C., Wienold, J., & Bodart, M. (2018). Review of factors influencing discomfort glare perception from daylight. Leukos, 14(3), 111-148. https://doi.org/10.1080/15502724.2018.1428617

Radziszewski, K., & Waczyńska, M. (2018). Machine learning algorithm-based tool and digital framework for substituting daylight simulations in early-stage architectural design evaluation. In Proceedings of the Symposium on Simulation for Architecture and Urban Design (pp. 1-7).

Xie, F., Song, H., & Zhang, H. (2023). Research on Light Comfort of Waiting Hall of High-Speed Railway Station in Cold Region Based on Interpretable Machine Learning. Buildings, 13(4), 1105. https://doi.org/10.3390/buildings13041105

Xie, J., & Sawyer, A. O. (2021). Simulation-assisted data-driven method for glare control with automated shading systems in office buildings. Building and Environment, 196, 107801. https://doi.org/10.1016/j.buildenv.2021.107801

Wu, W., Fu, X. M., Tang, R., Wang, Y., Qi, Y. H., & Liu, L. (2019). Data-driven interior plan generation for residential buildings. ACM Transactions on Graphics (TOG), 38(6), 1-12. https://doi.org/10.1145/3355089.3356556

ACCELERATING FUTURE SCENARIO DEVELOPMENT FOR CONCEPT DESIGN WITH TEXT-BASED GENAI (CHATGPT)

LINUS TAN[1] and THOM LUKE[2]
[1] *Swinburne University of Technology*
[2] *RMIT University*
[1] *linustan@swin.edu.au, 0000-0002-5705-0493*
[2] *thom.luke@rmit.edu.au, 0009-0003-6317-5879*

Abstract. This case study describes the integration of Generative Artificial Intelligence (GenAI) into a design workflow that envisions future scenarios for concept development. While image-based GenAI tools like MidJourney and Stable Diffusion have garnered attention from designers for their ability to visualise ideas rapidly, integrating textual GenAI, like ChatGPT-3.5, in design workflows has been less explored. This case study investigates how future thinking techniques can be digitized and accelerated using ChatGPT-3.5 to create a textual GenAI-embedded design workflow. Next, we test the workflow with postgraduate design students to speculate future scenarios, substantiate scenarios with existing circumstantial evidence, and develop a concept design based on the scenario. The outcomes highlight that GenAI suggested social changes from a range of disciplines, and designers still need to search for the source to clarify and evidence the changes manually. The case study concludes by describing the benefits of using textual GenAI in design workflows, and future research needed to strengthen the use of textual GenAI as a tool for design concept development.

Keywords. Future scenario, Futures thinking, Horizon Scanning, Signal, Futures Wheel, Generative AI, ChatGPT, Concept design

1. Introduction

Speculating scenarios for architecture design helps to deliver a vision of the future that responds to emerging trends, technological advancements, and environmental challenges. These speculations also can suggest how architecture are crucial catalysts of those visions. Not only can we find examples of these from the past, such as Peter Cook's Plug-in City (1965) and Superstudio's The Continuous Movement (1970), but similar speculative architectural projects still emerge today, such as Liam Young's Planet City (2021) and Lateral Office's State of Disassembly (2017). These examples, while diverse, share a common theme: the provocative proposals offer novel spatial experiences and, more importantly, spark discourses about the relationship between societal and architectural development.

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 39-48. © 2024 and published by the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA), Hong Kong.

Design researchers have also integrated future thinking and scenarios as briefs and inspirations to guide their design processes. These mixed approaches include Critical Design (Dunne, 2008), Speculative Design (Dunne & Raby, 2013; Mitrović et al., 2021), and Design Fiction (Bleecker, 2009; Bleecker et al., 2023). In these design approaches, the goal is not to produce practical, implementable solutions but to provoke thought, spark conversations, and inspire new perspectives on the future. These methods encourage designers to step beyond the constraints of the present and consider the long-term consequences of design choices in an environment of societal change. By engaging with speculative scenarios, designers can offer unique insights into the potential impacts of emerging technologies, social changes, and cultural shifts, ultimately contributing to a more informed and imaginative approach to design.

This case study builds on such approaches by introducing architectural design to a design workflow using textual Generative Artificial Intelligence (GenAI) to generate future scenarios that creatively provoke architectural and design discourse. The significance of this work lies in demonstrating how textual GenAI augments the ideation phase of the architectural design process and how it helps architectural designers consider future thinking in their own approach.

2. Background

2.1. FUTURE SCENARIOS

Future scenarios offer individuals insights into potential future directions and opportunities. These described visions of the future help us expand our understanding of what might lie ahead. To speculate future scenarios, designers often gather data on current trends in technology, society, and culture and extrapolate how these trends might evolve. This is different to simply imagining future environments, which does not always include examining existing societal behaviours and trends. For example, imagining a floating city, without considering if current societal demands or behaviours indicate a development towards living on water. Below we expand on two of its components, Futures Thinking and Scenarios.

Futures Thinking is a broad concept that emerged from strategic planning, foresight, and futurism. This approach provides designers a structured way to think about the future (Evans & Sommerville, 2007; Jonas, 2001; Raymond, 2003). It encompasses future-oriented techniques that not only help individuals build practical foresight skills but at its conceptual level, seeks to challenge our perceived notions of the future. As Inayatullah (2008, p. 6) pointed out, "it is not so much predicting [the future] correctly... but about enhancing our confidence to create futures that we desire. Futures methods thus decolonize the world we think we may want."

Scenarios, in the field of future studies, are narrative representations of a possible future development, spotlighting key moments that do not exist currently (Meinert, 2014). Speculative scenarios, particularly those in film and science fiction, "represent highly implausible and impractical situations and technologies… with a sheen of plausibility" (Kirby, 2010, p. 46). Examples include Peter Cook’s Plug-in City (1965) and Liam Young’s Planet City (2021). On the other hand, future scenarios offer a more grounded perspective on the possibilities ahead. According to Sardesai et al., (2021, p.

62), such future scenarios "portray causal relationships that explain how, from the vantage point of the present, a particular future in a certain story setting has been derived." By grounding future scenarios with existing observations, it enables designers to explore the potential consequences of current societal actions and trajectories of its collective endeavour.

2.2. GENAI TECHNOLOGY IN ARCHITECTURAL DESIGN

Architectural designers and researchers have begun to integrate GenAI into their architectural design development, to generate design concepts (As et al., 2018), develop architectural forms (del Campo et al., 2019), develop spatial compositions (Dzieduszyński, 2022), introduce glitches and defamiliarisation to spark new approaches to design (del Campo & Manninger, 2022), and use food imagery to inspire architectural designs (Koh, 2023), to name a few. Many architectural design projects that use GenAI, such as these examples, often use visual GenAI. These image-based GenAI, like MidJourney and Stable Diffusion, have received significant attention in the design industry, and have the potential to reshape how designers conceptualise spatial designs. These text-to-image GenAI tools often rely on deep learning models to almost generate realistic images instantly from a few keyword descriptions, accelerating the designer's ability to visualise their ideas.

3. Research Objective

The novelty of our research is integrating textual GenAI, as opposed to visual GenAI, into the design process. While AI-powered text expanders have shown their potential in various domains, like creative writing and content generation, they have not received equal attention in the architectural community as a design tool. This case study documents our exploration of future scenario development using textual GenAI to spark discussions about future environments through the three objectives:

- Explore how textual GenAI can accelerate future design scenario development. We use GenAI to digitise future design techniques into a digital workflow.
- Evaluate the impact of using our GenAI-augmented digital workflow in generating design narratives that convey the future user experience as architectural proposals.
- Explore architecture-related future scenarios of using our GenAI-augmented digital workflow.

3.1. RESEARCH APPROACH

In some forms of architectural design, the design process is seen as a dynamic and interconnected series of tasks (e.g. Generative Architecture). This design process references Alexander's (1968, p. 605) *Systems Generating Systems* perspective, which posits, "a generating system… is a kit of parts, with rules about the way these parts may be combined." We use his "kit-of-parts" perspective to develop our design workflow, then test it with designers to see its effectiveness in their concept design process. While Tan (2024) has implemented a textual GenAI workflow to help designers extract installation concepts from historical stories, our GenAI workflow focuses on value-adding future thinking approaches to architectural design process.

To develop our digital workflow, we connect different priming and prompting tasks in GenAI to form a system that serves as the framework for generating future narrative (we describe this development in the section below). In Grasshopper, a popular generative design modelling software, components are linked up into a Grasshopper definition that can produce a 3D model based on a few inputs. In our research, priming and prompting are like Grasshopper components, and our workflow is the Grasshopper definition that enables designers to come up with a future scenario based on a set of inputs. The value to designers in adopting design workflows is for them to assess their overall creative design process and pinpoint aspects where they might substitute digital technologies to augment their creativity (see e.g. Yousif & Vermisso, 2022). Here, the prospect of elevating their creativity is using our digitised future thinking process.

To implement our workflow, we rely on Yang et al.'s (2018) two approaches to teach students about designing with AI. The first way is to teach them the technical details of how the AI technology works so they understand the capabilities prior to design. The second way is to let students learn-by-doing. Our research combined both approaches; we first demonstrated the workflow by generating an example scenario to postgraduate design students, then tasked them to experiment, evaluate and iterate the workflow. These designers are not future designers, and most are novices at using GenAI. The research lasted four weeks. In these four weeks, we instructed students to experiment with the workflow, identify issues, and suggest improvements through their revised workflow. We then evaluated our digital workflow with theirs to provide a synthesised textual GenAI model for designers. After demonstrating their refined digital workflow and early scenario, the postgraduate design students spent six weeks to refine and visualise their scenarios.

4. Results

4.1. THE GENAI DIGITAL WORKFLOW

In this phase, we used textual GenAI to digitise and accelerate future design techniques. Specifically, we used ChatGPT-3.5 and digitised 1) Horizon Scanning for Signals, and 2) Futures Wheel of Consequences. We then stacked these digitised techniques to form a GenAI-embedded design workflow for scenario development (see Figure 1), similar to linking Grasshopper components to create a Grasshopper definition for designers to generate a design.

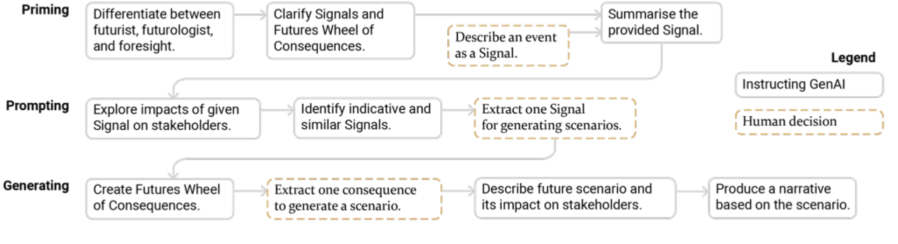


Figure 1 Textual GenAI digital workflow for developing future scenario

4.1.1. Future design techniques

We describe only Horizon Scanning for Signals and Futures Wheel of Consequences in the array of futures thinking techniques as they were chosen for our digital workflow. Reviewing all the future design techniques is outside the scope of this paper. Additionally, Evans & Sommerville (2007) summarised and contextualised a list of futures thinking techniques for design education, whereas Palmer & Ward (2010) demonstrated futures thinking integration with architectural pedagogy.

Horizon Scanning is searching for information from various sources that may indicate changes in the environment (Cuhls, 2020). This information is better known as "signals" and the collection of changes can suggest trends, which may help individuals anticipate potential futures (Rossel, 2012; Saul, 2006; Voros, 2003). We postulated that ChatGPT-3.5, pre-trained on a large corpus of internet sources, may accelerate Horizon Scanning by rapidly providing a range of signals from different disciplines to help designers indicate a changing trend. Futures Wheel of Consequences organises thinking and questioning about the future (Glenn, 2009) through a structured mind mapping process. We employed ChatGPT-3.5's language pattern recognition capabilities to analyse our Signals input (from Horizon Scanning), imagine different possibilities, and describe them as unique scenarios.

4.1.2. Generative AI ChatGPT-3.5

Textual GenAI, or Conversational GenAI, are large language models (LLMs) that leverage advanced natural language processing techniques and machine learning algorithms to simulate human-like conversations. It interprets and generates text-based responses to user's input, mimicking interactive dialogues with the user on a wide range of topics and tasks. Common LLMs include OpenAI's ChatGPT, Microsoft's Bing Chat (powered by ChatGPT), Google's Gemini, Anthropic's Claude 2, and Meta's Llama 2. We chose ChatGPT as it has the highest accuracy and lowest hallucination rate compared to other LLMs (Hughes et al., 2023/2023) based on Vectara's Hallucination Evaluation Model (Hughes, 2023). We chose ChatGPT-3.5 over ChatGPT-4 as it was free and available to public and thus more accessible to designers.

Since ChatGPT-3.5 recognises language patterns and relies on contextual information to respond, we primed ChatGPT-3.5 to enhance the accuracy and relevance of its responses. Additionally, priming enabled us to verify that ChatGPT-3.5 is outputting contextually relevant information. We primed ChatGPT-3.5 by asking it to differentiate similar keywords prior to prompting with those keywords. In our workflow, we asked ChatGPT-3.5 to differentiate between futurist, futurologist, and foresight practitioner, then prompted ChatGPT-3.5 to provide responses like a futurist to start our digital workflow.

4.2. GENAI DIGITAL WORKFLOW EVALUATION

Figure 2 shows samples of the students' workflows at the end of the four-week hands-on experimentation, whereas Figure 3 shows our revised workflow.

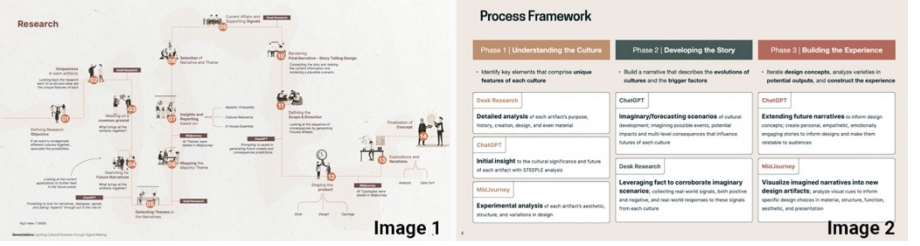


Figure 2 Iterated digital workflows (Mohanasunder et al., 2023; Yadav et al., 2023)

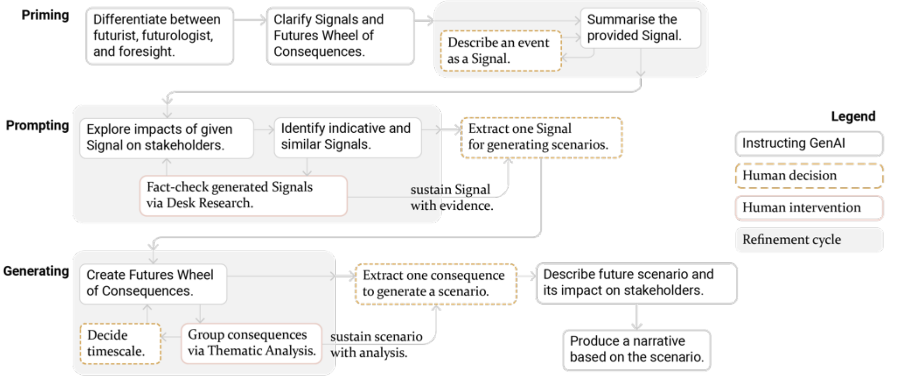


Figure 3 Revised digital workflow

In our revised workflow, we implemented three refinement cycles. This is for designers to verify the outputs provided by ChatGPT-3.5 and sharpen their inputs to produce more nuanced outcomes. During priming, we noticed that ChatGPT-3.5 did not always summarise the given Signal accurately and so it is important to refine the Signal before proceeding on to prompting. While ChatGPT-3.5 provided suitable responses for developing scenarios, the dialogue between ChatGPT-3.5 and the users played a significant role in steering the concept direction. During Horizon Scanning for Signals (i.e. the prompting episodes), ChatGPT-3.5 suggested likely but not specific evidence of Signals, and unsurprisingly, manual fact-checking on the internet to verify its evidence was required. However, this web-based evidence enabled the designers to justify their selection of a Signal for scenario development. Additionally, the value of using ChatGPT-3.5 was its ability to identify potential Signals from disciplines outside design, which increased the breath of where the Signals were found. The final refinement cycle was for Futures Wheel of Consequences. Adding a timescale to the scaffold the prompting influenced how divergent the resulting Futures Wheel of Consequences were from the original Signal. The shorter the timescale, the smaller the deviation the consequences were from the Signal. Below, we chose two generated scenarios, based on maximum variability within the range of design outcomes, to expand on how the scenarios developed into an architectural and a place-making design outcome.

4.3. SCENARIO OUTCOMES FOR CONCEPT DESIGN

Figure 4 shows two different scenarios developed by our workflow. Image 1 shows a scenario, 40 years into the future, about the consequences of using AI-augmented design algorithms on vernacular architecture in Riyadh, Saudi Arabia. Image 2 shows a scenario, 20 years into the future, about protecting culture through public artefacts that generate energy in India. We describe both scenarios below.

Figure 4 Scenario developments (Mohanasunder et al., 2023; Yadav et al., 2023)

In Scenario 1 (i.e. Figure 4 Image 1), the Signal begins with Saudi Arabia strengthening their ties with India (in 2023). The speculation is that some Indians re-locate to Saudi Arabia by 2025, cultural integration gradually occurs by 2030, and so did cultural misrepresentations by 2035. Both demographics attempt to design architecture that represents both cultures by 2050. GenAI's datasets become tainted by 2055, due to existing cultural misrepresentations, and thus future architecture in Riyadh, Saudi Arabia diverges to either be culturally accurate or misrepresented by 2060. Figure 5 Image 1 shows a comparison of the culturally accurate and inaccurate design elements, Image 2 shows how the culturally accurate design appears in architecture, and Image 3 shows how culture in architectural design may be misrepresented.

Figure 5 Scenario 1 as concept design development (Yadav et al., 2023)

In Scenario 2 (i.e. Figure 4 Image 2), the Signal begins with local communities in Native America looking to generate sustainable energy in 2023. The future thinking is that these communities will implement energy creation devices such as solar panels, wind turbines, and piezoelectric floor panels by 2026. Their ability to generate electricity while simultaneously integrating cultural design inspired other nations like India to adopt their technology by 2030, which then began to integrate their cultural stories for learning in the piezoelectric floor tiles by 2035. This becomes an initiative for public cultural learning by 2040. Figure 6 shows how this 2040 scenario may look (Image 1), how it may be used by the community (Image 2) and a prototype of the piezoelectric floor tile (Image 3).

Figure 6 Scenario 2 as concept design development (Mohanasunder et al., 2023)

5. Discussion

This research uses textual GenAI to develop and test a digital workflow for designers to create future scenarios for concept development. Scenarios like Peter Cook's Plug-in City (1965) and Liam Young's Planet City (2021) are visually provocative. While our GenAI workflow guides designers to produce concepts that are less evocative than the speculative examples we gave, it differs from mere speculations for, based on Kirby's (2010) description, our outcomes represent plausible and practical situations and technologies to an extent. Revisiting both concepts described above, the technologies and design appear attainable within their proposed timeframes. This can be attributed to how we intentionally integrated futures thinking approaches into our textual GenAI digital workflow. This is evidenced through the scenarios' timeframes, which sought to visualise imagined relationships that lead from the present to their future. Apart from introducing futures thinking to architectural concept development, our case study also contributes to the increasing works on GenAI in architectural design. The novelty of our study is using textual GenAI and not visual GenAI to develop concepts for the built environment.

A key limitation of this study is the lack of a distinct architectural concept corpora to train ChatGPT-3.5. While ChatGPT-3.5's internet-derived training data provided the breadth for suggesting Signals in other fields during Horizon Scanning, being able to restrict ChatGPT-3.5's training data to architectural concept corpora for the Futures Wheel stage may yield scenarios that are architectural-specific. Despite this limitation, our research demonstrated the potential of using textual GenAI in architectural concept development, opening up new avenues for innovative design approaches.

To further understand the value of this introduced concept technique, future research can conduct a comparative study between the difference between manual concept design development and this digital workflow. Another future research would be to integrate visual GenAI within our proposed workflow. Augmenting each generated scenario with a GenAI image may influence which scenario designers choose for development. This image creation approach may also influence designers to select more speculative scenarios than future scenarios.

6. Conclusion

GenAI can potentially shape how architects develop concepts, and our case study shows how it can help designers develop future concepts of the built environment. While many recent research focus on the image-based GenAI in exploring architectural concepts, we instead demonstrated how textual GenAI can help designers use future thinking techniques to imagine future built environments. Our case study shows how we replicated futures thinking techniques in textual GenAI to create a digital workflow. Then, we tested the workflow with postgraduate design students to generate future scenarios of the built environment.

References

Alexander, C. (1968). Systems Generating Systems. *Architectural Design*, *38*(Dec 1968), 605–610.

As, I., Pal, S., & Basu, P. (2018). Artificial intelligence in architecture: Generating conceptual design via deep learning. *International Journal of Architectural Computing*, *16*(4), 306–327. https://doi.org/10.1177/1478077118800982

Bleecker, J. (2009). *Design Fiction: A Short Essay on Design, Science, Fact, and Fiction* (1st ed.). Near Future Laboratory. https://doi.org/10.1002/9781119815075.ch47

Bleecker, J., Foster, N., Girardin, F., & Nova, N. (2023). *The Manual of Design Fiction*. Near Future Laboratory. https://nearfuturelaboratory.com/the-manual-of-design-fiction

Cook, P. (1965). *Plug-in City* [Portfolio]. Archigram. https://www.archigram.net/portfolio.html

Cuhls, K. E. (2020). Horizon Scanning in Foresight – Why Horizon Scanning is only a part of the game. *Futures & Foresight Science*, *2*(1), e23. https://doi.org/10.1002/ffo2.23

del Campo, M., & Manninger, S. (2022). Strange, But Familiar Enough: The Design Ecology of Neural Architecture. *Architectural Design*, *92*(3), 38–45. https://doi.org/10.1002/ad.2811

del Campo, M., Manninger, S., Sanche, M., & Wang, L. (2019). *The Church of AI - An examination of architecture in a posthuman design ecology*. 767–772. https://doi.org/10.52842/conf.caadria.2019.2.767

Dunne, A. (2008). *Hertzian Tales: Electronic Products, Aesthetic Experience, and Critical Design*. The MIT Press. https://mitpress.mit.edu/9780262541992/hertzian-tales/

Dunne, A., & Raby, F. (2013). *Speculative everything: Design, fiction, and social dreaming*. MIT Press.

Dzieduszyński, T. (2022). Machine learning and complex compositional principles in architecture. *International Journal of Architectural Computing*, *20*(2), 196–215. https://doi.org/10.1177/14780771211066877

Evans, M., & Sommerville, S. (2007, July 9). Design Led Futures: Futures Thinking in the Design Curriculum. *2007 International Conference on Design Education*. Connected 2007, Sydney, Australia.

Glenn, J. C. (2009). Chapter 6: The Futures Wheel. In J. C. Glenn & T. J. Gordon (Eds.), *Futures Research Methodology—Version 3.0*. The Millenium Project: Global Futures Studies & Research. https://www.millennium-project.org/publications-2/futures-research-methodology-version-3-0/

Hughes, S. (2023, November 6). Cut the Bull…. Detecting Hallucinations in Large Language Models. *Vectara*. https://vectara.com/cut-the-bull

Hughes, S., Mendelevitch, O., Bae, M., Connelly, S., & Ashimine, I. E. (2023). *Hallucination Leaderboard* [Computer software]. Vectara. https://github.com/vectara/hallucination-leaderboard (Original work published 2023)

Inayatullah, S. (2008). Six pillars: Futures thinking for transforming. *Foresight*, *10*(1), 4–21. https://doi.org/10.1108/14636680810855991

Jonas, W. (2001). A Scenario for Design. *Design Issues*, *17*(2), 64–80. https://www.jstor.org/stable/1511876

Kirby, D. (2010). The Future is Now: Diegetic Prototypes and the Role of Popular Films in Generating Real-world Technological Development. *Social Studies of Science*, *40*(1), 41–70. https://doi.org/10.1177/0306312709338325

Koh, I. (2023). AI-Bewitched architecture of Hansel and Gretel. *Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA)*, *1*, 9–18.

Lateral Office. (2017, September 2). *State of Disasembly: Speculations on Architecture as part of the E-Waste Cycle* [Portfolio]. Lateral Office. https://lateraloffice.com/states-of-disassembly-2017

Meinert, S. (2014). *Field manual Scenario building*. European Trade Union Institute. https://www.etui.org/publications/guides/field-manual-scenario-building

Mitrović, I., Auger, J., Hanna, J., & Helgason, I. (Eds.). (2021). *Beyond Speculative Design: Past - Present - Future*. SpeculativeEdu.

Mohanasunder, P., Prakash, H., Chachria, S., & Chen, Y. (2023). *Energizing Change: Preserving Tradition and Culture Within Infrastructure and Energy Solutions*. Issuu. https://issuu.com/phurnneemohan/docs/arch1417_as3_energizingchange

Palmer, J. S., & Ward, S. P. (2010). Futures Thinking: Towards Innovation and Sustainability in Architecture. *Proceedings of ConnectED 2010 2nd International Conference on Design Education*. 2nd International Conference on Design Education, Sydney, Australia.

Raymond, M. (2003). *The Tomorrow People: Future Consumers and How to Read Them*. Pearson Education.

Rossel, P. (2012). Early detection, warnings, weak signals and seeds of change: A turbulent domain of futures studies. *Futures*, *44*(3), 229–239. https://doi.org/10.1016/j.futures.2011.10.005

Sardesai, S., Stute, M., Fornasiero, R., Kalaitzi, D., Barros, A. C., Multu, C., & Muerza, V. (2021). Future Scenario Settings for Supply Chains. In R. Fornasiero, S. Sardesai, A. C. Barros, & A. Matopoulos (Eds.), *Next Generation Supply Chains* (pp. 61–78). Springer International Publishing. https://doi.org/10.1007/978-3-030-63505-3_3

Saul, P. (2006). Seeing the Future in Weak Signals. *Journal of Futures Studies*, *10*(3), 93–102.

Superstudio. (1970, May 7). The Continuous Movement. *Cristiano Toraldo di Francia*. https://www.cristianotoraldodifrancia.it/continuous-monument/

Tan, L. (2024). Using textual GenAI (ChatGPT) to extract design concepts from stories. *LearnxDesign Conference 2023*. LearnxDesign 2023: Futures of Design Education, online.

Voros, J. (2003). A generic foresight process framework. *Foresight*, *5*(3), 10–21. https://doi.org/10.1108/14636680310698379

Yadav, T., Rajamani, V., & Althubaiti, N. (2023). *The Unified Deception: Cultural Fusion and AI Tales*. Issuu. https://issuu.com/trishlayadav/docs/assignment_02_concept_architectural_group_noha_tri

Yang, Q., Scuito, A., Zimmerman, J., Forlizzi, J., & Steinfeld, A. (2018). Investigating How Experienced UX Designers Effectively Work with Machine Learning. *Proceedings of the 2018 Designing Interactive Systems Conference*, 585–596. https://doi.org/10.1145/3196709.3196730

Young, L. (2021). *Planet City* [Portfolio]. https://liamyoung.org/projects/planet-city

Yousif, S., & Vermisso, E. (2022). *Towards AI-Assisted Design Workflows for an Expanded Design Space*. 335–344. https://doi.org/10.52842/conf.caadria.2022.2.335

ADAPTING THE SOFTWARE DESIGN PATTERN MODEL FOR AI-ENABLED DESIGN COMPUTING

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. Exponential AI development requires an adaptation to new technology by traditionally reluctant architects and allied practitioners. This paper examines the potential of the software design pattern (SDP) model, used in software engineering to capture and reapply designs, as one underpinning. Patterns have creativity and pedagogical benefits in parametric modelling, yet consideration of AI and broader design computing as well as the derivation and versatility implied by an SDP model are underexamined. This research questions how, in an AI context, new patterns may evolve for varied AI levels and non-geometrical features. It is undertaken in the Unity game engine with critical application of two prominent extant patterns as a computational workflow design response to a real-world citizen engagement scenario. A novel, feature-agnostic pattern is derived with a simple AI model and is verified for other AI models. The work concludes design computing patterns can abstract existing pattern knowledge to flexibly evolve and apply across rapidly changing AI-enabled design computing contexts and thereby assist practitioners to positively respond to AI advances.

Keywords. Artificial Intelligence, Computational Design, Software Design Patterns, Architectural Practice, Unity 3D, Intelligent Agents.

1. Introduction

Computer scientist and entrepreneur Andrew Ng posits that “AI is the new electricity” and is developing exponentially (Ng, 2017). Yet as architect and Yale University emeritus professor Phillip Bernstein notes (2022, p. 131), architects typically maintain a sceptical resistance to the adoption of “new, disruptive technologies”, and an enduring lag exists between AI development and its integration in architectural practice (Kimm & Burry, 2021). AI is continually evolving within a rapidly changing techno-savvy society and is in a persistent state of novelty (ibid.); since Ng spoke in 2017, broad AI models, like Midjourney or ChatGPT, and emerging industry software solutions in domains from early site analysis (see Sidewalk Labs' Delve) to generative design (see Evolve Lab’s VERAS), have demonstrated potential to supplant architects’ and urban designers’ (“practitioners”) roles in workflow aspects. Yet full

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 49-58.

consideration of diverse project criteria requires the creative analysis and synthesis of the practitioner. These two contradictions indicate a need for support to practitioners when creating their own AI-enabled workflows. This paper examines the model of Software Design Patterns (SDPs) and assesses how it may be adapted to aid practitioners – typically inexpert programmers – in their own construction of AI-enabled design computing workflows and tools. "Design computing" is here used in the broad sense of Gero (1998) as use of "software with the relevant features and utilities to support some aspects of design activity."

2. Background and Context

The formal concept of a pattern as a reusable solution to a common problem, given independent of a specific context, was popularised with the publication of *A Pattern Language: Towns, Buildings, Construction* (Alexander et al., 1977). These Alexander patterns (APs) were adopted by the programming community, including in the seminal book *Design Patterns: Elements of Reusable Object-Oriented Software* which catalogued 23 SDPs (Gamma et al., 1995). As illustration, the *singleton* SDP ensures only one instance of a class may exist, thereby avoiding problems that can arise from concurrent access of resources such as database connections. Patterns help developers to leverage insights of past designers to avoid common pitfalls or reinventing the wheel, and to adopt a higher-level, abstract perspective to remain "in the design" for longer rather than prematurely contesting with the intricate specifics of implementing code (Freeman et al., 2020, Chapter 1; Shalloway & Trott, 2004, Chapter 5).

A systematic literature search showed limitations in use of pattern concepts in design computing. Pattern models are typically explicitly rooted in APs oriented by subsequent pattern-based approaches, and SDP consideration, if any, is narrowly based on the work of Gamma et al. As *gap 1*, although some publications referenced AI tangentially – for example Globa or Steenson (Globa, 2015, p. 61; Steenson, 2017, p. 76) – none focused on the application of patterns to the use of AI in design computing.

Foremost in the literature are the 13 Woodbury patterns (WPs) that help designers create geometry through interacting with the particulars of the parametric design software package GenerativeComponents and, more generally, as Woodbury (2010, p. 186) writes in his book *Elements of Parametric Design*, "help designers learn and use propagation-based parametric modeling systems". In the *controller pattern* (CP), for instance, a full geometric model may be controlled by a simpler linked model so that a user may interact with a model in a "clear and simple way" (ibid. p.191).

Qian (2009, p. 191) demonstrated that WPs support self-directed and formal learning, script development and the potential to introduce novel solutions, and collaborative design and practitioner communication. Globa (2015) affirms and extends the findings of Qian, although proceeds from Woodbury's book and not Qian directly. Globa tested design support through reuse of knowledge by empirical comparison of WPs, representing an abstraction reuse approach, with reuse of specific scripting solutions. The outcomes of workshop-based studies, in which participants used Grasshopper, showed that abstract programming solutions facilitate design scripting: they aid user understanding of the essential logic and modelling processes of scripting, encourage production of complex script and model outputs, and assist translating ideas into designs and exploration of novelty. Additionally, Globa (2015, pp. 280-282) reasons these benefits would extend to other design computing domains.

Yu and Gero (2015) provided other empirical evidence of patterns by coding observed workshop design activities to the distinct stages of a design process ontology, thus inferring pattern existence though ready transitions between design process states. They conclude that if inferred patterns are extant or developed on the fly is unanalysed in their work, and urge that “if we can generalise these transitions to design patterns it would be of assistance to architects in conceptualizing their scripting process.”

The model of SDPs is heterogeneous and there are a number of gaps between it and the reviewed literature. This paper focuses on two such: how patterns arise, and their flexibility across computational environments.

Gap 2 concerns pattern derivation. An SDP is a response to a “recurring” design problem (Buschmann et al., 1996, Chapter 1; Freeman et al., 2020, Chapter 13) and should express “successful designs” (Gamma et al., 1995, p. 1). The part of experience of software designers in deriving patterns is clear: patterns come from the “collective experience of skilled software engineers” and “prior experience” of software designers that is “learned by designers and users over the years” (Buschmann et al., 1996, Chapter 1; Gamma et al., 1995, p. 1; Shalloway & Trott, 2004, pt. 1).

Qian demonstrated pattern derivation through forming a large corpus of design examples in participant observer workshops followed by interview coding and other analysis. The process, although sound, cannot serve to identify pattern use in rapidly evolving contexts or in the informal settings of practice. The method of Yu and Gero could in principle identify pattern existence in even a singular workflow, but does not define patterns. Woodbury (2010, p. 189) asserts the importance of the experience of the designer: “To write a pattern is to listen to yourself and your colleagues.” In the works surveyed, the particulars of how an individual designer may derive a pattern are left unstated. As Qian (2009, p. 191) concluded, the question of “how a pattern… can be polished and enriched by designers' active involvement and communication” remains open. The “pioneering spirit” of Pemberton and Griffith (1998), who derived patterns for collaborative workspaces and technology, offers a possible response – a pattern once found could prompt the identification of another pattern per new needs.

Gap 3 concerns pattern versatility. An SDP typically has enduring use across diverse computational environments; the singleton SDP, for instance, is as applicable to C++, released in 1985, as it is to Rust, released in 2015. In contrast, idiomatic SDPs are particular to a given language (Buschmann et al., 1996, sec. 1.3). An idiom useful for one programming paradigm is not necessarily applicable to another paradigm.

The patterns models surveyed tend towards idiomacy. Globa extended WPs from GenerativeComponents to Grasshopper but they are still within the domain of propagation-based parametric modelling systems. Furthermore, focus is on geometrical primitives rather than on manipulation of more general built environment features such as structural members, circulation paths, houses, or styles.

3. Research Questions and Method

The lead author designed and conducted a research-through-design (RtD) investigation to test SDP potential at those three gaps. RtD is concrete research through the action of the practitioner to create an experimental object within and regulated by a specific practice context (Herriott, 2019). The object is a means to uncover and convey knowledge, not the end goal. Two complementary research questions on pattern derivation and versatility in an AI-attuned design computing practice context were considered. On derivation, what potential exists to evolve new patterns from an existing pattern? On versatility, may such patterns be applicable to different types of AI and features other than geometry alone?

An extensible model of AI is essential to address the aspect of practitioner use of diverse AI. This research uses an adaptation to design of a computer science model of intelligent agents (IAs) that perceive, decide, and act (Kimm, 2022). Beyond the agent scale, the model considers AI from a teleological perspective of the designer providing value to the client rather than analysing the underlying AI technical mechanics directly. It offers a framework of increasing AI sophistication that, rather than imposing an arbitrary, inflexible partitioning of AI and not-AI, sets out a continuum from AI that reflexively responds to stimuli to AI with a complex internal world model.

Initial patterns are taken from WPs for their well-described, prominent, 13-strong pattern catalogue of demonstrated practitioner benefit. Two are selected in the next section for their model clarification and abstraction ability.

The practice context which enables consideration of features other than geometry is an engagement with Aurecon – a global design, engineering, and advisory company – and its online community engagement software tool commissioned by a government client which sought citizen stakeholder preferences for the design of a harbourside park in central Sydney, Australia (https://harbourpark.sydney/). Website users may select and position on a bare site three to five 2D park feature graphics from a library of fourteen cultural, health/safety, landscaping, pedestrian, or recreational features (Figure 1). An aim, per the tool website, was to permit a user to create a design that "will contribute valuable information that will help us determine how the park will look and feel based on the features and activities that matter to you." Graphics are animated on short loops – for example, a person at an exercise station is shown using it – and resize to give a sense of perspective. The graphics are otherwise invariant: the user is presented with little scope to indicate any preferences in this static framework.

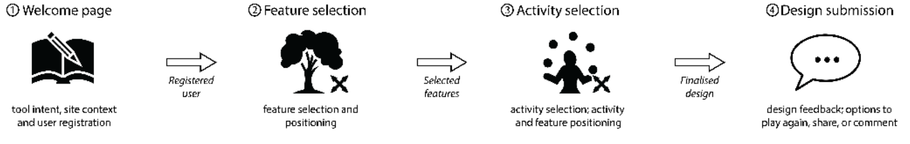


Figure 1 The workflow of the harbourside park app.

The research computational starting point is an initial study on capturing citizen street tree planting preference, only imprecisely given, and then making a definite move in a search space of appropriate trees. A user interacting with this may trace a desired tree massing profile on the screen. The prototype then analyses the path of the user's indicative 2D line, using the *$1 Unistroke Recognizer* algorithm (Wobbrock et al., 2007), and suggests one option from a curated library of 3D precedents (Figure 2).

A design goal was thus formulated for reimaging the harbourside park app (HPA):

How may the user have more latitude to convey their preferences imprecisely or incompletely, and how may they receive guidance as they develop their design?

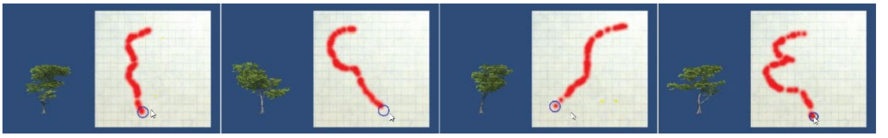


Figure 2 The base computational exploration.

The Unity game engine was selected as the research platform for its industry use, capacity for design analysis, and accessibility to non-programmers (Huang et al., 2021). It supports non-parametric modelling and, via C#, flexible exploration of AI.

Prototype evaluation was by interactive demonstration to and interview of Aurecon employees. As the research focus is the pattern exploration process, rather than high prototype utility and usability, a response to the design goal need not be unconditionally successful but must demonstrate in its seriousness that the investigation took place in a bona fide design computing practice scenario. Prototype evaluation is thus restricted.

4. Results

Two WPs were selected as pattern exploration seeds for their capacity to map user design intent expression to some point in the search space of a design computing tool, as implied by the design goal. The controller pattern (CP) was described above; in the *reactor pattern* (RP), an object responds to another object as an "interactor point" within a common modelling environment. Each simplifies the set of inputs to a model and is conceptually abstract rather than narrowly geometry oriented. Nonetheless, they, although applicable to placement of park features, for instance, do not fully address the design goal that deals with the non-concrete domain of flexible preference expression.

4.1. APPLYING THE WOODBURY CONTROLLER PATTERN

The initial computational study embodies the CP to a limited degree: the specification of a tree massing as a direct one-to-one mapping of sketched profile preference to sampled tree elevation fits the *abstracting* CP classification of Woodbury (2010, p. 192) of a "simple version of the main model that suppresses unneeded detail". The *transforming* mode of the CP, which "changes the way you interact with a [full] model" (ibid. p192), suggests a brokering model may be inserted into the full model to allow nuanced expression of the feature preferences within the wider environment.

Accordingly, the initial prototype was updated such that a user could express a tree preference – shading provided by a tree – distally by interacting with the model environment rather than proximately by direct selection. The processing of the library of trees was redeveloped so that it may be searched by shading profile rather than elevation profile as in the initial prototype. In the iterated tree workflow (Figure 3), a user clicking on a grass canvas places a new tree and can draw the desired shading profile. The system displays the best-match 3D tree and the profile stored within the $1 gesture recogniser. Any instantiated tree can be selected and repositioned by the user.

The resulting prototype has limitations as a simplified model of environment

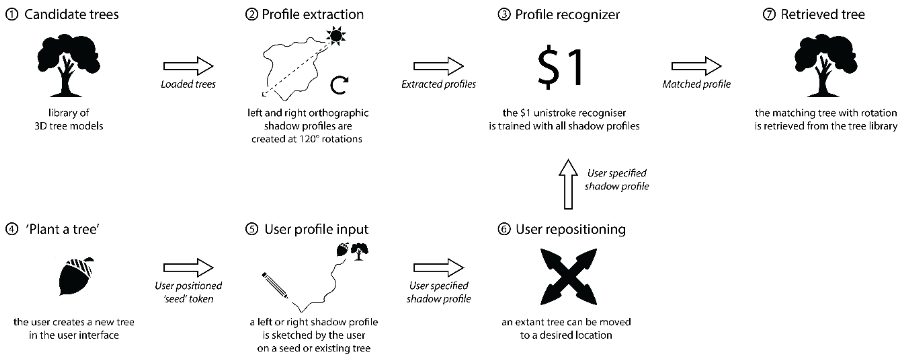


Figure 3 The workflow of the prototype following consideration of the controller pattern.

conditions and user interaction. First, shadow sketching is supported only for drawing on flat ground. Second, due to preprocessing, only one time of day can be supported. Third, the sketching method permits a user to specify one tree feature preference with a greater dimensionality than can be captured readily in a list of choices, yet the treatment within the prototype, once that sketch is analysed by the gesture recogniser, is precise – the prototype never provides options where ambiguity of choice may exist.

4.2. APPLYING THE WOODBURY REACTOR PATTERN

The HPA provides guidance only on submission of a design and a user cannot change any aspect of a feature to a preferred state by indirect or direct interaction. To test the RP for this, a matrix was developed to analyse the potential for RP application if taking features, rather than geometry, as the material of that pattern. The matrix sampled tree, light, and path features from the HPA to model how a feature may respond to the proximity of a RP interactor point which, rather than being an abstract position in space with no other associated quality, is itself a feature in this treatment. The RP, rather than operating on geometry that responds in its position, direction, or scale, here provides a schema for features to dynamically moderate their design role.

Matrix feature response analysis reveals a deficiency of direct use of the RP. If treating the RP interactor point as a feature, rather than as an abstract position in space, the interactor point must itself change. As example from the matrix, a path feature as an interactor correctly induces in nearby trees a change to low branchfall risk. Conversely, a tree moving as an interactor near to a path causes a path change, such as to preserve a clear root buffer zone, yet is itself not changed: it takes on no attribute of low branchfall risk, and thus a risk to public safety is not improved. Branchfall in this example is a proxy for any interactor feature attribute that requires moderation.

Therefore, an adaptation of the RP must be reciprocal: any feature engaged in the model should operate both as an interactor and a reactive result. Counterintuitively, given the criticism of one-sided interaction directly above, this symmetry indicates that exploration of reactive response in the reactor pattern might, for a suitable subset of features, be realised in the developing prototype through a reactive response in either of the moving interactor or the stationary features. Moreover, a focus on the reactive response of a moving feature facilitates consideration of features that are static for reason of being preexisting site conditions (e.g. overshadowing) or of being fixed in an earlier stage of the design process (e.g. primary pedestrian circulations).

The workflow was updated so that a tree species is selected for a site location by its appropriateness (Figure 4). Each tree library tree is thus tagged with its characteristics. In this test, the attribute set is limited to shade tolerance, soil condition, and branchfall risk. These three measures represent the many possible factors and potential trade-offs that could be incorporated, including considerations of hydrology, heat island effect, or bird life. The notional tags do not reflect the actual species requirements; however, tag selection was curated to create a crude congruence between the tagged requirements and the appearance of the tree models. The Unity implementation is seen in Figure 5.

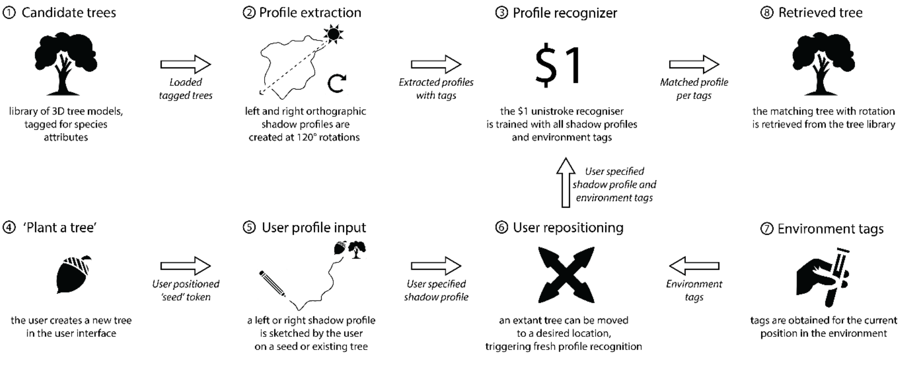


Figure 4 The workflow of the prototype following consideration of the reactor pattern.

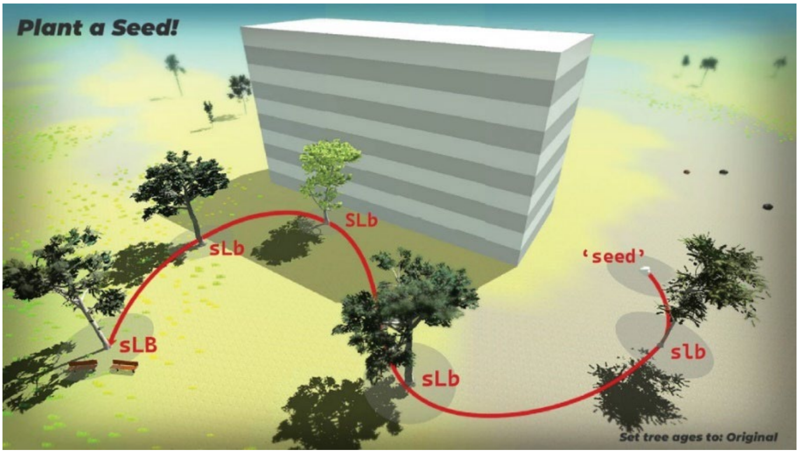


Figure 5 The progress of one tree in the final prototype using the workflow of Figure 4.

Despite limitations noted previously, workshop feedback from two Aurecon leaders in digital innovation and engagement indicated the outcome of consideration of the CP and RP in the prototype design process met the design goal and could usefully extend the HPA workflow, particularly for diverse government-based projects and establishing social licence. Nonetheless, separate informed comment by an academic and landscape architect formally interviewed on the CP stage results warned against uncritical use: “design is much more comprehensive than just placement of objects.”

4.3. A SYNTHETIC NOVEL PATTERN

Critical application of the encapsulated knowledge of two WPs to the design problem evolved a novel method. That solution is here expressed as a pattern using the four prime descriptors urged by Woodbury: “Name, What, When and How” (2010, p. 189).

Name: Polymorphic Feature (PFP).

What: Allow a feature to change its own nature according to its context.

When: Useful if feature solution space position information is initially incomplete.

How: Model a feature with an IA interface. Develop one IA to update model attributes per dynamic environment conditions, bounding the solution space search by domain knowledge and any original user intent. Implement new IA models as required.

4.4. CROSS-DOMAIN APPLICATION OF THE NOVEL PATTERN

To test the PFP and its capacity to harness diverse AI, the simple, reflexively responsive IA of the prototype was replaced with an IA of a more complex internal world model. This, as a proof of concept, was carried out manually using commercial large language model and text-to-image model services. For each new tree position, the prototype output a ChatGPT prompt customised with context specific tags (in italics below):

> Determine a commonly used tree species for a *sunny* city park location with *loam* soil in Sydney, Australia. You must consider factors of urban planning, botany, landscape architecture, and other relevant disciplines. Then write a text-to-image (Midjourney) prompt of a *mature* example of that species. The prompt should follow this formula: “/imagine prompt: [subject], [your descriptive keywords or phrases], [background] [art style] --ar [width]:[height]”. Use a low poly style, and a plain white removable background is essential. Adjust the aspect ratio (“--ar”) as appropriate for the dimensions of your selected tree.

Figure 6 Midjourney outputs for the quoted prompt (two leftmost images), prompted juvenile examples (two centre images), and a reference blue gum photo (rightmost image).

An example generated Midjourney prompt for a recommendation of *Eucalyptus globulus* (blue gum), which also appears in the City of Sydney tree species list (swi.nu/trees), is below with output and variations in Figure 6:

> /imagine prompt: Towering Blue Gum Tree, silver-blue leaves shimmering in the sunlight, providing an iconic silhouette in the city park. The low poly style enhances its grandeur. Plain white removable background for transparency. --ar 4:5

5. Discussion and Conclusion

This investigation has shown an adaptation of the Software Design Pattern (SDP) model for the construction of AI-enabled design computing workflows and tools. Aspects of SDPs have been applied by others in design computing and in depth for the subfield of propagation-based parametric modelling, most prominently as Woodbury patterns (WPs). Those uses have clear design space exploration and pedagogical benefits but consider the SDP model as one precedent of many and, as one gap, do not specifically consider AI. Consequently, further gaps exist of which two were questioned here: the derivation and versatility of patterns in an AI context.

Adapting the SDP model, the lead author demonstrated pattern derivation utilising the intuition and experience of the designer. Critical application of two WPs, as conceptual framings for a design computing tool, guided evolution of an initial computational study. Significantly, this direction occurred though clear congruence with the design problem as well as points of divergence. Incompatibilities prompted exploration of novel reaches of the design solution space, in contrast to the pattern chaining of Pemberton and Griffith in which a pattern is linked to its ancestors by its consonance alone. The resultant polymorphic feature pattern (PFP) is not asserted to be novel in the sense of Boden's H-creativity and it does not need to be: patterns capture past experience. Nonetheless, it is a P-creative, novel addition to this design computing practice scenario. This work is a complement to that of Yu and Gero and their call to generalise to patterns the design process transitions they inferred. It is, however, a single datapoint that illustrates a generalisation mechanism; group participant studies could offer valuable empirical evidence but the challenges of unadulteratedly recording design process decisions to the degree done so here would be considerable.

The investigation exhibited versatility of patterns in respect to technology domains and to built environment considerations. The former had two aspects. In the selection of WPs, the design process exploited the encapsulated knowledge of past designers in parametric computing and applied it to the distinct Unity environment. Further, it demonstrated how a pattern may apply across an AI continuum of intelligent agents (IAs) by developing the PFP with an IA model of mere instinctual responsiveness and then showing how it can apply with contemporary, sophisticated AI models. While those outputs were not perfect – see the juvenile blue gum images of Figure 6 – they do provide a robust proof of concept. Regarding built environment considerations, this research shows how a pattern can operate not only on geometric elements, as in the precedents, but also broadly on actual built environment features. The derived novel PFP hence aids a freedom of action in expressing a feature with subsequent guidance – it facilitates exploring what-if scenarios on a micro scale.

This pattern investigation and framework developed in a real-world project context. Stakeholder workshop feedback substantiated a reasonable design computing practice scenario took place. Although the primary research focus was at a design thinking level, the pattern found has a measure of external validity despite its narrow spectrum of designer experience, insofar stakeholders favourably judged it in terms of its method and the design goal. The generative AI reuse also supports validity, as does background PFP use for the prototyping within a past study of different focus (Kimm et al., 2023).

The adaptation of the SDP model in this paper signposts a way to further, broader adoption of patterns in industry by which design computing practitioners can be supported in their professional service to clients and society. They, with such patterns and as AI continually evolves, can apply existing knowledge flexibly across computing contexts and in consideration of features other than bare geometry. This research enlivens the toolkit available to practitioners to construct their own design computing workflows. It proposes a viable, enriching direction for further research into how architects may positively respond to AI advances and sustain their creative practice.

References

Alexander, C., Ishikawa, S., & Silverstein, M. (1977). *A pattern language*. Oxford.

Bernstein, P. (2022). *Machine learning: Architecture in the age of AI*. Routledge.

Buschmann, F., Meunier, R., Rohnert, H., Sommerlad, P., & Stal, M. (1996). *Pattern-oriented software architecture. [Vol. 1], A system of patterns*. Wiley.

Freeman, E., Robson, E., Sierra, K., & Bates, B. (2020). *Head first design patterns*. O'Reilly.

Gamma, E., Helm, R., Johnson, R., & Vlissides, J. (1995). *Design patterns: Elements of reusable object-oriented software*. Pearson Deutschland GmbH.

Gero, J. S. (1998). Research in design computing: An artificial intelligence framework. *International Conference on Artificial Intelligence for Engineering*, (pp. 5-12).

Globa, A. (2015). *Supporting the use of algorithmic design in architecture: An empirical study of reuse of design knowledge*. PhD thesis, Victoria University of Wellington.

Herriott, R. (2019). What kind of research is research through design. *IASDR 2019*.

Huang, X., Kimm, G., & Burry, M. (2021). Exploiting game development environments for responsive urban design by non-programmers. *26th CAADRIA*, 2, (pp. 689-698).

Kimm, G. (2022). Classes of AI tools, techniques, and methods. In I. As, P. Basu, & P. Talwar (Eds.) *Artificial Intelligence in Urban Planning and Design* (pp. 61-83). Elsevier.

Kimm, G., & Burry, M. (2021). Steering into the Skid: Design Augmentation to Arbitrage Human and Artificial Intelligences. *40th ACADIA*, 1, (pp. 698-707).

Kimm, G., White, M., & Burry, M. (2023). Extending Visuospatial Analysis in Design Computing. 28th CAADRIA, (pp. 655-664).

Ng, A. (2017). 'AI is the new electricity'. *AI Frontiers Conference*, Santa Clara, Calif, 3 Nov. Available at: www.youtube.com/watch?v=JsGPh-HOqjY (Accessed: 3 Sept. 2023).

Pemberton, L., & Griffiths, R. N. (1998). The timeless way: Making living cooperative buildings with design patterns. *CoBuild'98*, 1, (pp. 142-15).

Qian, Z. C. (2009). *Design patterns: Augmenting design practice in parametric CAD systems*. PhD thesis, Simon Fraser University.

Shalloway, A., & Trott, J. (2004). *Design Patterns Explained: A New Perspective on Object-Oriented Design*, Second Edition (2nd edition.). Addison-Wesley Professional.

Steenson, M. W. (2017). Christopher Alexander: Patterns, Order, and Software. In *AI: How Designers and Architects Created the Digital Landscape* (pp. 21-76). MIT Press.

Wobbrock, J. O., Wilson, A. D., & Li, Y. (2007). Gestures without libraries, toolkits or training. *20th ACM Symp. on User Interface Software and Technology* (pp. 159-168).

Woodbury, R. (2010). *Elements of parametric design*. Routledge.

Yu, R., & Gero, J. (2015). An Empirical Foundation for Design Patterns in Parametric Design. *20th CAADRIA* (pp. 551-560).

AI-ENHANCED PERFORMATIVE BUILDING DESIGN OPTIMIZATION AND EXPLORATION

A design framework combining computational design optimization and generative AI

CHUWEN ZHONG[1], YI'AN SHI[2], LOK HANG CHEUNG[3] and LIKAI WANG[4]
[1,2,3,4]*Xi'an Jiaotong-Liverpool University*
[1]*chuwen.zhong20@student.xjtlu.edu.cn, 0009-0006-3483-5948*
[2]*yian.shi21@student.xjtlu.edu.cn, 0009-0003-3769-9548*
[3]*lokhang.cheung19@student.xjtlu.edu.cn, 0009-0001-2911-3733*
[4]*wang.likai@outlook.com, 0000-0003-4054-649X*

Abstract. When using computational optimization for early-stage architectural design, most optimization applications often produce abstract design geometries with minimal details and information in relation to architectural design, such as design languages and styles. Meanwhile, Generative AI (GAI), including Natural Language Processing (NLP) and Computer Vision (CV), hold great potential to assist designers in efficiently exploring architectural design references, but the generated images are often blamed for having limited relevance to the context and building performance. To address the limitation in computational optimization and leverage the capability of GAI in design exploration, this study proposes a design framework that incorporates Performative/Performance-based Design Optimization (PDO) and GAI programs for early-stage architectural design. A case study is demonstrated by designing a high-rise mixed-use residential tower in Hong Kong. The result shows that the PDO-GAI approach can help designers efficiently proceed with both diverging exploration and converging development.

Keywords. Building Performance, Computational Optimization, Design Exploration, Generative AI, Architectural Style, Façade Language.

1. Introduction

Buildings greatly impact urban environments, and performative/performance-based building design plays a critical role in achieving sustainable urban development. Recently, a number of studies have employed computational optimization to assist in such design tasks to improve environmental and energy performance of the design, such as daylighting, passive heating/cooling, and energy consumption. Computational optimization can automatically generate and evaluate a large number of design variants

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 59-68.

and search for desirable ones with competitive performance using parametric models, performance simulations, and evolutionary algorithms (Li et al., 2020). However, due to the fact that using parametric modeling techniques to generate building design requires a significant degree of design abstraction and simplification, most parametric models developed for computational optimization only generate designs with minimal details and information in relation to architectural design, such as architectural styles and facade languages (Wang, 2022a; Wang et al., 2022). This undermines its integration into architects' design ideation and exploration process.

Regarding architectural design exploration, recent advancements in Generative AI (GAI), including Natural Language Processing (NLP) and Computer Vision (CV) techniques, hold great potential to assist architects in efficiently exploring architectural references with different design styles and languages (Zhu & Luo, 2023). However, solely using GAI is blamed as it can only generate designs with limited relevance to the context and building performance. Most current relevant applications are focused on creating photorealistic rendering images. In contrast, some other recent studies also investigated collaborative applications of different GAIs as a team of design partners in the architectural design process (Cheung & Dall'Asta, 2023), highlighting the potential of GAI in enhancing and strengthening architects' ideation and exploration. This also shows the potential of such applications in more dynamic and abstract design exploration tasks beyond creating appealing architectural images.

Considering the strengths and the weaknesses of computational optimization and GAI, on the one hand, the strength of GAI in rapidly producing architectural images with various styles and languages can be leveraged to address the limitations of the lack of architectural details in conventional computational optimization in performance-based building design. On the other hand, connecting GAI with computational optimization also helps address the issue of the contextualization inherited in GAI. These two aspects reveal a research gap regarding the combination of computational optimization and GAI for early-stage architectural design exploration for both performance and architectural styles and languages.

Departing from this, this study proposes a design framework that incorporates computational optimization with GAI for early-stage architectural design. The framework includes the use of computational optimization to evolve design populations and search for optimized building massing forms ((Wang, 2022b)). Meanwhile, several multimodal AI recognition and generation tools, including Visual-Question-Answering (VQA) AI such as BLIP2, large language model such as GPT, and AI arts generation tools, such as Midjourney and Stable Diffusion, are explored to enhance the design exploration and feedback regarding architectural styles and language. Instead of using them as usual standalone applications, they were experimented with and examined iteratively to conduct design exploration back and forth with the computational optimization tool. Hence, the AI and optimization tools transform from computational tools into design collaborators.

Following this, the main body of the paper describes a design workflow established to integrate computational optimization and GAI for iterative design optimization, ideation, and reflection. At the same time, a case study is presented to demonstrate the proposed workflow and its utility in the architectural design process. The paper concludes by discussing the relevance of this study to the research discourse of computational optimization and GAI while pointing out the future research direction based on the limitations we identified in this study.

2. Method

2.1. DESIGN WORKFLOW

Figure 1 illustrates the overall working system, which mainly consists of two parts: Performance-based Design Optimization (PDO) and GAI program.

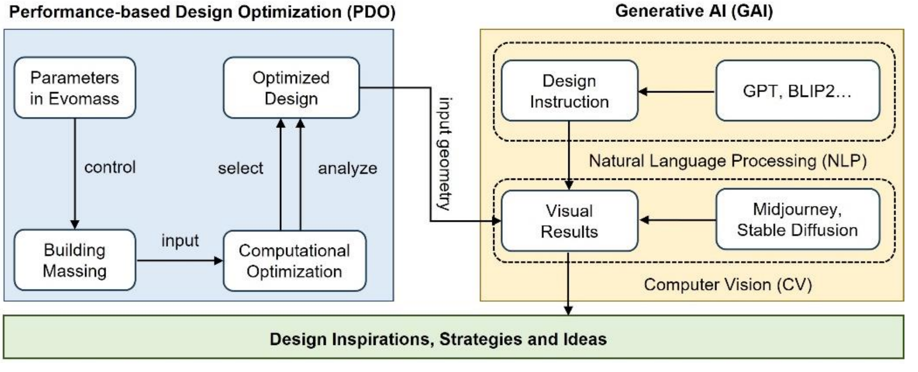


Figure 1. Overall Design Workflow

The workflow starts with the PDO, which can produce several optimized building massing designs through multiple iterations of optimization runs. Subsequently, the GAI process assists in the exploration and visualization of architectural style and design concepts according to the designers' inputs. Finally, designers extract design strategies, ideas, and inspirations related to architectural styles and languages by synthesizing the building massing form with visualized conceptual images, which can guide further design scheme development.

The PDO stage is based on using EvoMass on the Rhino-grasshopper platform (Wang et al., 2020), while EvoMass can also be replaced with other parametric models to achieve similar building massing optimization results. The GAI process involves the use of NLP and CV tools. The NLP tool includes GPT and BLIP2, while CV tools include Midjourney and Stable Diffusion. The application methods and implementation process of EvoMass and related GAI tools are explained in the following demonstration section.

2.2. BUILDING MASSING DESIGN OPTIMIZATION AND EXPLORATION

In the building massing design optimization and exploration stage, EvoMass is used to search for desirable building massing designs with competitive performance. With the use of EvoMass, designers first need to customize it to generate satisfying building massing designs through different user-defined parameters. After the parameter settings are completed, the Steady-State Island Evolutionary Algorithm (SSIEA) embedded in EvoMass is utilized to evolve the design population. The optimization can be driven by various building performance factors such as solar radiation, view, annual sunlight, etc. In Grasshopper, these building performance objectives can be simulated using Ladybug and/or ClimateStudio to measure the performance of the generated designs. With the optimized design, designers need to interpret the formal features revealed by the optimized designs and extract the design implications related to building performance, which can also further inform AI-assisted design exploration.

Moreover, the use of EvoMass also allows for multiple runs of optimization focusing on different performance objectives or using different user-defined parameters to generate building massing designs with different formal characteristics. In this way, the designer can collect more feedback pertinent to different design and performance aspects or iteratively refine the optimization setup and result.

2.3. ARCHITECTURAL LANGUAGE AND STYLE EXPLORATION

In the architectural language and style exploration stage, the NLP tools are first used for textual conversation for design ideation, and it does not necessarily occur after the completion of the PDO stage. The conversation with the NLP tools primarily extends the designer's perception of the design context and understanding of potential design styles into a conversational reflection. Several potential design ideas are fed back to NLP tools such as BLIP2 and GPT to provide further iterative feedback and optimize design guidance instructions.

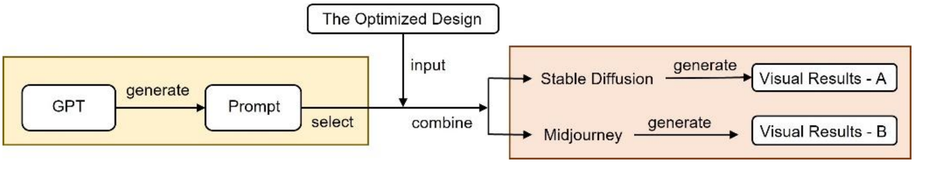


Figure 2. Design Workflow of the GAI Stage

CV tools such as Midjourney and Stable Diffusion are applied based on the optimal solutions obtained from the PDO stage. Combined with the design instructions fed back by NLP tools in the above-mentioned stage and the images of the massing forms provided by PDO, it can provide variations of visual feedback that envision design development possibilities and inspirations. The generated images are the consequences of the iterative process combined with the designer's design intent, performance optimization, and NLP tools' feedback. For example, it could provide variations of sustainable building design based on the input massing model that has considered the site environment in the first place. This workflow integrates site context consideration into GAI, greatly enhancing the practicality of applying GAI to generate controlled and contextualized variations of potential design development.

3. Demonstration

To demonstrate the efficacy of the proposed workflow, a case study is presented, describing the design of a high-rise mixed-use residential tower in Apleichau, Hong Kong, one of the most densely populated islands in the world. Figure 3 demonstrates the location of the selected site, which is alongside Main Street and currently has a

Municipal Building and four old tenement houses that are about to be demolished or renovated. The developer originally intended to renovate the old tenement houses into a "toothpick" building for residential purposes. However, toothpick buildings have the characteristics of obstructed lighting, high thermal radiation, and limited visibility due to their high-density forms, which are not suitable for living.

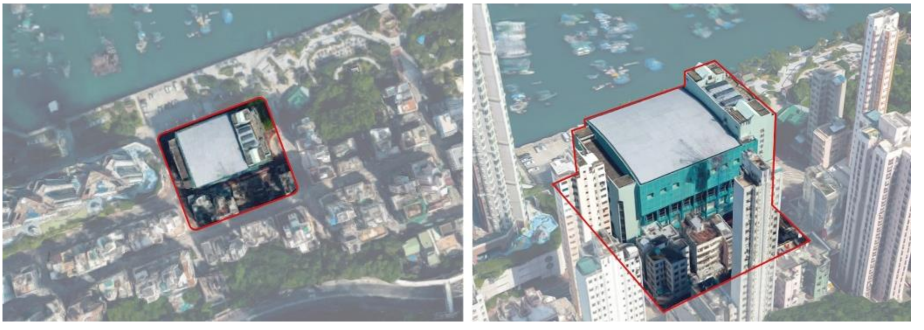


Figure 3. Site Condition and Surrounding Urban Context

Over recent years, the frequency of extremely high temperatures during Hong Kong's summers has consistently risen, accompanied by a deterioration in the urban heat island effect. Considering these, the proposed design workflow is applied to the case-study with the focus on alleviating the physical environmental challenges of the design of this block, as well as investigating the potential of using building and environmental performance as a driving factor in the design development process.

3.1. PERFORMANCE-BASED DESIGN OPTIMIZATION

In order to optimize the environmental performance of this urban massing design, two optimization objectives were selected: view and summer radiation. To obtain solutions that can simultaneously correspond to the two objectives, a single-objective optimization mode was used, and the two objectives were integrated into one fitness evaluation function using a weight-sum approach. To achieve a solution with more intricate and architectural appealing formal characteristics, the additive algorithm within EvoMass was used for design generation, which produces building massing forms by aggregating multiple sub-volumes within a predefined spatial boundary.

To obtain desirable solutions, three optimization runs were conducted at the PDO stage in this case study. The first and the second optimization runs were conducted in parallel but using different user-defined parameter settings to explore different directions in terms of formal characteristics. The first optimization uses an approach to controlling the overall size of the aggregated volumes, while the second optimization uses a more detailed control of the size of each sub-volume. The differences can be exhibited in the optimization results shown in the first two rows of Figure 4.

The first two rows list five high-performing designs from the first two optimization runs. Regarding the designs from the first optimization run (Group 1), the designs are monolithic and does not coordinate well with the surrounding urban context. In addition, the volumetric configuration in these designs is rather unclear, which makes it difficult to interpret the performance implications related to the building massing form. In contrast, the designs from the second optimization run (Group 2) display interlocking volumes and clear tendencies in terms of volumetric compositions. The slender upper part of the building can reduce solar heat gain while achieving a greater unobstructed view. Additionally, the solid bottom part of the building can make most of the volume within the shadow cast by the surrounding buildings to achieve a greater passive cooling effect. Nevertheless, the bottom part lacks a connection to the surrounding environment.

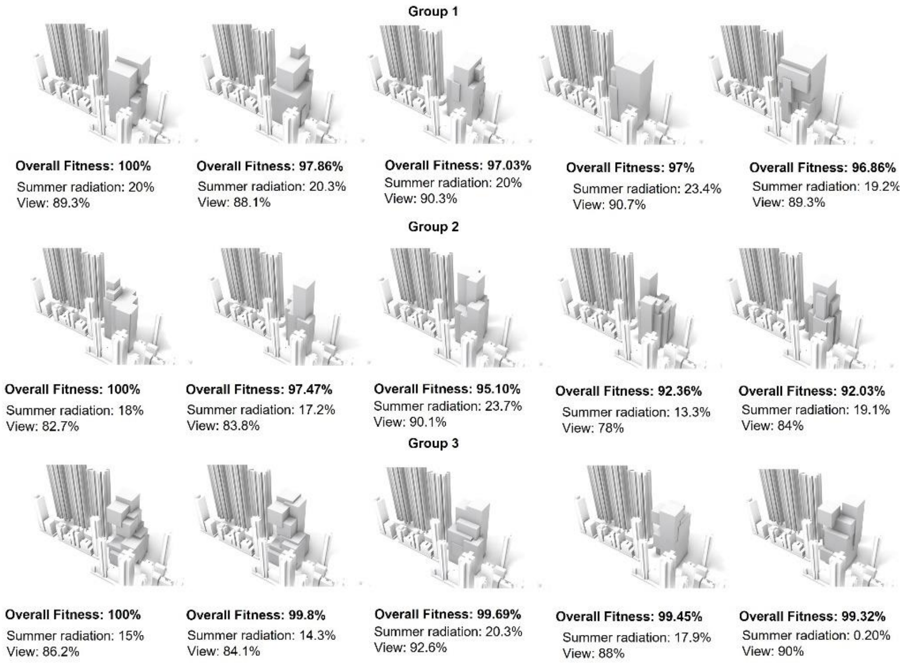


Figure 4. Building Massing and Performance Calculation Results

Based on the issues identified from the first two optimization runs, the third iteration of optimization was conducted, and user-defined parameters were changed to enhance the more even distribution of sub-volumes. The result shows that most optimized designs can overcome the shortcomings in the prior optimized designs. Within these designs, one solution achieving a desirable balance and compromise between performance and building design feasibility was selected (Fig. 5) as the basis for subsequent design exploration using GAI.

As shown in Figure 5, the selected design has multiple cantilevered blocks that enhance the façade surface with an unobstructed view. At the same time, these blocks also self-shade a large proportion of the building façade from excessive solar exposure, which lowers the overall solar radiance received by the building facade surfaces. The above analysis not only can help designers better understand the performance implications of the design strategies revealed by the optimized design, but can also provide hints for the subsequent design exploration using GAI.

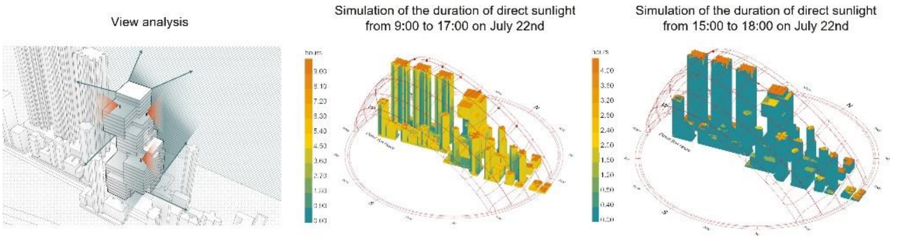


Figure 5. Analysis of the Causes of the Morphological Trend

3.2. ARCHITECTURE LANGUAGE AND STYLE EXPLORATION

In this stage, a large language model, GPT, is first chosen to obtain descriptions in terms of architectural languages and styles, which serves as the input for the architectural image generation in the next step. Since the GPT only receives and produces textual content, a prompt question template is formulated to guide the GPT in generating descriptions that can be interpreted in the image generation step. The prompt question template:

> "According to the design process, provide five sets of prompt words in different directions. Design progress: The current design is optimized according to x and y, hoping to design mixed-use buildings in a densely built city. Next, we want to start deepening the facade of the building, the surrounding landscaping, etc. Please refer to and follow the following format and examples for the prompt words of the AI drawings. Format: Project type + Architecture style + Architects Name + Facade + Rendering Style + View + Environment + Light. Example: museum, modernism, OMA Architecture, glass facade, white light inside, facing the road in downtown Hong Kong, mir rendering, bird's-eye view, natural light."

Prompts generation results from GPT:

- Public office building, Ecological style, Sustainable Architecture Studio, Green facade, Modern rendering style, Urban park landscape, Natural lighting.
- Hotel, Postmodernism, Green Architecture Firm, Adjustable shading devices, Natural ventilation design, Beach landscape, Natural light.
- Apartment building, Ecological style, Sustainable Design Group, Vertical garden facade, Stone exterior, City skyline view, Passive solar energy utilization, Soft lighting.
- Commercial complex, Modernism, Renewable Architects, Glass curtain wall, Green roof garden, City center street view, Natural ventilation, Bright natural light.
- Community center, Futurism, Eco-Friendly Design Studio, Adjustable light control facade, Technological rendering style, Surrounding Park landscape, Optimized lighting, and ventilation.

Finally, the third one is chosen in this demonstration, with a selected style of "Zaha Hadid architects" included manually.

In the next step, the selected optimized solution generated by EvoMass and the prompt obtained from the GPT is used as the input for architectural image generation. In this study, two CV tools, including Midjourney and Stable Diffusion, are used for architectural language and style exploration, which also investigates the differences between these two tools for image generation.

Stable Diffusion implements image generation control through different ControlNet models (Fig. 6), including MLSD (straight line detection), Depth (3D depth detection), and Canny (edge detection). It also integrates the weight of the prompt, mainly divided into 1) "balanced", 2) "the prompt is more important", and 3) "the ControlNet is more important". After experiments and comparison, Stable Diffusion 1.5 and MLSD model were adopted, with two weights, including "balanced" and "the prompt is more important".

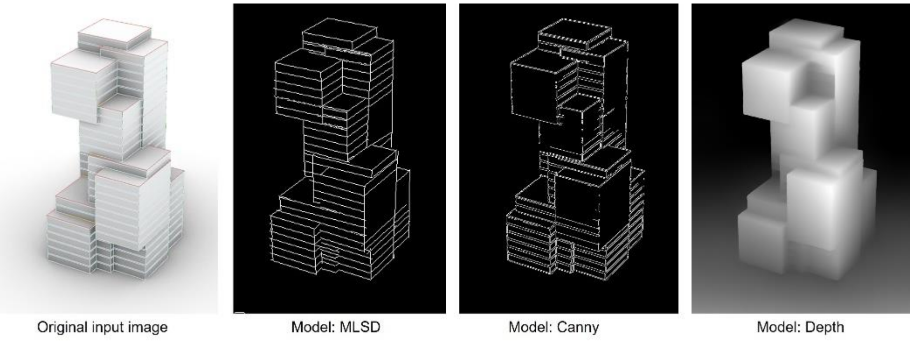


Figure 6. The Original Image and Filtered Images of the ControlNet Models

MLSD model provides precise detection of this building model outline, which is all composed of straight lines. Regarding the ControlNet, the "balanced" mode is able to provide a high consistency of the massing model, and number of floors, while the "the prompt is more important" option creates less-controlled variations such as generating buildings with round corners. Still, the perspective, and overall building height and main gestures of extrusions could be observed. Therefore, these two ControlNet models provide design variations with different consistency depending on how many differences there are from the original model the designers are looking for.

Differing from Stable Diffusion, the control of the generated architectural form of Midjourney shows less relevance to the input images, while the generated design images have significantly more variability. In the design examples illustrated in Figure 7, the v5.2 version of Midjourney is being experimented with three different image weights (0.5, default (1), 2) to compare the design variability of the generated design images. When the image weight was set to 2, the outputs had the greatest similarity with the input image, which is suitable when the massing model is preferred to be developed with little variation. When image weight was set to default (1), it displayed similar results with more creative generation in the context. For the lowest degree of image weight (0.5), it transformed the massing model most creatively with a more photorealistic background generated. Therefore, low image weight is useful when designers look for more creative options. In contrast, high image weight is more suitable for getting controlled variations while trying to keep the form of the original massing model.

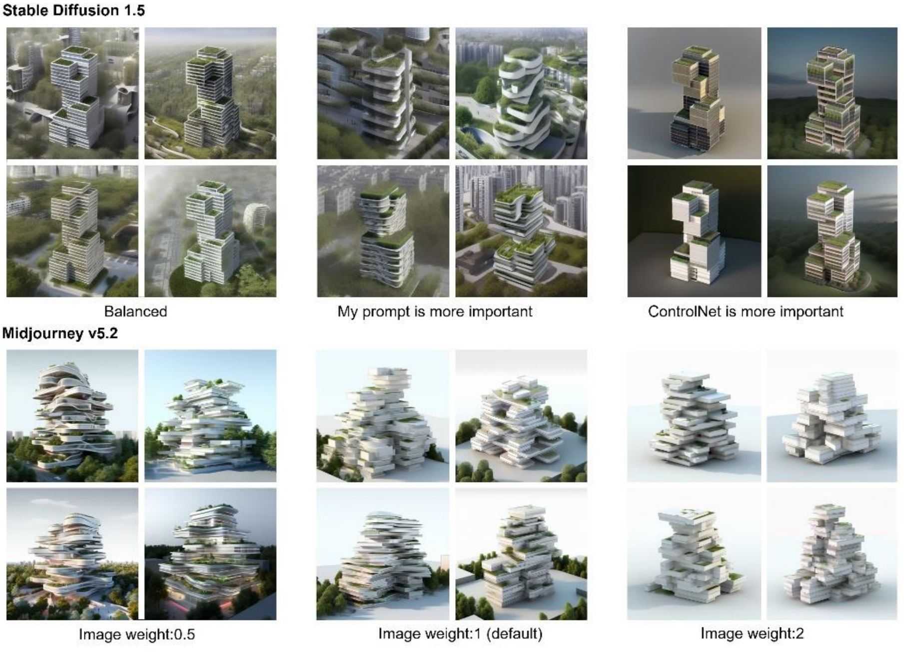


Figure 7. Image Generations Based on Optimized Massing Model

4. Discussion and Conclusion

The above demonstration showcases the potential of the integration of computational optimization and GAI in assisting early-stage architectural design exploration. The integration addresses the shortcomings inherited within the two techniques and achieves greater collective utility by synthesizing their strengths for design optimization and exploration.

As shown in the case study, the proposed design workflow allows for a rapid decision-making approach for early-stage design ideation and conceptual development. Although the application solely relying on computational optimization can produce a sizeable number of design options for design space exploration, it is often challenging and overwhelming for designers, especially inexperienced ones, to efficiently evaluate the potential of different design options in other architectural design aspects and select a viable one for subsequent design development. In this regard, the proposed workflow can facilitate designers to quickly visualize the architectural feasibility of the abstract building geometry generated by computational design optimization. At the same time, this approach also extends the utility of GAI for performance-based architectural design by providing more contextualized solutions for the designers.

Despite its potential utility, certain aspects need further improvement and development. First, the design produced by the GAI may contradict the design generated by computational design optimization in terms of design strategies or features associated with building performance. For example, the effectiveness of the self-shading revealed by the computational optimization can be undermined by the large transparent façade surfaces potentially created by the GAI. Thus, the awareness of the performance implications shown in the optimized designs, no matter from human designers or GAI, is a critical step to generate designs more consistent with the optimized design regarding the building performance. Apart from this, the user-friendliness of the design tool and the automation of the design process are also critical future research directions.

To conclude, this paper presents a study investigating the integration of PDO and GAI for early-stage architectural design exploration. The proposed design workflow is established by integrating EvoMass and various GAI tools, and its efficacy is demonstrated through a case-study design. The result shows that the use of the approach allows for rapid and performance-pertinent exploration of building massing and architectural languages and styles. Finally, the relevance of the study is discussed, with limitations and future research directions identified.

Acknowledgements

This study is supported by Xi'an Jiaotong - Liverpool University Research Development Fund (RDF-23-01-107) and Summer Undergraduate Research Fellowship (SURF-2023-0024).

References

Cheung, L. H., & Dall’Asta, J. C. (2023). Exploring a Collaborative and Intuitive Framework for Combined Application of AI Art Generation Tools in Architectural Design Process. In Bertie MullerBertie Muller (Ed.), Proceedings of the AISB Convention 2023 (pp. 122–130).

Li, S., Liu, L., & Peng, C. (2020). A Review of Performance-Oriented Architectural Design and Optimization in the Context of Sustainability: Dividends and Challenges. Sustainability, 12(4), 1427. https://doi.org/10.3390/su12041427

Wang, L. (2022a). Understanding the Span of Design Spaces. In D. Gerber, E. Pantazis, B. Bogosian, A. Nahmad, & C. Miltiadis (Eds.), CAAD Futures 2021: Computer-Aided Architectural Design. Design Imperatives: The Future is Now (Vol. 1465, pp. 288–297). Springer Singapore. https://doi.org/10.1007/978-981-19-1280-1_18

Wang, L. (2022b). 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., Chen, K. W., Janssen, P., & Ji, G. (2020). Enabling Optimisation-based Exploration for Building Massing Design - A Coding-free Evolutionary Building Massing Design Toolkit in Rhino-Grasshopper. RE: Anthropocene, Design in the Age of Humans - Proceedings of the 25th International Conference on Computer-Aided Architectural Design Research in Asia, CAADRIA 2020, 1, 255–264. https://doi.org/10.52842/conf.caadria.2020.1.255

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), 761–780. https://doi.org/10.1016/j.foar.2022.02.002

Zhu, Q., & Luo, J. (2023). Generative Design Ideation: A Natural Language Generation Approach. In Design Computing and Cognition’22 (pp. 39–50). Springer International Publishing. https://doi.org/10.1007/978-3-031-20418-0_3

ARCHICLIP

Enhanced Contrastive Language–Image Pre-training Model With Architectural Prior Knowledge

SHENGTAO XIA[1], YIMING CHENG[2] and RUNJIA TIAN[3]
[1] *Architectural Association School of Architecture*
[2] *Xi'an Jiaotong-Liverpool University*
[3] *Harvard University*
[1] *shengtao.xia@aaschool.ac.uk , 0009-0007-3106-521X*
[2] *y1mingc201@gmail.com, 0009-0008-6101-8710*
[3] *runjiatian@gmail.com, 0000-0002-5983-9754*

Abstract. In the rapidly evolving field of Generative AI, architects and designers increasingly rely on generative models for their workflows. While previous efforts focused on functional or building performance aspects, designers often prioritize novelty in architectural design, necessitating machines to evaluate abstract qualities. This article aims to enhance architectural style classification using CLIP, a Contrastive Language–Image Pre-training method. The proposed workflow involves fine-tuning the CLIP model on a dataset of over 1 million architecture-specific image-text pairs. The dataset includes project descriptions and tags, aiming at capturing spatial quality. Fine-tuned CLIP models outperform pre-trained ones in architecture-specific tasks, showcasing potential applications in training diffusion models, guiding generative models, and developing specialized search engines for architecture. Although the dataset awaits human designer review, this research offers a promising avenue for advancing generative tools in architectural design.

Keywords. Machine Learning, Generative Design, Contrastive Language-Image Pre-Training, Artificial Intelligence.

1. Introduction

With the fast development of Generative AI, generative tools are taking up an increasing share of day-to-day workflows of architects and designers. Usually, generative models use algorithms to evaluate the differences between generated outputs and desired ground truth, using this feedback to improve the generation quality. Many attempts have been made to assess architecture with algorithms, but most focus on functional or performance aspects, such as structural stabilities. For example, Zheng et al. utilises an iterative machine learning algorithm to enhance the topological design exploration of compression-only shell structures, considering structural performance

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 69-78. © 2024 and published by the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA), Hong Kong.

and construction limitations (Zheng et al., 2020).

However, designers often value spatial quality more than other aspects of generative tools in the context of architecture design. This leads to the demand for algorithms to assess more conceptual and abstract beyond pragmatic qualities in architectural design, such as spatial qualities. In the realm of architecture, a profound understanding of spatial quality is essential, directly shaping individuals' perception, experiences, and interactions. Space, extending beyond mere functional dimensions, acts as a canvas for emotions, culture, and social elements. Thoughtful spatial design guides individuals through movement, crafting unique genius loci Spatial cognition plays a central role in architectural design, influencing how individuals comprehend and interact with their surroundings. By exploring how people perceive and engage with space, AI-generated tools have the potential to produce outcomes that are both aesthetically meaningful and practical.

2. Related Work

2.1. RULE-BASED ASSESSMENT

Previous work focused on CNN/DPM finesse to achieve the purpose (Xu et al., 2014).

2.2. MACHINE LEARNING BASED ASSESSMENT

As an alternative, several state-of-the-art generative models, such as Stable Diffusions, a Latent Diffusion model that generates images from text inputs, use CLIP, a Contrastive Language–Image Pre-training method proposed by OpenAI, as a text embedding to condition the generation (Radford et al., 2021). CLIP can evaluate the semantic similarity of given text and image pairs and perform classification tasks in a zero-shot manner. Zero-shot learning offers a promising avenue to broaden the scope of cognitive capabilities in the design process, transcending limitations tied to specific application-related problems (Larochelle et al., 2008).

However, the original CLIP model performs poorly on architecture-specific classification tasks due to its dataset's lack of prior knowledge and source of truth, causing latent diffusion models trained with the default CLIP pipeline to underperform in generation tasks in the architecture domain. Improving the prior architectural knowledge in these pre-trained models is the key to enhancing the generation quality of latent diffusion models conditioned on these CLIP models. A specific model requires task-specific modifications and training from scratch for downstream tasks; fine-tuning would be a more effective way (Howard & Ruder, 2018). The finalization could be used to classify architectural components, conduct assessments throughout the process, and even directly influence design decisions.

3. Methodology

Therefore, we propose a workflow to compile architecture-specific data and leverage it for fine-tuning the CLIP model while preserving its zero-shot capabilities.

CLIP is a multi-modal model based on contrastive learning. Unlike some contrastive learning methods in computer vision, CLIP's training data consists of text-image pairs, where each pair consists of an image and its corresponding text description. The aim is for the model, through contrastive learning, to learn the matching relationship between these text-image pairs. CLIP comprises two models: Text Encoder and Image Encoder. The Text Encoder is utilized to extract text features and can employ the text transformer model commonly used in Natural Language Processing (NLP), while the Image Encoder is employed to extract image features and can use common Convolutional Neural Network (CNN) models or vision transformers.In this experiment, both the Text Encoder and Image Encoder used are vanilla models. A custom dataset was employed to train and fine-tune CLIP models using various prompt methods. Multiple checkpoints were fine-tuned using OpenCLIP (Cherti et al., 2023), an open-source implementation of the OpenAI CLIP model. Some checkpoints are focused solely on project tags, while others include artistic descriptions, capturing the essence and atmosphere of the creative endeavour. The impact of fine-tuning was assessed by conducting inference on image classification tasks and image-text-pair correlation tasks, comparing the results of our fine-tuned checkpoints with the original pre-trained checkpoints from OpenCLIP.

3.1. DATASET

We construct a comprehensive image-text dataset comprising over 1 million images paired with corresponding textual descriptions. This dataset is curated by sourcing image and text pairs from publicly accessible internet resources. The images exclusively consist of photographs of recently completed projects, while the text descriptions encompass project-related details. These descriptions often convey the essence of projects in an abstract or artistic manner. Additionally, the information includes project tags, specifying components present in the images, such as building materials or architectural elements.

3.2. PROMPT

As CLIP evaluates the similarity between a sentence and an image, the training of the CLIP model with a classification dataset often involves prompt engineering (Radford et al., 2021). Typically, tags or classes are formed into sentences using determiners.

In our work, we employ a diverse array of tags to describe images from various perspectives. These include:

- MediaTags: Type of graphic media, e.g., 'photograph', 'architectural drawing'.
- UtilityTags: Describing spatial utilities of the building, e.g., 'bathroom', 'courtyard'.
- ProjectTags: Indicating the program of the project, e.g., 'residential', 'public', 'office'.
- ElementTags: Detailing architectural elements shown in the image, e.g., 'stairs', 'arch', 'handrail'.

Notably, the images we collected may not encompass all tag building types. This aspect introduces certain limitations to our dataset, which we will discuss further in Sections 5.1 and 5.3.

We developed a procedure to convert these mixed tags into sentences, prioritizing them based on their significance in the image, thereby setting up image-text pairs.

For an image encompassing all four tag types, the format is as follows:

"This is a {MediaTag} of {UtilityTag} in {ProjectTag}, showing {ElementTag}."

In cases where an image lacks certain tag types, for instance, only project and element tags, the format adapts:

"This is {ProjectTag}, including details of {ElementTag}."

Additionally, if project information collected includes the design studio and description, these elements are also incorporated:

"{TagPromptedSegment}, designed by {DesignStudio}, with the following description: '{Description}'."

A few fully prompted text-image pairs are displayed below in Figure 1.

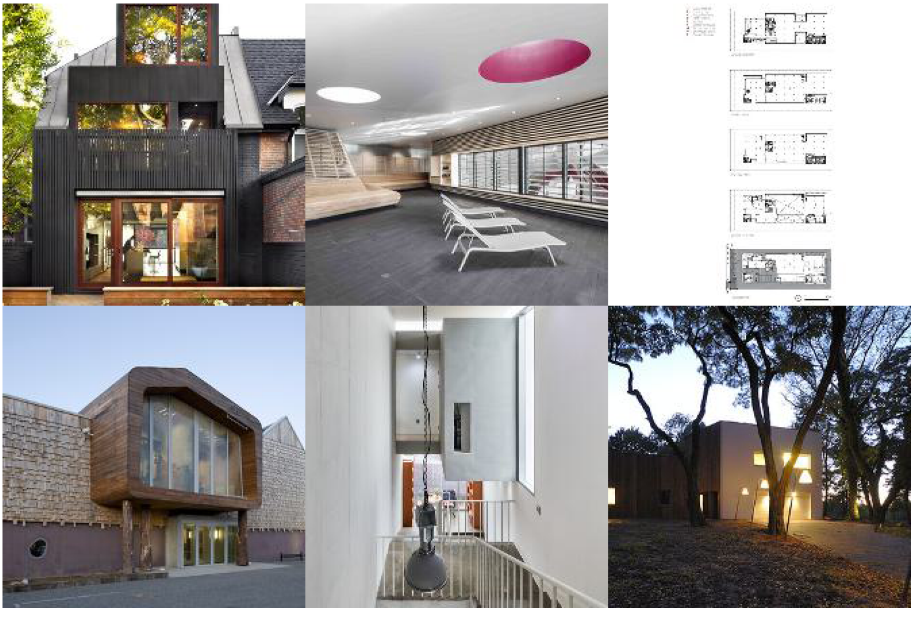


Figure 1. 6 images extracted from training dataset with their corresponding text listed below (from left to right, top to bottom). Images were upscaled from 224x244

- this is a garden in a house, include details of doors, facades, lighting, designed by plus tongtong, that has the following descriptions: "+tongtong Transforms Traditional Toronto House into Tasteful Modern Home that Honors East-end Neighborhood"
- this is an image of an architecture project, designed by auer weber, that has the following descriptions: "Sourcane is a new sports, leisure and wellness swimming hall in Douai, in the North of France. Due to its location, the aquatic centre will be an essential part of the future eco-quarter of Le Raquet, a new city district, at whose

heart will be a landscaped park. The new swimming hall lies at its northeastern end, at the interface between the artificial landscape of the park and the urban structure of the city. The project is oriented towards the future tramway station and a central urban square."

- this is an architectural drawing of a factory, designed by dp architects, that has the following descriptions: "The building housing the new headquarters of Sunray Woodcraft Construction is one of the first to be completed as part of the newly positioned International Furniture Hub in Sungei Kadut, Singapore. It presents an opportunity to look afresh at the light industrial factory type, stacking production processes in order to optimise working conditions."
- this is a museum, include details of facades, doors, designed by guinee et potin architects, that has the following descriptions: "French architectural photographer Stphane Chalmeau shared with us the Rennes Metropole Museum by french architects Guine et Potin."
- this is a kitchen in an apartment, include details of stairs, facades, handrails, lighting, designed by leau design, that has the following descriptions: "The apartment estate market which seemed not to be withered is cooling now. As the imprudent house-poor, who borrows the money to buy a house even if the interest rate is low, disappears, the fever of becoming to farming for a moment with the longing for the life in a country house is passing like a wind. And as changing the interest to the investment of the profitable real estate, the interest about a leasing profitable building of neighbourhood living facilities with an integration of a habitation house is increasing."
- this is a house, include details of facades, designed by neostudio architekci, that has the following descriptions: "This project is located on a picturesque plot that originally was a home for seed drying installation of Agricultural University - and with its magnificent Acacia trees plantation and natural splendor was a design challenge for us."

3.3. GENERAL PURPOSE DATASETS AND PRE-TRAINED MODEL SELECTION

We developed two distinct types of image-text paired datasets: the first comprising solely prompted tags, and the second encompassing both the prompted tags and project descriptions.

Each dataset underwent a split, allocating 60% for the training dataset and 40% for use as a validation/benchmark dataset. The 60% designated for training was then combined with two extensive image-text-pair datasets: LAION 400M (Schuhmann et al., 2021) and CC3M (Sharma et al., 2018). These datasets are widely recognized for their broad range of general content and are typically utilized in CLIP training.

Owing to limitations in computational resources, we incorporated only a small portion of the LAION 400M dataset and one-fifth of the CC3M dataset. This selective integration, in combination with our training dataset, was employed to fine-tune our model.

For the pre-training model, we utilized ViT-B/16 LAION400M (Release Pretrained Weights · mlfoundations/open_clip. (n.d.). GitHub. 2023), a model previously trained by OpenCLIP (Cherti et al., 2023).

Model	Process	Archiclip Dataset (ours)				Laion-400m		Cc3m	Training dataset size
		tag and descrip-tions		only tag					
		60%	40%	60%	40%	0.15 %	100 %	18.1 %	
ArchiCLIP-tagdcsp (ours)	finetune	~600k				~600k		~600k	~1.8m
	validation/ benchmark		~400k						
ArchiCLIP-tagonly (ours)	finetune			~600k		~600k		~600k	~1.8m
	validation/ benchmark				~400k				
ViT-B/16 LAION400M (open-clip)	training (open-clip)						~413m		~413m
	benchmark		~400k		~400k				

Table 1. Table showing the datasets we used and how they were merged before proceed with training

4. Experiments

We fine-tuned the ViT-B/16 LAION400M model on our two merged datasets for 8 epochs. The initial learning rate was set to 1e-5 with a learning rate scheduler featuring cosine decay to optimize the gradient descent.

The finetuning process was carried out on an Nvidia A100 80GB graphic card with an batch size of 600 to maximize the utilisation of GPU memory.

4.1. TRAINING

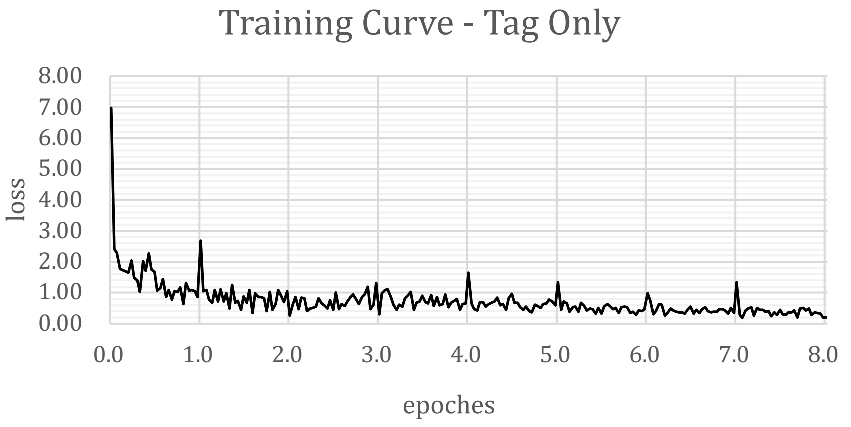


Figure 2. The training curve of tag only dataset showing CLIP loss against epoches

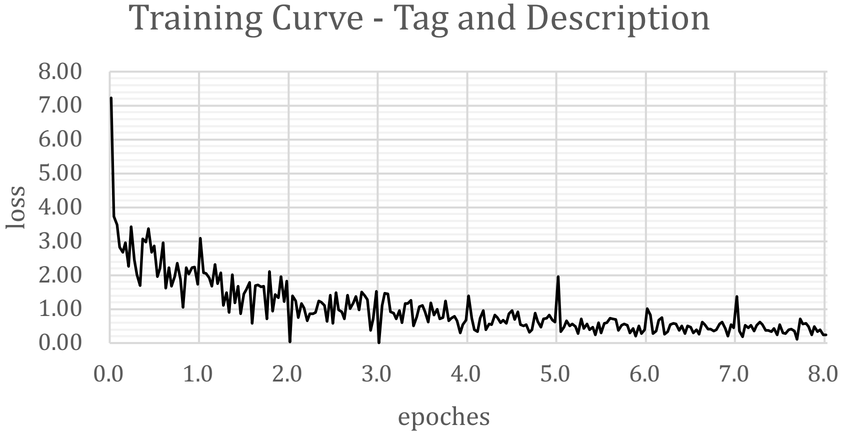


Figure 3. The training curve of tag and description dataset showing CLIP loss against epoches

Figure 2 and Figure 3 demonstrate the training curve during fine-tuning. Model already started showing traces of overfitting as the training curve flatten around 6th epoch. We believe that this is partially due to the size of our dataset, further discussed in 5.1

4.2. EVALUATION

Due to our dataset not being a single-label classification dataset, as mentioned earlier in section 3.2, the common benchmarks for CLIP models, which typically evaluate zero-shot classification tasks, cannot be applied to ArchiCLIP.

Therefore, we developed our own benchmark. For each image, this benchmark calculates a CLIP correlation against all possible tags. The tags with the highest correlation scores are selected for comparison with the ground truth in the dataset. The number of tags chosen matches the count of tags in the ground truth. For each correct tag identified, the model earns a score equal to the reciprocal of the number of ground truth tags in that corresponding tag category.

$$S_{tagtype} = score\ of\ model\ achieved\ in\ evaluating\ \{tagtype\}\ tags$$
$$N_{tagtype} = counts\ of\ images\ that\ have\ at\ least\ one\ \{tagtype\}\ tags$$
$$n_{i,tagtype} = counts\ of\ \{tagtype\}\ tags\ belong\ to\ image\ \{i\}$$
$$c_{i,tagtype} = correct\ \{tagtype\}\ predictions\ belong\ to\ image\ \{i\}$$

$$S_{tagtype} = \frac{1}{N_{tagtype}} \sum_{i=1}^{N_{tagtype}} \frac{c_{i,tagtype}}{n_{i,tagtype}}$$

Checkpoint	epoch	MediaTag	UtilityTag	ProjectTag	ElementTag
ViT-B-16 (open-clip)	n/a	0.620	0.450	0.245	0.467
ArchiCLIP description (ours)	5	0.752	0.698	0.135	0.644
	6	0.510	0.677	0.090	0.559
	7	0.508	0.640	0.141	0.627
	8	0.610	0.679	0.116	0.575
ArchiCLIP tagonly (ours)	5	0.575	0.763	0.204	0.690
	6	0.563	0.762	0.240	0.707
	7	0.573	0.742	0.231	0.735
	8	0.541	0.783	0.275	0.707

5. Limitations and Future Improvements

When evaluating our finetuned model, we identified several limitations affecting its performance.

5.1. DATASET

As mentioned in section 3.2, not all images in our dataset contain all four types of tags. Upon further manual examination, we found that in many cases, the tags sourced from the internet are a subset of what is depicted in the images. For instance, an image may feature a 'staircase' that is not reflected in the tags. This discrepancy likely arises from sourcing images from the internet, where the information on webpages may not fully capture all elements in the image.

This inconsistency impacts the model's performance. The loss function of the CLIP model is defined as the difference between the cosine similarity matrix and the identity matrix (Radford et al., 2021). Text that fails to fully describe the image contributes to loss, even if the model's prediction is accurate, thus making the dataset less reliable.

Moreover, most CLIP models are trained on much larger datasets with significant computational power. Studies have shown that the performance of CLIP models generally improves with larger datasets (Gadre et al., 2023). For comparison, the pre-trained model we used for fine-tuning, ViT-B-16 LAION400M, was initially trained on the Laion400M dataset, comprising approximately 413 million data points (Cherti et al., 2023). Our dataset, combined with a portion of Laion400M and CC3M, only reached 1.8 million data points, less than 0.5% of the initial pre-trained model's size. This substantial difference contributes to the performance loss.

5.2. BENCH MARK AND EVALUATION

CLIP models are generally adept at evaluating correlations between objects in an image and descriptive text. However, their ability to assess the correlation between an image and abstract descriptions remains undetermined. To our knowledge, no existing research focuses on evaluating these correlations in the field of architecture. This lack of precedent makes it challenging to gauge our model's success in this area.

5.3. AREAS OF IMPROVEMENTS

We believe that with targeted improvements, the model's performance can be further enhanced.

- Manual Labeling: With sufficient architects involved, we could more accurately relabel our dataset, thereby improving its accuracy and optimizing tag comprehensiveness and consistency. This addresses the limitation mentioned in 5.1
- Human Benchmark: Engaging a sufficient number of architects, we could establish a human benchmark for assessing the correlation between images and abstract descriptions. This would provide deeper insights into standard performance for such image-text tasks, addressing the limitation mentioned in 5.2
- Classification Tasks: One of the primary research areas in computer vision is image classification. The CLIP model, a sub-field of this area, has made significant strides in processing natural language and in zero-shot learning. Nonetheless, due to the unique format of its training dataset, certain methods that improve image classification accuracy are not applicable for enhancing the accuracy of CLIP's image-text pair matching tasks. By reformating our dataset, we might transform the training dataset into multiple image classification datasets and apply existing methods (Wortsman et al., 2022) from related fields to optimize the model's performance

6. Conclusion

Compared to the pre-trained checkpoints, our fine-tuned model performs better in architecture-specific image classification tasks and demonstrates some capabilities in relating abstract and artistic descriptions with architectural photographs. In a practical context, the post-optimized CLIP model can find application in various ways. For example, it can train diffusion models, serve as a guiding discriminator for other generative models, and even be harnessed to develop a specialised semantic-based searching engine explicitly tailored for architectural purposes. It is foreseeable that ArchiCLIP will play a critical role in AI-generated process. Limited by our resources, our dataset has yet to undergo review nor filtering by human designers, leaving room for potential improvements in dataset qualities.

References

Cherti, M., Beaumont, R., Wightman, R., Wortsman, M., Ilharco, G., Gordon, C., ... & Jitsev, J. (2023). Reproducible scaling laws for contrastive language-image learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 2818-2829).

Howard, J., & Ruder, S. (2018). Universal language model fine-tuning for text classification. arXiv preprint *arXiv preprint arXiv:1801.06146*.

Larochelle, H., Erhan, D., & Bengio, Y. (2008, July). Zero-data learning of new tasks. In *AAAI* (Vol. 1, No. 2, p. 3).

Phelan, N., Davis, D., & Anderson, C. (2017, May). Evaluating architectural layouts with neural networks. In *Proceedings of the Symposium on Simulation for Architecture and Urban Design* (pp. 1-7).

Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., ... & Sutskever, I. (2021, July). Learning transferable visual models from natural language supervision. In *International conference on machine learning* (pp. 8748-8763). PMLR.

Xu, Z., Tao, D., Zhang, Y., Wu, J., & Tsoi, A. C. (2014). Architectural style classification using multinomial latent logistic regression. In *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part I 13* (pp. 600-615). Springer International Publishing.

Zheng, H., Moosavi, V., & Akbarzadeh, M. (2020). Machine learning assisted evaluations in structural design and construction. *Automation in Construction*, 119, 103346.

Schuhmann, C., Vencu, R., Beaumont, R., Kaczmarczyk, R., Mullis, C., Katta, A., Coombes, T., Jitsev, J., & Komatsuzaki, A. (2021). LAION-400M: Open Dataset of CLIP-Filtered 400 Million Image-Text Pairs. *ArXiv:2111.02114* [Cs]. https://arxiv.org/abs/2111.02114

Sharma, P., Ding, N., Goodman, S., & Soricut, R. (2018). Conceptual Captions: A Cleaned, Hypernymed, Image Alt-text Dataset For Automatic Image Captioning. *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics* (Volume 1: Long Papers). https://doi.org/10.18653/v1/p18-1238

Release Pretrained Weights · mlfoundations/open_clip. (n.d.). GitHub. Retrieved September 2023, from https://github.com/mlfoundations/open_clip/releases/tag/v0.2-weights

Gadre, S. Y., Ilharco, G., Fang, A., Hayase, J., Smyrnis, G., Nguyen, T., Marten, R., Wortsman, M., Ghosh, D., Zhang, J., Orgad, E., Entezari, R., Daras, G., Pratt, S., Ramanujan, V., Bitton, Y., Marathe, K., Mussmann, S., Vencu, R., & Cherti, M. (2023, October 20). DataComp: In search of the next generation of multimodal datasets. *ArXiv.org*. https://doi.org/10.48550/arXiv.2304.14108

Wortsman, M., Ilharco, G., Kim, J. W., Li, M., Kornblith, S., Roelofs, R., Gontijo-Lopes, R., Hajishirzi, H., Farhadi, A., Namkoong, H., & Schmidt, L. (2022). Robust fine-tuning of zero-shot models. *ArXiv:2109.01903* [Cs]. https://arxiv.org/abs/2109.01903

ARCHITECTURAL GENERATIVE MODEL EVALUATION METHODS: IMAGE QUALITY ASSESSMENT METRICS AND EXPERT-BASED APPROACH

Taking the Chinese Campus Layout as an Examples

YING LIN[1] and FEI YE[2]
[1,2] Xi'an University of Architecture and Technology.
[1]lynn813910@gmail.com, 0009-0008-0132-381X
[2]feiye@xauat.edu.cn, 0009-0002-1444-2419

Abstract. The feasibility of using machine learning methods to generative architectural design solutions has been widely recognized as an effective in enhancing innovation, diversity, and efficiency of solutions. However, in generative design methods, the accuracy and quality of design results often rely on empirical evaluation of expert, which is challenging to evaluate and quantify by unified standards. This paper proposes a comprehensive method for evaluating model performance in architectural design tasks. The evaluation is based on computational criteria (i.e., FID, IS, SIMM indicators) and expert system criteria. The computational metrics will measure the distance, diversity, and similarity between the feature vectors of the real image and the generated image. In contrast, the expert criteria will measure the accuracy, intentionality, and rationality of the layout scheme. This study applies this framework to evaluate three widely used generative models in architectural design: GANs, Diffusion Models, and VAE. The framework also guides the optimization of generative models in architectural applications and assists architects in validating generative outcomes with more efficient workflows.

Keywords. Deep Learning, Generate Design, Evaluation Metrics, Campus Planning.

1. Introduction

Recently, with the development of computing power and technological breakthroughs, artificial intelligence and its application scenarios have experienced significant growth, especially in the fields of NLP (Natural Language Processing) and CV (Computer Vision). The widespread use of deep learning technology is driving architecture towards a new paradigm of digital design, playing an important role in the field of architecture and urban planning. This interdisciplinary approach uses complex neural networks to analysis, predict, and improve aspects of building and urban environment. Key application areas include design generation and optimization, urban data analysis,

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 79-88.

prediction model, building performance simulation, historic preservation and reconstruction, construction management, and smart cities.

However, the rapid iteration of Artificial Intelligence Generated Content (AIGC) technology has significantly improved the performance of generative models, leading to the emergence of many variants with superior capabilities. These advances have come from typical generative models widely used in architecture research, such as GANs (Generative Adversarial Networks) (Goodfellow et al., 2014), Diffusion models (Ho et al., 2020), VAEs (Variational Auto-Encoders) (Kingma and Welling. et al., 2013), and Transformers (Vaswani et al., 2017). They often employ specific evaluation methods that are unique to computer science. Expert systems must empirically evaluate the accuracy and quality of results produced by generative design models. However, there is currently a lack of standardized evaluation frameworks in architectural research, which remains unavailable for objective measurement and quantification through consistent criteria. This research gap highlights the need to develop a comprehensive, interdisciplinary evaluation framework that can effectively combine computer science methods with the practical and aesthetic factors inherent in architectural design. Such a framework would facilitate not only the rigorous evaluation of generative models but also their practical application in architecture and urban planning.

1.1. GENERATIVE MODELS LITERATURE REVIEW

Generating new image from qualifying conditions is one of the challenging tasks in CV. So it have received a lot of attention in machine learning for their ability to generate new data instances that mimic real-world data distributions.

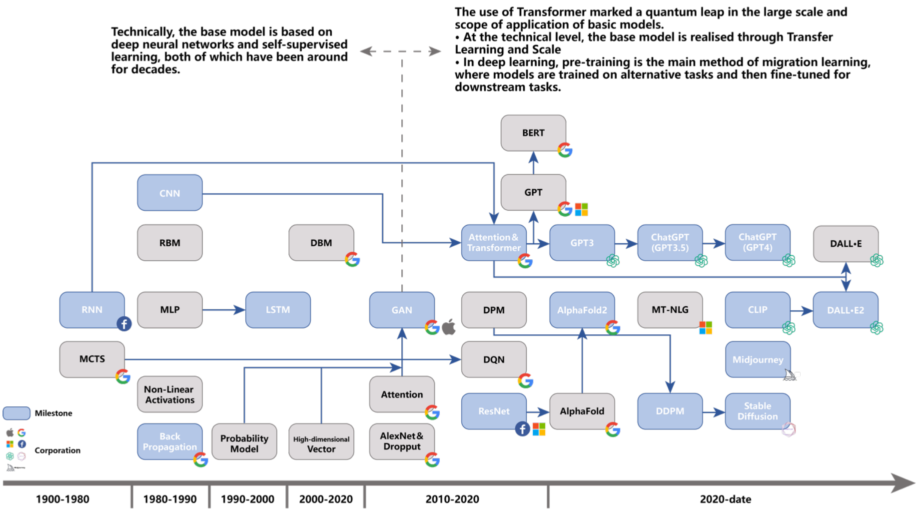


Figure 1. Generation model development process

With the proliferation of generative model technology (Figure 1), architectural generative design is currently experiencing a significant era of opportunity, particularly in the realms of production tools and innovative methodologies. The application of artificial neural networks to the field of architectural design generation can be traced back to GANs, with several scholars undertaking studies on the transference of architectural facade styles and the creation of architectural plans. In June 2020, diffusion models garnered widespread attention. Subsequently, OpenAI's research, for the first time, demonstrated superiority over GANs (Dhariwal and Nichol, 2021). In 2022, OpenAI released DALL-E 2, Google introduced Imagen, and Stability AI made Stable Diffusion open source, marking the formal advent of a new era in the field of image generation. In 2023, research on diffusion models continues to optimize model training methods and multiple-modality. According to George Guidai, the latest advances in NLP and Diffusion Models will lead to significant changes in the way of architectural design (Guida, 2023).

1.2. RIDING ARTIFICIAL INTELLIGENCE WAVE IN ARCHITECTURE

In recent years, architectural image generation research has advanced significantly with the aid of cutting-edge computer technology, enabling architects to explore design possibilities, despite not achieving full autonomy.In the early stages of schematic design, Huang and Zheng (Huang and Zheng, 2018) and Chaillou (Chaillou, 2020) used pix2pix or other modified GANs to generated floor plans of apartments, gradually learning the locations of doors and windows. Sun (Sun et al., 2022) and Ali (Ali and Lee, 2023) explored architectural facade generation. Technological advances have also facilitated architectural rendering and representation studies, for example Wang (Wang et al., 2023) and Meng (Meng, 2022). Due to the complexity of the physical space of buildings, 3D generation has been the focus of research in the field of architecture, Zheng (Zheng, 2019) and Pang (Pang and Biljecki, 2022) have explored the methods of generating high-rise and street building 3D models respectively. YOUSIF (YOUSIF and BOLOJAN, 2021) studied the application of pix2pix model in automated building performance simulation, and SSIM index was introduced to evaluate the results, achieving a score of 0.94. The average similarity between Jia's (Jia, 2021) predicted daylight autonomy maps and the simulation results is as high as 91.51%.

1.3 EVALUATE THE GENERATION MODEL

When using generative models for architectural image generation tasks, the aim is to obtain high-quality generated images, mainly considering the following quantitative and qualitative aspects:

- The quality of the image itself, such as whether it is clear, whether it is realistic, whether it is diverse;
- The expert opinion that is similar to the ground truth, such as whether the layout is reasonable and whether the requirements of the specification are taken into account, among others.

In order to make a fair comparison between the three main categories and five sub-categories of models, we use pre-trained classifiers for computational purposes. Specifically, there is a class of automatic evaluation criteria within the computational metrics, that can be used to quickly measure the quality of the generated images, including IS for measuring the diversity and quality of the generated images; FID for comparing the similarity between the distribution of the generated images and the distribution of the real images. Furthermore, we measure the structural similarity between generated and real images using the SSIM.

2. Methodology

With reference to previous research experience, the evaluation of generative models in this paper revolves around four parts: 1. Model Training, 2. Image Generation, 3. Evaluation, 4. Analysis of Results (Figure 2).

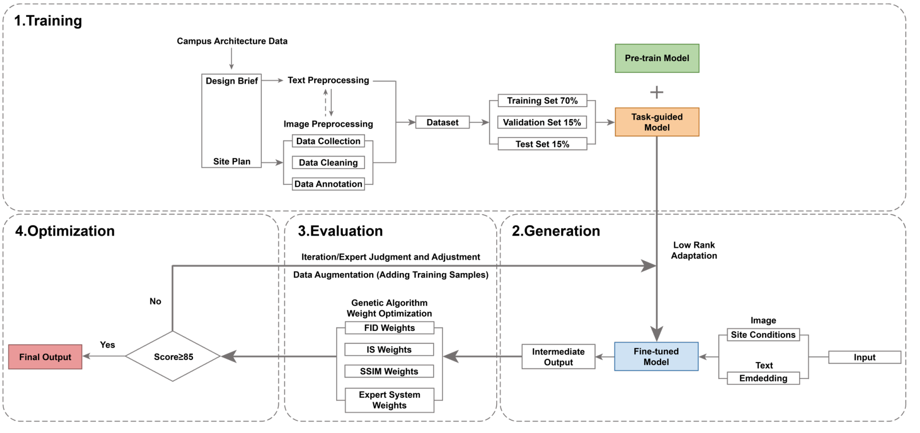


Figure 2. Workflow of the methodology

2.1. DATASET CONSTRUCTION

The dataset of this study, focusing on Chinese Campus Layout research, has three parts as data sources: Mapbox, public cases from ArchDaily, and design projects from our subject group. The publicly available master plans were collected manually and through Python crawling, and the valid cases were filtered. Specific screening rules are: 1. The zoning of the building master plan is clearly visible, and the compass is clear. 2. There is an independent standard playground runway (except those stacked vertically with other functional zones). 3. Classrooms are arranged unilaterally, and the spacing of the teaching buildings meets the requirements of the specification. 4. The neighbouring roads and land classification are clear. We screened a total of 155 valid samples this time.

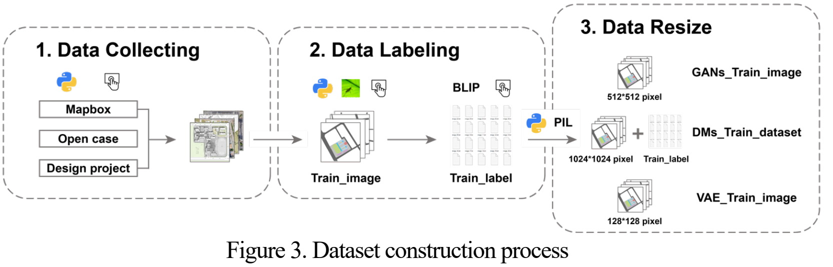
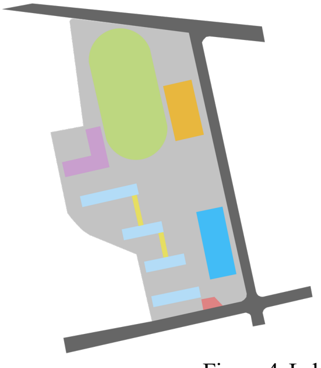


Figure 3. Dataset construction process

Next, we marked and labelled the samples (Figure 3). Here we use the Python Imaging Library (PIL) to process the image, extract the outline of the teaching building through OpenCV threshold segmentation and extraction of land boundary information, and store the coordinate information of the outline points, read the document information in Grasshopper, and use Polyline to connect lines to form a closed polygon, so that the vectorization of the site boundary and building profile is completed. Threshold segmentation of the elements in the general plane is labelled with RGB values (Figure 4), completing the image sample processing. Finally, using the general label model BLIP, ensuring one-to-one correspondence of the labelled image to generated TXT files. The sample label processing is finished,



Color	RGB Value	Functional Sub-areas
	(189, 215, 128)	Playground
	(201, 159, 206)	Restaurant and kitchen
	(151, 206, 229)	Teaching building
	(230, 222, 98)	Corridor
	(101, 101, 101)	Rood
	(194, 194, 194)	Site Location
	(232, 183, 62)	Gymnasium
	(66, 188, 246)	General Classroom
	(223, 133, 131)	Entrance Square

Figure 4. Labelling rule

2.2. MODEL CONSTRUCTION

The idea of GAN confrontation is inspired by Game Theory, where the generator is trained, at the same time, a discriminator is applied to determine whether the input is a real image or a generated image, the two are getting stronger by playing each other in a Zero-sum Game, and a large number of realistic images can be generated by sampling (Figure 5a). The Diffusion Models defines two processes, forward and backward, both of each sample from the real data distribution and gradually add Gaussian noise to the samples, and generate a series of noisy samples, and the noise addition process can be controlled by the variance parameter (Figure 5b). VAE is a variant of auto-encoder, its purpose is to train neural networks in an unsupervised way, including Encoder and Decoder: Encoder process is to compress the original data into low-dimensional vectors, and Decoder is to restore the low-dimensional vectors to the original data (Figure 5c).

In recent years, GAN and VAE have shown great potential in the task of sampling a given data distribution to generate a new one. GAN learns the sampling procedure of complex distributions in an adversarial manner to learned them, while VAE seeks a model that is high likelihood to the distribution of data samples. Although the satisfactory performance of these models in producing high-quality images, they have some limitations of their own. Due to the adversarial of training, GANs tend to suffer from training mode collapse and less distribution coverage, so therefore inferior to SOTA model likelihood-based VAE models in terms of diversity. VAEs can capture more diversity and are often easier to scale and train than GANs, but still fall short in terms of visual sample quality and sampling efficiency (Chai et al., 2023). Recently, diffusion models such as DDPM have emerged as another powerful class of generative models capable of producing high-quality images comparable to GANs (Dhariwal and

Nichol, 2021), with desirable properties such as strong sample diversity, realistic probability distributions, adaptability to different training goals, and ease of scaling. This means that Diffusion Models are well suited for learning complex and diverse data, which motivates us to explore the potential of Diffusion-based generative models for architectural image.

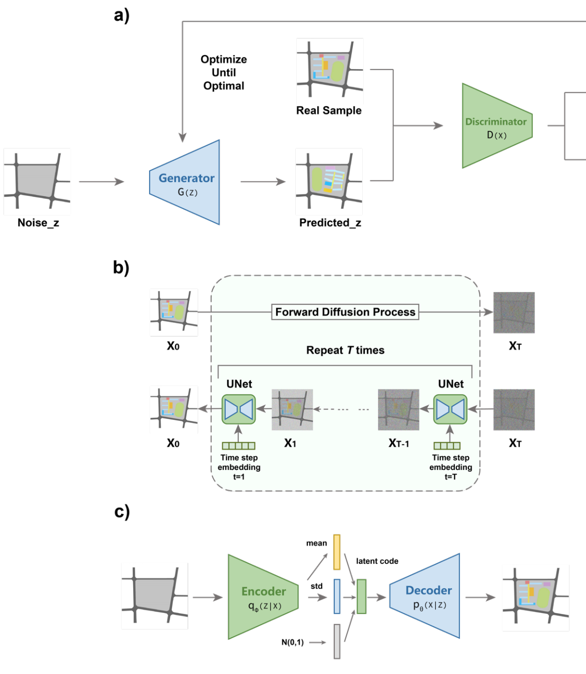


Figure 5. a) Workflow of GAN; b) Workflow of Diffusion Model; c) Workflow of VAE

In this article, three major categories and five subcategories of generative models are selected for evaluation: GANs (pix2pix and cyclegan), DMs (Stable Diffusion, SDXL), and VAE (VQ-VAE), input the real site boundaries and neighbouring road conditions to output the corresponding campus layout images. The models are trained to learn the layout rules from the real campus, and the general plan is automatically generated. All experiments are performed on NVIDIA GeForce RTX 4060 GPU (8 GB).

2.3. EXPERIMENTAL SETTINGS

	Model	sample number	resolution	batch_size	learning rate	epoch
GANs	pix2pix	134	512*512	8	0.0002	220
	cyclegan	134	256*256	8	0.0002	180
DMs	SD	48	512*512	12	0.0001	100
	SDXL	48	1024*1024	10	0.0001	60
VAE	VQ-VAE	155	128*128	32	0.0002	150

Table 1. Implementing details of model training.

Initially, we produce corresponding datasets for each of the five different models based on their requirements for images and text label. For instance, the GANs model necessitates paired input and real images, the diffusion model demands one-to-one correspondence between the input image and text label, and the VAE model's training duration is directly proportional to the pixel count of the image. Consequently, the VAE dataset is divided to optimize computational costs. Upon reviewing the training data from prior studies and conducting numerous experiments, we selected the optimal batch size, learning rate, and number of epochs to enhance training (Table 1).

2.4. GENERATING RESULTS

After training, the neural network can be used to generate a new layout based on the input site boundaries and road conditions (Figure 6). Based on the results, the generated images are compared with ground truth to assess similarity.

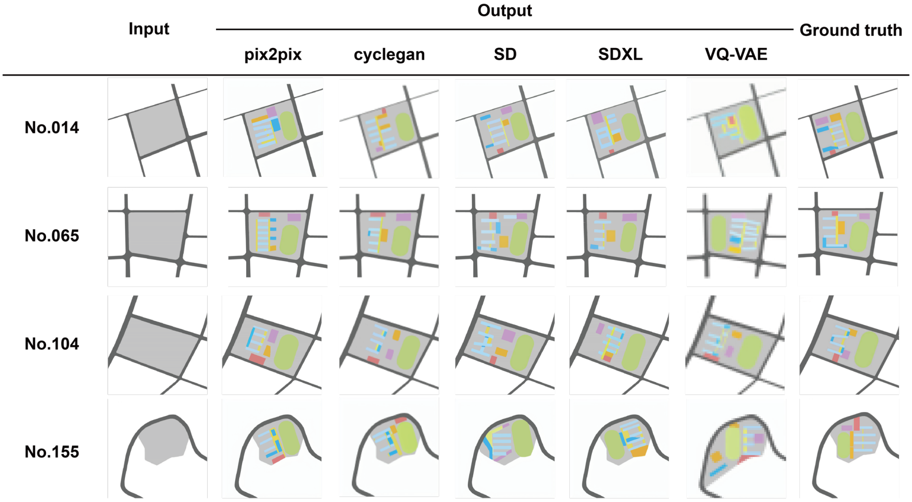


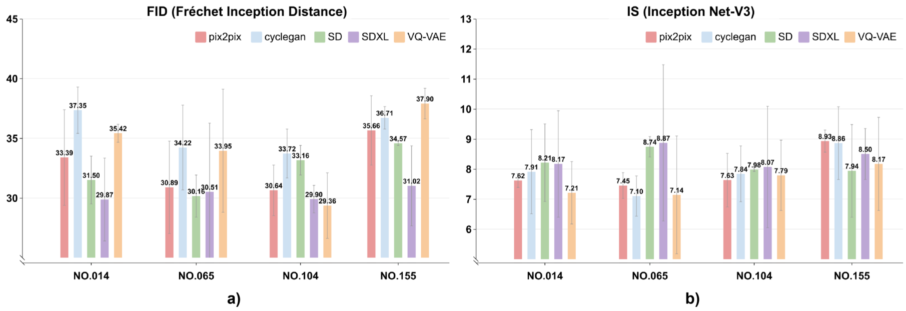
Figure 6. Generated results buy different Generative *Models*

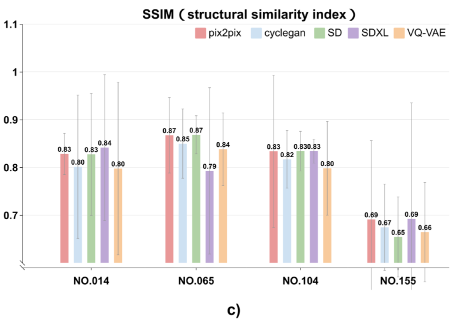
The experimental results show that our trained model somewhat learns the layout rules of the middle school campus. The selection of test samples focuses on rectangular site, with irregular terrain also chosen as an example. In No. 014, 065 and 155, GANs and DMs learned the layout rules of the teaching buildings and the playground well. Regarding functional layout, the results generated mainly by SD (Stable Diffusion) and SDXL are highly consistent with the original plan. The relationship between corridors, general classrooms, and teaching building, as well as between the gymnasium and playground, and the entrance square and the road, are more reasonable. Almost all test samples successfully output the orientation of the classroom buildings and the long-axis direction of the playground. The final result is as expected. In sample No. 155, the overall orientation of the layout is skewed, but it's adapted to terrain. After a preliminary analysis, it appears that the dataset is insufficient, particularly regarding the number of samples of irregularly curved terrain, resulting in the model not learning the relationship between the road, the terrain, and the layout very well.

3. Evaluating discussion

3.1. QUANTITATIVE EVALUATION

First, we evaluated the FID metrics on images generated by each generative model. A lower FID corresponds that generated images are more similar to the real images. Figure 7a shows that the DMs (SD and SDXL) consistently show lower FID scores, in tests Nos. 014, 065, and 155, indicating they often produce images closest to the real image distribution, SDXL in particular performed better. Pix2pix also has lower FID scores in tasks with similar delicate structures, due to the model properties applied to paired images,it has the second-highest generation quality after the DMs in all four tests, shows great potential in architectural image generation.





	Evaluation metrics			Expert study	
	FID	IS	SSIM	Completion	Referenceable
pix2pix	32.65	7.91	0.8549	8.50	8.45
cyclegan	37.99	7.93	0.8396	8.20	8.40
SD	31.60	8.22	0.8273	8.50	8.58
SDXL	32.55	8.40	0.8596	8.25	8.48
VQ-VAE	34.16	7.58	0.8245	7.92	8.40

d)

Figure 7. Evaluation results; a) FID; b) IS; c) SSIM; d) Evaluation metrics and Expert study

In this paper, the Inception V3 model is used to evaluate IS indicators, which has better classification performance, widely used in image recognition and classification tasks (Figure 7b). The high scores indicate that the model generates diverse and reasonably clear images. The SD consistently scores the highest across all samples, indicating it generally produces clearer and more diverse images compared to the other models. The VQ-VAE is generally good at accurately reconstructing or interpolating between existing images, thus, which might be due to less diversity in the images it generates.

The SSIM measures the similarity between two images and takes values in the range of [0,1]. The larger the value, the closer in structure to ground truth. Figure 7c shows the images generated by each model and the ground truth input function. Except for test NO. 155, pix2pix consistently performed the best across all samples in other datasets. The DMs showed strong performance but were not always superior to other models, possibly due to the samples' specificity or the model's training.

3.2. QUALITATIVE COMPARISONS

We distributed questionnaires through online and offline methods, targeting mainly architects, architectural faculty and students, designers, and other related practitioners. A total of 67 pieces of valid data were returned, each subject compared four pairs of layouts, sampled from five generative models that generate images, with the ground truth.and the statistical results are shown in Figure 7d. Expert Study with a range of [0,10], calculating the mean of each score and tallying. The metrics 'Completion' and 'Referenceable' are subjective measures evaluated by professionals proficient in the relevant field for the generated images.

Both pix2pix and SD scored 8.50 in Completion, the highest in the category. This indicates that they were able to generate fully realised images for the assigned tasks. Pix2pix is particularly good at processing detailed structural information due to its paired training data, making it well suited for targeted explicit tasks. SD's high score indicates its ability to generate complete images by understanding and reproducing complex image features. SD had the highest reference score of 8.58, indicating that experts believed it generated the most accurate and reliable images relative to real ones. SDXL followed closely at 8.48, indicating that it also generated high-quality, reference-worthy images.

4. Conclusion

In this paper, we discuss the training and evaluation of architectural image generation models, different metrics lead to different trade-offs, and different evaluation scores will benefit different models. Therefore, it is important to train and evaluate according to the specific situation of the target application. In addition, when selecting models for architectural image generation tasks, we should be careful not to take good performance in one task as evidence of good performance in another application.

According to the evaluation results of the campus master plan layout generation experiment, in terms of computational criteria, pix2pix generates images closer to the ground truth, while the latest diffusion models SD and SDXL are more enlightening and diverse. Expert indicators show that the images generated by pix2pix and SD models are more complete, while SDXL is more informative and innovative. In summary, within the model evaluation framework proposed by this paper, the diffusion model excels, demonstrating superior computational performance and receiving high expert evaluations. This indicates that the diffusion model is adept at executing the building floor plan generation task, thereby informing future collective building layout planning. Although our proposed evaluation framework is impressive compared with existing architectural layout generation methods, it still has some limitations. For example, when architects want to apply the framework to filter the generated images as their main reference, they need to deploy the evaluation system manually. This greatly increases the learning cost and is not conducive to the promotion of the method.

References

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. *Advances in neural information processing systems, 27.*

Ho, J., Jain, A., & Abbeel, P. (2020). Denoising diffusion probabilistic models. *Advances in neural information processing systems, 33,* 6840-6851.

Kingma, D. P., & Welling, M. (2013). Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114.*

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

Dhariwal, P., & Nichol, A. (2021). Diffusion models beat gans on image synthesis. Advances in neural information processing systems, 34, 8780-8794.

Guida, G. E. O. R. G. E. (2023, March). Multimodal Architecture: Applications of Language in a Machine Learning Aided Design Process. In *HUMAN-CENTRIC-Proceedings of the 28th CAADRIA Conference. Ahmedabad* (pp. 18-24).

Huang, W., & Zheng, H. (2018, October). Architectural drawings recognition and generation through machine learning. In *Proceedings of the 38th annual conference of the*

Chaillou, S. (2020, September). Archigan: Artificial intelligence x architecture. In *Architectural Intelligence: Selected Papers from the 1st International Conference on Computational Design and Robotic Fabrication (CDRF 2019)* (pp. 117-127). Singapore: Springer Nature Singapore. *association for computer aided design in architecture, Mexico City, Mexico* (pp. 18-20).

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.

Ali, A. K., & Lee, O. J. (2023). Facade style mixing using artificial intelligence for urban infill. *Architecture, 3*(2), 258-269.

Wang, B., Zhang, S., Zhang, J., & Cai, Z. (2023). Architectural style classification based on CNN and channel–spatial attention. *Signal, Image and Video Processing, 17*(1), 99-107.

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 *Proceedings of the 2021 DigitalFUTURES: The 3rd International Conference on Computational Design and Robotic Fabrication (CDRF 2021) 3* (pp. 55-68). Springer Singapore.

Zhang, H. (2019). 3D model generation on architectural plan and section training through machine learning. *Technologies, 7*(4), 82.

Pang, H. E., & Biljecki, F. (2022). 3D building reconstruction from single street view images using deep learning. *International Journal of Applied Earth Observation and Geoinformation, 112,* 102859.

YOUSIF, S., & BOLOJAN, D. (2022). Deep learning-based surrogate modeling for performance-driven generative design systems. In *Proc of the 27th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA), CAADRIA* (pp. 363-372).

Jia, M. (2021). Daylight prediction using Gan: General workflow, tool development and case study on Manhattan, New York.

Chai, S., Zhuang, L., & Yan, F. (2023). LayoutDM: Transformer-based Diffusion Model for Layout Generation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 18349-18358).

CAN GENERATIVE AI MODELS COUNT?

Finetuning Stable Diffusion for Architecture Image Generation with Designated Floor Numbers Using a Small Dataset

WEISHUN XU[1], MINGMING LI[2] and XUYOU YANG[3]
[1] *Zhejiang University College of Engineering and Architecture*
[2] *Zhejiang University College of Computer Science and Technology*
[3] *1DesignLab*
[1]*xuweishun@zju.edu.cn, 0009-0000-4489-7858*
[2]*limingming@zju.edu.cn, 0000-0001-8916-6485*
[3]*xuyou.yang.92@gmail.com, 0000-0002-6294-378X*

Abstract. Despite the increasing popularity of off-the-shelf text-to-image generative artificial intelligence models in early-stage architectural design practices, general-purpose models are challenged in domain-specific tasks such as generating buildings with the correct number of floors. We hypothesise that this problem is mainly caused by the lack of floor number information in standard training sets. To overcome the often-dodged problem in creating a text-image pair dataset large enough for finetuning the original model in design research, we propose to use BLIP method for both understanding and generation based automated labelling and captioning with online images. A small dataset of 25,172 text-image pairs created with this method is used to finetune an off-the-shelf Stable Diffusion model for 10 epochs with affordable computing power. Compared to the base model with a less than 20% chance to generate the correct number of floors, the finetuned model has an over 50% overall chance for correct floor number and 87.3% change to control the floor count discrepancy within 1 storey.

Keywords. Text-to-Image Generation, Model Finetuning, Stable Diffusion, Automated Labelling.

1. Introduction

Architectural design practices have seen a recent rise in the use of text-to-image (T2I) generative AI (GAI) models due to their improved accessibility and performance. Commercially available cutting-edge T2I GAI tools such as DALL-E-2 (Ramesh et al., 2022), MidJourney, and Stable Diffusion (Rombach et al., 2022) are based on large diffusion models that can generate high-quality images that provide design inspirations in only minutes and thus largely accelerate the design process.

Typically, T2I GAIs have been adopted in early design phases such as ideation

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 89-98. © 2024 and published by the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA), Hong Kong.

(Stigsen et al., 2023), digital sketching (Ploennigs & Berger, 2023) and style exploration (Chen et al., 2023). However, off-the-shelf general models are less efficient in domain-specific tasks such as architectural design, creating inconsistency in communication through text and difficulty in interpreting the results (Turchi et al., 2023). Domain-specific tasks often require extra field knowledge and respective descriptive language, leading to adaptation efforts of general models such as prompt-engineering and context-specific tuning to gain more detailed control over specific aspects of the generative process.

One of the persisting issues in adopting GAI models pre-trained on general purpose in early-stage architectural design is the generation of images with the correct number of floors. For example, when the prompt describes a building of three floors, very often the results generated by the off-the-shelf T2I GAI tools reflect a building with a wrong number of floors (Figure 1). Because prevailing building codes often associate the number of floors or building height to building typology, this problem limits the potential of application of general GAI models in early conceptualisation.

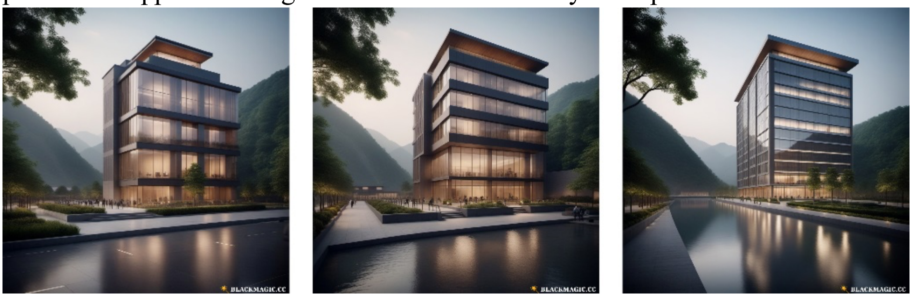


Figure 1. From left to right: images generated by Stable Diffusion with prompts as "Rendering of a Modernism office building of x storeys in mountains in China" where x is 3, 4, and 7 respectively.

To enhance the practicality of using large pre-trained T2I GAI models as an effective early-stage conceptual design tool for architecture with specific requirements, this research proposes a finetuning method using a small dataset to allow for professional specifications by generating images that match with the prompt more accurately, taking the number of floors as a case study.

2. Background

Since changing the number of floors in the prompts for off-the-shelf models does change the generated floor count only with unacceptable accuracy, we hypothesise that the poor performance is caused by the lack of floor count information in the model's training data. Hence, injecting text-image pairs with correct information to the model may be a solution. Therefore, a literature review is conducted in optimised methods for T2I GAI models to learn new knowledge for design purposes, and how training data can be prepared respectively.

2.1. ADAPTING GAI FOR EARLY-STAGE DESIGN

To better apply large pre-trained GAI models in specific downstream tasks, there have been several common methods. Overall, existing methods that require smaller training data for design applications are often used in styling or personalisation while leaving most of the original model intact. For example, prompt engineering is a process to create and optimise input text to instruct the models to perform specific tasks, which can be used in better extracting knowledge from a GAI model to help design ideation (Deshpande, 2023). However, this method relies on the capability of the original model without injecting new knowledge. Another method to bridge these models with downstream tasks is Low-Rank Adaptation, or LoRA (Hu et al., 2021), which freezes the pretrained model weights and injects trainable rank decomposition matrices, greatly decreasing the number of trainable parameters. LoRA can be useful in styling, applying texture and rendering, but with the original model left untouched, its ability to inject structural knowledge is limited (Kuru, 2023). Specifically targeted at diffusion models, Dreambooth allows users to inject custom objects with a few images as additional training data to an existing class (Ruiz et al., 2023). Text-inversion uses a pseudo word to embed a set of vectors for highly personalised results learned from only a few input images, but mostly of the same object (Gal et al., 2022).

Yet domain-specific problems such as the counting of floors require that the original model generate the correct visual semantics with a new class of text inputs. This suggests finetuning the entire model, an approach of transfer learning, in which the weights of a pre-trained model are trained on new text-image data pairs (Goodfellow et al., 2016). Previous work in architectural design has successfully associated structured vocabulary in five classifications to architectural forms by finetuning U-Net of Stable-Diffusion v1.4 with 1001 manually labelled images (Kim, 2023). Yet the process of such data curation is demanding, requiring both labour and professional consistency in reading the form and tedious preparation of graphically similar images all in isometric views.

2.2. AUTOMATED DATA LABELLING

As can be observed from the precedents above, one of the main reasons why finetuning a T2I GAI model through re-training is often avoided in domain-specific tasks is the lack of structured data, as re-training with trivial datasets often lead to severe loss of previous knowledge in a large model (Li et al., 2022). Since accessible online images often do not have associated captions with domain-specific knowledge, researches then must rely on either open datasets that are manually curated for general purposes, or find ways work with noisy data with expensive post-processing steps (Jia et al., 2021). However, recent developments in scaling pre-trained representations for T2I GAI models with noisy data have allowed us to consider a methodological framework with readily available resources.

One of the commonly referred to state-of-the-art visual-language pretraining frameworks, BLIP, can perform a wide range of downstream tasks that are both understanding-based and generation-based (Li et al., 2022). Pretrained with datasets generated with CapFilt, BLIP has strong abilities in understanding images and generate texts that accurately match with the given images. For our task of floor counting with T2I GAI, the structure of standard training data needs to involve a text-image pair which involve both an understanding of an original image in the form of the number of floors, and a generated natural language caption. Therefore, the BLIP model can be a viable option based on our review.

3. Method

Since domain-specific tasks in architectural design can vary from case to case performed by entities with varied computing power and labour, our goal of this early-stage research is to establish a finetuning framework with publicly available data, automated labelling and captioning with relatively small dataset to enhance its accessibility.

3.1. IMAGE DATA ACQUISITION, FILTERING AND LABELLING

3.1.1. Data Acquisition

To label images of architecture for the number of floors, source images should be exterior photos or renderings that cover the entire height of the building, and preferably the entire building volume to learn about the building typology as well. In addition, other than image resolution and clarity, photos of buildings in the dataset should cover a variety of styles, which helps the model generalise better, by covering a wider range of other descriptive language in the prompt.

The 'Exterior Photography for Architects' image label on one of the widely accessed websites for architectural design, ArchDaily (ArchDaily, n.d.), provides an ideal source for our data acquisition demand. Despite some misplacements of interior, urban design and infrastructural projects, the label collects high-quality photos of building exteriors of designs with varied styles, including materials and typology. We acquired a raw dataset total of 33,269 images from this label for filtering and processing.

3.1.2. Data Filtering

As the raw dataset still contained images from which the number of floors cannot be counted, directly applying such data for caption labelling may adversely affect the performance of our subsequent model finetuning, causing the model to learn irrelevant features. Empirically, such photos are often associated with wrong project types or views. Therefore, we needed to filter such photos based on automated recognition of image subjects.

Conventional image filtering methods often use pre-trained image classification models or object detection models. In our case, such methods may not be flexible enough considering the complexity of the background in architectural photos. In addition, since our filter standard was empirically based on observation, we needed a method that adapts to varied needs described in natural language.

We used a zero-shot classification method based on BLIP to categorise and filter images. In our method, the BLIP model and its pre-processor were loaded to extract image and text features respectively, and then the similarity between them was calculated using the Softmax function. This way, the subject of the image can be converted into probabilities of pre-determined project types based on our initial empirical review of the dataset, with the highest probability category extracted as the final type of the subject.

Based on preliminary visual analysis, we determined that subjects which need to be filtered out belong to "bridge", "city", "public square", " "top view" and "indoor space" categories respectively. As can be seen, this process was both targeted at unwanted perspectives and subject types. In the end, we finalised the categories of classification as "building", "bridge", "city", "public square", "top view", and "indoor space". Only images with a probability of being a building higher than 60% were retained. Through this method, we filtered out approximately 8k non-building images from the original data, and the final dataset size was 25172 images for labelling (Figure 2).

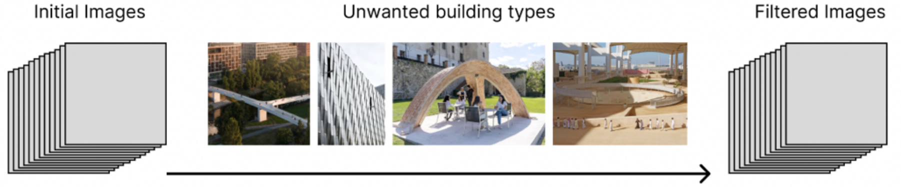


Figure 2. Examples of filtered image types

3.1.3. Data Labelling

To create the text-image pair required for finetuning, the caption we associate with each image needs to both include our main training goal which is the number of floors, and some natural language description about the general scene to respond to other key words that may appear in the prompts.

Based on BLIP's capability in understanding and generation-based tasks, this process could be automated. We used the BLIP-VQA (visual question answering) model to get the number of building floors by asking "How many floors does this building have?". Then, we used the BLIP caption base model to generate an overall description of the image. As shown in Figure 3, for the image "192.jpg", the final output for the label was "a two-storey building. a white building with a staircase going up to it", which contains the number of floors and a description of the built environment. "20.jpg" is labelled as "a 10-storey building. a tall building with plants on the balconies", highlighting the number of floors and architectural details.

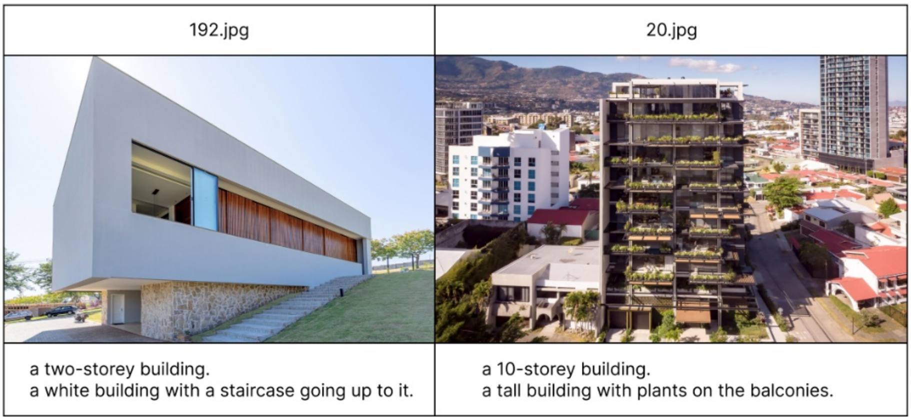


Figure 3. Examples of understanding and generation based captioning tasks by BLIP.

This dual annotation method played a key role in our training set, providing the model with comprehensive semantic information while simultaneously accentuating the number of floors. Such labels help improve the model's understanding of the building structure and appearance, enhancing the model's generalisation capability. In the finetuning stage, these labels make it easier for the model to adjust weights to meet specific task requirements that utilise domain-specific knowledge, while reducing the ambiguity in the model's understanding of labels.

3.2. MODEL FINETUNING

As our research proposal targets at the practicality of injecting domain-specific knowledge for architectural practices, we prioritised accessible off-the-shelf models that require less computation power for finetuning.

Diffusion models are a class of deep generative models often used in computer vision tasks including image generation based on two stages: a forward diffusion stage and a reverse diffusion stage. Compared to conventional diffusion models, Latent Diffusion Models (LDMs) execute the diffusion process on the latent instead of pixel space and subsequently require much less computation resource. Stable Diffusion is one of the latest versions of LDMs for T2I tasks accessible under the CreativeML Open RAIL-M license, and the Stable Diffusion v1-4 model was chosen for our finetuning process.

With the 25172 text-image pairs, the finetuning was executed following Stable Diffusion's official finetuning documentation on a single NVIDIA GeForce RTX 4090 graphics card for 10 epochs with a learning rate of 1e-5. The finetuning process took a total of 71 hours and 31 minutes. To further verify the efficiency of our proposed methodological framework, we also made an alternative finetuned model with the same data but for only 1 epoch as comparison.

4. Results

4.1. VALIDATION METHOD

To test the performance of our finetuned model against the original model and empirically evaluate the difference between training epochs, we designed an evaluation process based on 1000 512px-by-768px image sets generated from the original Stable Diffusion model (referred to as the base model in the following text), the finetuned model with 1 epoch learning (1epoch), and the final finetuned model with 10 epochs (10epoch) with the same seeds and prompts.

For each generation across all three models, a seed was randomly chosen images from our training data to generate the latent space of the models. Then, the prompt was automatically generated with randomly picked parameters which can describe the number of building floors, building style, and building type. The range of floor numbers was from 1 to 7, and the styles included "modern", "contemporary", "traditional", "industrial", "Art Deco", and "minimalist". Building types were chosen from "single-family house", "apartment building", "villa", "shopping mall", "retail store", "restaurant", "café", "office building", "museum", "art gallery", "library", "concert hall", "school building", "hospital", "historical building", and "church". The prompt then was generated as a natural language text with standard descriptive structure, such as "A modern 2-storey single-family house". We also added key phrases to each prompt, namely "whole building" to designate the subject of the generated image, "human perspective" to set the view angle, and "façade with well-defined boundaries between each storey" to help generate subjects that are easy for humans to count the number of floors to quantify the performance. Due to our prompt generation method, resultant buildings in the generated images mostly have either a single volume or several volumes of similar heights, which affects our validation method described below.

We then manually counted the number of floors on each generated image for different models respectively. If the subject has a changing number of floors in the image, only the largest portion of the building was taken into the final count. For instance, for the image on the left in Figure 4, the building changes from 9 floors to 8 and a floor number of 8 was taken. In the image in the middle, the floor number of 2 was taken. In rare cases where the façades generated could not clearly separate the floors visually, other evidence such as vertically stacked windows or openings will be used as a clue for estimation, such as in the image on the right in Figure 4. All counts were executed by graduate students who received their undergraduate training in architecture to ensure that industry knowledge was involved.

Figure 4. Examples of floor count rules where professional judgement was applied.

4.2. OVERALL MODEL PERFORMANCE

We treated the number of floors given in the prompt as the ground truth, and calculated the overall percentage of correctness when the number of floors in the generated image matches the ground truth. We also calculated the average difference between the number of floors generated and the ground truth. To evaluate the models' tolerance, we noted the percentages of images when the floor count differs from ground truth by no more than 1 floor and by no more than 2 floors. The results for each model are shown in Table 1.

Model	**Correct (%)**	**Avg. floor difference**	**Differ within 1 storey (%)**	**Differ within 2 storeys (%)**
base	19.4	1.587	55	79.1
1epoch	41.3	1.048	69.8	87.3
10epoch	51.2	0.663	87.3	95.8

Table 1. Overall model performance

As the result shows, finetuned models exhibit majorly improved performance in the number of floors. 19.4% of the images generated by the base model have the correct number of floors, while the percentage of correctness is increased to 41.3% after 1 epoch of finetuning, and subsequently to 51.2% after 10 epochs. Improved performance is also observed in controlling the floor number gap between generated images and the prompt. The average floor difference is reduced from 1.587 in the base model to 0.663 in the finetuned 10epoch model. In addition, 10epoch has a 95.8% chance of containing the floor difference within 2 storeys compared to 79.1% for the base model. A visual comparison of some generated results is shown in Figure 5.

Figure 5. 4 sets of visual comparison between images generated by base (left) and 10epoch (right)

4.3. PERFORMANCE BY NUMBER OF FLOORS

To further analyse applicable use scenarios for the models, we compared the difference between 10epoch and base model for their performance by floor count. The result of the comparison is shown in Table 2 below:

Floor Count	**Correct (%)**		**Differ within 1 storey (%)**		**Differ within 2 storeys (%)**	
	base	**10epoch**	**base**	**10epoch**	**base**	**10epoch**
1	6.6	44.1	36.8	90.4	71.3	95.6
2	28.6	73.5	73.6	92.9	91.4	100
3	32.5	58.4	81.1	92.9	94.8	98.1
4	24.6	54.6	79.2	83.8	94.6	96.2
5	12.5	46.7	48	85.5	84.9	96.1
6	9	34.6	40.4	81.4	75	94.2
7	7.6	47.7	33.3	84.1	65.2	90.2

Table 2. Base model and 10epoch performance comparison by floor counts

As demonstrated, with the number of floor accuracy lower than 30% almost across all floor counts, the base model faces general challenges but particularly in single storey buildings and buildings higher than 5 floors. The model also struggles to contain the difference in floor counts within 1 storey, which renders its usage as an ideation source in early-stage design limited, because buildings of 1 or 6 to 7 levels are often more sensitive to floor number changes due to planning and fire code regulations.

Meanwhile, the 10epoch model exhibited a performative increase by 30% in almost every floor count category. The table demonstrates that despite the overall enhanced performance, the finetuned model has a similar pattern of performing better in generating buildings of 2 to 4 storeys. However, the finetuned model shows a more stable and elevated performance of containing the floor count difference within 1 storey. This feature greatly improves the model's practicality in early-stage designs.

5. Discussion

Our research demonstrates the feasibility of using a small dataset (25172 images), automated labelling and limited computing power, it is possible to finetune a publicly accessible T2I GAI model for majorly improved performance in early-stage architectural design tasks, such as generating ideation images with the correct number of floors, which involve domain-specific knowledge. The result also partially supports our hypothesis, which is that one critical reason for current general purpose T2I GAI models to fail basic-level early-stage architectural design tasks is the lack of well-structured training data that includes building information.

Meanwhile, we understand that as early-stage research, our methodological framework has its limitations. Firstly, our evaluation for both the base model and the finetuned models does not necessarily reflect their actual performance in practice. The automated prompt generation script can create realistically improbable descriptions such as "a minimalist 1-storey hospital building" or "a contemporary 5-storey historical building". Therefore, both models, before and after finetuning, may perform better in actual architectural design practices. The other limitation caused by the prompt generation script is that it is unable to describe complex buildings with varied volume heights, which we intend to address in future research.

In addition, we did not quantitatively validate the accuracy of the automatic labelling process, which may have led to less optimised training results. However, having observed that even a small dataset with automated labelling can dramatically increase the performance of the pre-trained model, we think that more could be done with well-structured captions that contain more comprehensive building information. Further research should expand on matching what can be visually learned and what should early-stage design address in generating architectural images for designers, so that training datasets for architectural AI can be crowd-sourced.

Acknowledgements

This research is supported by the National Natural Science Foundation of China under Grant 52208036, China's Research and Development Project of Ministry of Housing and Urban-Rural Development under Grant 2022-K-004, and Center for Balance Architecture at Zhejiang University.

References

Chen, J., Wang, D., Shao, Z., Zhang, X., Ruan, M., Li, H., & Li, J. (2023). Using Artificial Intelligence to Generate Master-Quality Architectural Designs from Text Descriptions. *Buildings,* 13(9), Article 9. https://doi.org/10.3390/buildings13092285

Deshpande, R. (2023). Generative Pre-Trained Transformers for 15-Minute City Design. In *HUMAN-CENTRIC - Proceedings of the 28th CAADRIA Conference,* (pp. 595–604). https://doi.org/10.52842/conf.caadria.2023.1.595

Archdaily. (n.d.). ArchDaily. Retrieved August 28, 2023, from https://www.archdaily.com/search/images

Gal, R., Alaluf, Y., Atzmon, Y., Patashnik, O., Bermano, A. H., Chechik, G., & Cohen-Or, D. (2022). *An Image is Worth One Word: Personalizing Text-to-Image Generation using Textual Inversion* (arXiv:2208.01618). arXiv. http://arxiv.org/abs/2208.01618

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning.* MIT press.

Hu, E. J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., Wang, L., & Chen, W. (2021). *LoRA: Low-Rank Adaptation of Large Language Models* (arXiv:2106.09685). arXiv. http://arxiv.org/abs/2106.09685

Jia, C., Yang, Y., Xia, Y., Chen, Y.-T., Parekh, Z., Pham, H., Le, Q., Sung, Y.-H., Li, Z., & Duerig, T. (2021). Scaling Up Visual and Vision-Language Representation Learning With Noisy Text Supervision. *Proceedings of the 38th International Conference on Machine Learning,* 4904–4916. https://proceedings.mlr.press/v139/jia21b.html

Kim, F. C. (2023). Text2Form Diffusion: Framework for learning curated architectural vocabulary. *Digital Design Reconsidered - Proceedings of the 41st Conference on Education and Research in Computer Aided Architectural Design in Europe (eCAADe 2023)*, (Vol 1, pp. 79–88). https://doi.org/10.52842/conf.ecaade.2023.1.079

Kuru, J. (2023). *Training Non-Typical Character Models for Stable Diffusion Utilizing Open Sources AIS* [Honor Bachelor Thesis, University of Arizona]. https://repository.arizona.edu/handle/10150/668639

Li, J., Li, D., Xiong, C., & Hoi, S. (2022). BLIP: Bootstrapping Language-Image Pre-training for Unified Vision-Language Understanding and Generation. *Proceedings of the 39th International Conference on Machine Learning,* 12888–12900. https://proceedings.mlr.press/v162/li22n.html

Ploennigs, J., & Berger, M. (2023). AI art in architecture. *AI in Civil Engineering,* 2(1), 8. https://doi.org/10.1007/s43503-023-00018-y

Ramesh, A., Dhariwal, P., Nichol, A., Chu, C., & Chen, M. (2022). *Hierarchical Text-Conditional Image Generation with CLIP Latents* (arXiv:2204.06125). arXiv. https://doi.org/10.48550/arXiv.2204.06125

Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). High-Resolution Image Synthesis with Latent Diffusion Models. *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR),* 10674–10685. https://doi.org/10.1109/CVPR52688.2022.01042

Ruiz, N., Li, Y., Jampani, V., Pritch, Y., Rubinstein, M., & Aberman, K. (2023). DreamBooth: Fine Tuning Text-to-Image Diffusion Models for Subject-Driven Generation. *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR),* 22500–22510. https://doi.org/10.1109/CVPR52729.2023.02155

Stigsen, M. B., Moisi, A., Rasoulzadeh, S., Schinegger, K., & Rutzinger, S. (2023). AI Diffusion as Design Vocabulary—Investigating the use of AI image generation in early architectural design and education. *Digital Design Reconsidered—Proceedings of the 41st Conference on Education and Research in Computer Aided Architectural Design in Europe* (Vol. 2, pp. 587–596). https://doi.org/10.52842/conf.ecaade.2023.2.587

Turchi, T., Carta, S., Ambrosini, L., & Malizia, A. (2023). Human-AI Co-creation: Evaluating the Impact of Large-Scale Text-to-Image Generative Models on the Creative Process. In L. D. Spano, A. Schmidt, C. Santoro, & S. Stumpf (Eds.), *End-User Development* (pp. 35–51). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-34433-6_3

CHATDESIGN: BOOTSTRAPPING GENERATIVE FLOOR PLAN DESIGN WITH PRE-TRAINED LARGE LANGUAGE MODELS

JINMIN LI[1], YILU LUO[2], SHUAI LU[3], JINGYUN ZHANG[4], JUN WANG[5], RIZEN GUO[6] and SHAOMING WANG[7]
[1,2,3]Shenzhen International Graduate School, Tsinghua University
[1,4,5,6,7]Wechat Lab 33, Tencent
[1]ljm22@mails.tsinghua.edu.cn , 0009-0000-7098-2589
[2]luo-yl22@mails.tsinghua.edu.cn, 0009-0005-5448-9555
[3]shuai.lu@sz.tsinghua.edu.cn, 0000-0003-3772-1613
[4]zhang304973926@gmail.com
[5]earljwang@gmail.com
[6]rizenguo@tencent.com
[7]mangosmwang@tencent.com

Abstract. Large language models (LLMs) have achieved remarkable success in various domains, revolutionizing tasks such as language translation, text generation, and question-answering. However, generating floor plan designs poses a unique challenge that demands the fulfilment of intricate spatial and relational constraints. In this paper, we propose ChatDesign, an innovative approach that leverages the power of pre-trained LLMs to generate floor plan designs from natural language descriptions, while incorporating iterative modifications based on user interaction. By processing user input text through a pre-trained LLM and utilizing a decoder, we can generate regression parameters and floor plans that are precisely tailored to satisfy the specific needs of the user. Our approach incorporates an iterative refinement process, optimizing the model output by considering the input text and previous results. Throughout these interactions, we employ many strategic techniques to ensure the generated design images align precisely with the user's requirements. The proposed approach is extensively evaluated through rigorous experiments, including user studies, demonstrating its feasibility and efficacy. The empirical results consistently demonstrate the superiority of our method over existing approaches, showcasing its ability to generate floor plans that rival those created by human designer. Our code will be available at https://github.com/THU-Kingmin/ChatDesign.

Keywords. floor plan generation, large language models, user interactions, automatic design, deep learning, pre-train models

1. Introduction

Large language models (LLMs) have made significant strides in various domains, revolutionizing tasks such as language translation, text generation, and question-answering. These models, trained on vast amounts of textual data, have demonstrated the ability to comprehend and generate human-like language. In this paper, we propose ChatDesign, an innovative approach that harnesses the power of pre-trained LLMs to generate floor plan designs from natural language descriptions.

Generating floor plan designs presents a unique challenge that requires the fulfilment of intricate spatial and relational constraints. While existing methods have made progress in considering specific constraints such as room types, adjacencies, and boundaries, they often rely on predefined templates or manual parsing techniques. These approaches limit flexibility and adaptability, hindering the generation of floor plans that precisely meet the user's needs.

To address these limitations, we introduce ChatDesign, which incorporates iterative modifications based on user interaction. Our approach involves processing user input text through a pre-trained LLM and utilizing a decoder to generate regression parameters and floor plans tailored to the user's specific requirements. We employ an iterative refinement process that optimizes the model output by considering the input text and previous results, ensuring the generated design images align precisely with the user's needs.

Notable models like CogView (Ding et al., 2021) and Imagen (Saharia et al., 2022) have showcased the potential of pre-trained LLMs in transforming input data into meaningful representations, pushing the boundaries of AI-powered generation tasks. Building upon this progress, our proposed ChatDesign approach harnesses the power of pre-trained LLMs to generate floor plan designs. By doing so, we address the limitations of existing methods and highlight the potential of LLMs in automating and enhancing the floor plan design process.

In summary, our main contributions are as follows:

- We propose a novel method for the floor plan generation task, utilizing a pre-trained large language model to iteratively optimize the output.
- We introduce an innovative scheme for interactive floor plan generation, empowering users to interactively modify the generated floor plans in line with their changing needs.
- We implement a variety of interaction modalities, including text input, audio instructions, mouse clicks, and drag-and-drop operations.
- The empirical results consistently demonstrate the superiority of our method over existing approaches, showcasing its ability to generate floor plans that rival those created by human designer.

2. Related Work

2.1. FLOOR PLAN GENERATION

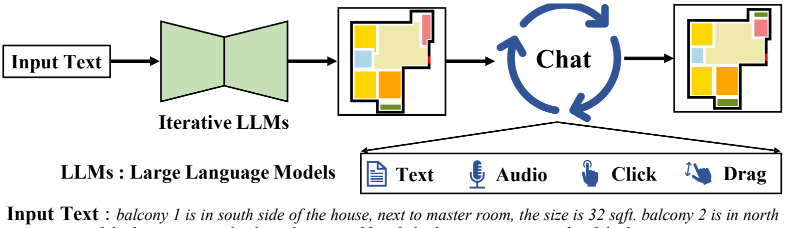
Several methods have been proposed for automatic floor plan design, with most of them taking into account specific constraints such as room types, adjacencies, and boundaries. For instance, Wu et al. (2019) introduced a CNN-based approach that utilizes boundary images as a constraint to determine the location of different rooms. Chen et al. (2020) provided a method where a small set of template-based artificial verbal commands is manually parsed into scene graphs to guide the generation process. Leng et al. (2023) contributed a new dataset called Tell2Design, which consists of over 80,000 floor plan designs accompanied by natural language instructions. Additionally, a benchmark model has been developed to evaluate the performance of different techniques in this field.

2.2. PRE-TRAIN LARGE MODELS

Pre-trained large models (Zhang et al., 2021, Liu et al., 2021) have made significant advancements in various domains. CogView (Ding et al., 2021) utilizes a pre-trained VQ-VAE to transform a target image into a sequence of image tokens. These image tokens are then combined with text tokens and input into a Transformer decoder to generate an image. Imagen (Saharia et al., 2022), on the other hand, is an advanced text-to-image generation model that leverages a large language model T5 (Raffel et al., 2020) for text comprehension and a diffusion model for generating high-fidelity images. Furthermore, Tell2Design (Leng et al., 2023) introduces a new dataset called T2D, which consists of over 80,000 floor plan designs paired with natural language instructions. This dataset enables the development of text-to-floor plan generation models using the powerful T5 language model, known for its exceptional language understanding capabilities. These models have demonstrated impressive results in their respective domains, showcasing the potential of pre-trained large models in pushing the boundaries of AI-powered generation tasks.

3. Methodology

3.1. OVERVIEW



Input Text : *balcony 1 is in south side of the house, next to master room, the size is 32 sqft. balcony 2 is in north east corner of the house, next to kitchen, the size is 32 sqft. bathroom is in west side of the house, next to common room 2 and living room, the size is 48 sqft. common room 1 is in south west corner of the house, next to master room and bathroom, the size is 140 sqft. common room 2 is in north west corner of the house, next to living room and bathroom, the size is 120 sqft. kitchen is in north east corner of the house, next to living room and balcony, the size is 64 sqft. living room is in north side of the house, next to common room 2, bathroom and kitchen room, the size is 220 sqft. master room is in south side of the house, next to common room 1 and balcony, the size is 120 sqft.*

Figure 1. An overview of our proposed ChatDesign.

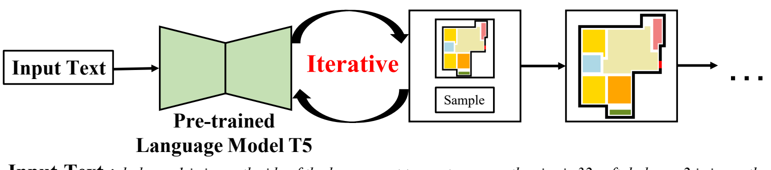
As shown in Figure 1, the methodology of ChatDesign involves an iterative process that leverages pre-trained LLMs to generate floor plan designs based on user input. The overall process can be summarized as follows:

- Input Text and Initial Floor Plan Generation: The user provides a human-described textual input describing their floor plan requirements. This input is processed by the Iterative LLMs, which generate an initial floor plan as a preliminary design. The Iterative LLMs utilize the power of pre-trained language models to comprehend and transform the textual input into a visual representation.
- Interactive Design Phase: Building upon the initial floor plan, the user enters an interactive design phase where they can further refine the design based on their evolving requirements. Various interactive modalities such as text input, voice commands, mouse clicks, and drag-and-drop interactions are supported. These interactions enable specific actions such as precise positioning of rooms, moving them within the layout, resizing rooms according to user preferences, as well as adding or removing rooms.
- Template-Based Interactions: In addition to the textual input, users also have the option to skip the initial step and directly select a template floor plan. Based on this template, users can propose new interactive design requirements, allowing them to modify and adapt the template to suit their specific needs.

The iterative nature of ChatDesign allows users to progressively refine the generated floor plans based on their evolving preferences and requirements. By incorporating interactive design features and providing flexibility in both initial input and template-based approaches, the methodology empowers users to actively participate in the floor plan generation process and achieve designs that align closely with their vision. The entire model is trained in a way and the objective function of the training follows the settings of baseline Tell2Design (Leng et al., 2023).

By doing so, ChatDesign enhances user engagement and ensures personalized output. It transforms the traditional design process into a dynamic interaction, understanding user preferences and delivering tailored results. This user-centric approach facilitates continuous improvement of the model, increasing its efficiency and accuracy over time.

3.2. ITERATIVE GENERATION



Input Text : *balcony 1 is in south side of the house, next to master room, the size is 32 sqft. balcony 2 is in north east corner of the house, next to kitchen, the size is 32 sqft. bathroom is in west side of the house, next to common room 2 and living room, the size is 48 sqft. common room 1 is in south west corner of the house, next to master room and bathroom, the size is 140 sqft. common room 2 is in north west corner of the house, next to living room and bathroom, the size is 120 sqft. kitchen is in north east corner of the house, next to living room and balcony, the size is 64 sqft. living room is in north side of the house, next to common room 2, bathroom and kitchen room, the size is 220 sqft. master room is in south side of the house, next to common room 1 and balcony, the size is 120 sqft.*

Sample1: *bathroom is in west side of the house, next to common room 2 and living room, the size is 48 sqft. common room 1 is in south west corner of the house, next to master room and bathroom, the size is 140 sqft.*

Sample2: *kitchen is in north east corner of the house, next to living room and balcony, the size is 64 sqft. living room is in north side of the house, next to common room 2, bathroom and kitchen room, the size is 220 sqft.*

… …

Figure 2. The framework of iterative LLMs

As shown in Figure 2, the iterative generation process in ChatDesign begins with the user's initial textual input, which serves as the foundation for generating an initial floor plan. To enhance the design further, the methodology utilizes random sampling to create modified textual inputs called Samples. These Samples are then used to generate alternative floor plans, and the most favourable ones are chosen based on evaluation metrics. By integrating these selected Samples into the iterative process, the generated floor plans undergo continuous refinement and improvement, ensuring a closer match to the user's specific requirements. The iterative approach in ChatDesign allows for continuous optimization, ensuring that the generated floor plans increasingly align with the user's vision. Through multiple iterations and refinements, the design is continuously improved based on user feedback and evolving requirements. This iterative process facilitates fine-tuning and adjustments, resulting in floor plans that better meet the user's specific needs and preferences.

3.3. DETAILS

3.3.1. Uniform Instructions Translation (UIT)

Figure 3 illustrates the integration of the Uniform Instructions Translator (UIT) within the ChatDesign system. The UIT is designed to handle various forms of user interactions, such as text, voice, mouse clicks, and drag-and-drop, by translating them into five core operations: Locate, Move, Resize, Add, and Delete. Each operation is accompanied by specific parameters, such as object type, target location, and dimensions, to fulfil different user requirements.

For example, when a user interacts through text input, the UIT can identify keywords and phrases to determine which core operation should be executed. Similarly, when a user interacts using mouse clicks or drag-and-drop gestures, the UIT can recognize the action and translate it into the corresponding core operation.

By translating different user interactions into these five core operations, the Uniform Instructions Translator provides a unified and standardized way to process user input. This allows for a seamless and intuitive interaction experience, enabling users to actively participate in the design process and accomplish their desired modifications or additions to the floor plan.

Furthermore, the UIT's ability to handle diverse input modalities ensures that ChatDesign can cater to a wide range of users with varying preferences and abilities. This inclusivity enhances the overall user experience and makes the design process more accessible to a broader audience.

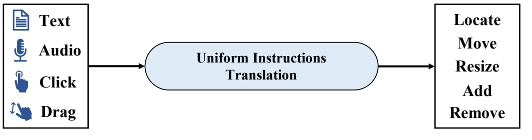


Figure 3. The functionality of Uniform Instructions Translation.

3.3.2. Locate

As shown in Figure 4 (a), the Locate operation is used to position elements within the floor plan. It allows users to specify the exact location or coordinates where a room or object should be placed. This precise control over element placement ensures that users can create floor plans that closely align with their vision and requirements, making the design process more personalized and satisfying.

3.3.3. Move

As shown in Figure 4 (b), the Move operation enables users to relocate rooms or objects within the layout. By providing new coordinates or indicating a target destination, users can easily move elements to desired positions.

3.3.4. Resize

As shown in Figure 4 (c), the Resize operation allows users to adjust the size and dimensions of rooms. Users can specify new dimensions or provide scaling factors to increase or decrease the size of a room. This functionality enables users to fine-tune the floor plan to accommodate specific space requirements, ensuring that the design is both functional and aesthetically pleasing.

3.3.5. Add

As shown in Figure 4 (d), the Add operation permits users to add new rooms or objects to the floor plan. Users can specify the type of room, its dimensions, and any additional details required. This capability allows users to expand and customize their design, adding new elements that cater to their unique needs and preferences, resulting in a more personalized floor plan.

3.3.6. Delete

As shown in Figure 4 (e), the Delete operation enables users to remove existing rooms or objects from the floor plan. By identifying the element to be deleted, users can easily eliminate unwanted components. This feature provides users with the freedom to declutter and simplify their design, ensuring that the final floor plan is efficient, practical, and tailored to their specific needs.

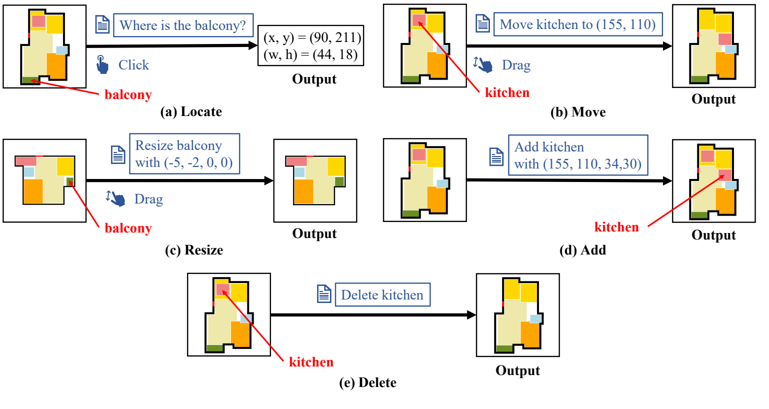


Figure 4. The operations of Locate, Move, Resize, Add and Delete.

In summary, the five core operations - Locate, Move, Resize, Add, and Delete - implemented in the Uniform Instructions Translator (UIT) provide users with a comprehensive toolset for customizing their floor plans. These operations allow users to exert precise control over the placement, size, and composition of their designs, facilitating a highly personalized and user-centric design process. By enabling users to interact with the system in a variety of ways, from text and voice to mouse clicks and drag-and-drop, the UIT ensures a seamless and intuitive user experience. This highlights the strength of ChatDesign in accommodating diverse user preferences and transforming complex design tasks into simple, manageable steps.

4. Experiments

4.1. SETTINGS

4.1.1. Datasets

We utilized the T2D dataset (Leng et al., 2023), which comprises 5,051 manually annotated language instructions and 75,737 AI-generated language instructions. For the T2D dataset, Amazon Mechanical Turk workers were hired and instructed to provide instructions for each room based on the given floor plan image. The requested instructions were expected to reflect the semantic, geometric, and topological information of the floor plan, allowing designers to ideally reproduce the layout based on the instructions. The remaining floor plans were used to generate artificial language instructions by the AI, following predefined templates. Additionally, human designers evaluated and audited the quality of the datasets to ensure their reliability and accuracy.

4.1.2. Evaluation Metrics

To perform evaluations, we employed Intersection over Union (IoU) scores at both macro and micro levels, measuring the overlap between pixel-level ground truth (GT) and generated floor plans. The definitions of these metrics are as follows:

$$\text{Micro IoU} = \frac{\sum_{r=1}^{R} I_r}{\sum_{r=1}^{R} U_r}, \qquad \text{Macro IoU} = \frac{1}{R}\sum_{r=1}^{R} \frac{I_r}{U_r}$$

where I_r and U_r represent the intersection and union of the ground truth and predicted rooms, respectively, for the r-th room type in the floor plan. R denotes the total number of room types. The Macro IoU calculates the average IoU across different room types, while the Micro IoU computes the global IoU by aggregating all rooms.

4.2. MAIN RESULTS

4.2.1. Comparison with state-of-the-art methods

As shown in Table 1, iterative indicates that we employ iterative optimisation with random sampling in the training of the model. Our proposed ChatDesign outperforms all other methods in terms of Micro-IoU. Additionally, ChatDesign-iterative surpasses all other methods in terms of Macro-IoU. These results indicate that our approach achieves superior performance compared to existing methods. Moreover, our methodology approaches the level of performance exhibited by human designers, indicating its effectiveness in generating floor plans that closely align with human expertise and design principles.

As shown in Table 2, we randomly selected 10 challenging examples from the T2D dataset (Leng et al., 2023). The second row represents the performance of Tell2Design, while the third row shows the results after interactive refinement using user feedback. It can be observed that ChatDesign demonstrates an improvement of nearly 10% compared to the original performance. This improvement underscores the value of integrating user feedback into the learning process of AI models, allowing them to better adapt and optimize their performance. Moreover, it highlights the potential of ChatDesign as a powerful tool for tackling complex tasks, demonstrating its capability to learn and improve from interactive refinement.

Method	Micro-IoU	Macro-IoU
Obj-GAN (Li et al., 2019)	10.68	8.44
CogView (Ding et al., 2021)	13.30	11.43
Imagen (Saharia et al., 2022)	12.17	14.96
Tell2Design (Leng et al., 2023)	54.34	53.30
ChatDesign (ours)	**58.31**	55.43
ChatDesign-iterative (ours)	54.57	**57.24**
Human	64.67	62.32

Table 1. IoU between generated floor plans and ground-truth for ChatDeisgn and other baselines.

Method	Micro-IoU	Macro-IoU
Tell2Design (Leng et al., 2023)	47.56	51.10
ChatDesign (ours)	**56.82**	**60.56**

Table 2. IoU between generated floor plans and ground-truth for ChatDeisgn and Tell2Design.

4.2.2. Visualisation Results

As shown in Figure 5, the floor plans generated by our method outperform all the comparative approaches, showcasing exceptional performance. Furthermore, it can be observed that the floor plans generated by ChatDesign are closer to the real-world layout compared to the floor plans hand-drawn by human designers. This suggests that our ChatDesign approach can produce floor plans that exhibit a high degree of realism and accuracy.

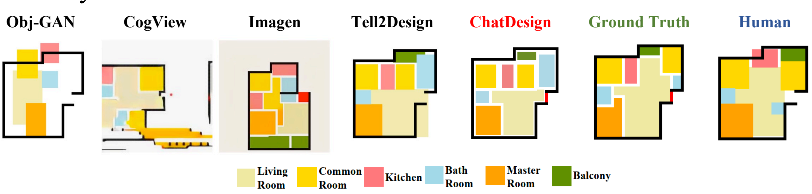


Figure 5. The visualisation results for ChatDesign and other baselines.

Figure 6 indicated that we further optimise the model-generated floor plans using post-processing to further approximate human needs. The main steps include Add door, Erase, Boundary

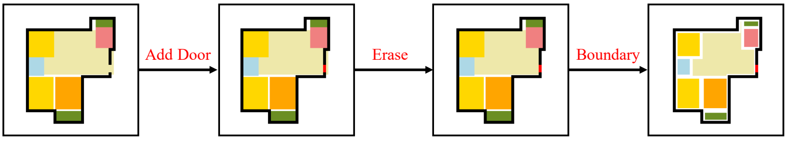


Figure 6. The post-process of model-generated floor plans.

The Add Door step adds entrance doors to the entire space, ensuring accessibility to the floor plan. The Erase step removes unnecessary elements outside the main walls, maintaining the design's integrity. The Boundary step generates boundaries between different rooms, defining individual spaces and enhancing the floor plan's readability and practicality. These post-processing steps refine the floor plan, making it functional and aesthetically pleasing. They contribute to a well-planned, coherent space that meets the user's specific needs and preferences.

4.2.3. Result Analysis

The experimental results demonstrate that our ChatDesign approach performs exceptionally well on the T2D dataset. Compared to other methods, ChatDesign achieves a significant improvement of nearly 10% in terms of performance. The visual results shown in Figure 5 indicate that the floor plans generated by ChatDesign outperform other methods and closely resemble real-world floor layouts.

4.3. DISCUSSION

4.3.1. The Role of the Human Designer

The human designer plays a crucial role in assisting the AI in the floor plan generation process. They provide floor plans as a reference for evaluating ChatDesign's performance, enabling the AI system to learn from their expertise and creativity. The designer also provides subjective ratings for the generated floor plans, serving as feedback for further improvement. This feedback loop allows the AI system to refine its performance and adapt to the designer's preferences. This collaborative approach enhances the abilities of both the human designer and the AI system, resulting in efficient and effective floor plan generation. By combining human creativity and AI automation, the partnership leads to a flexible, adaptive, and user-centric design process, catering to diverse user needs while reducing time and effort.

4.3.2. Future Research

In future work, we suggest exploring the incorporation of additional design constraints, such as building codes and accessibility requirements, to ensure regulatory compliance in the generated floor plans. Furthermore, investigating methods for generating 3D representations would enhance visualization and enable comprehensive design evaluations.

Expanding the training dataset and exploring transfer learning techniques could improve the model's understanding of domain-specific language and enhance its ability to generate diverse and accurate floor plan designs. Conducting user studies with professional architects and designers would provide valuable feedback on the usability and practicality of ChatDesign.

5. Conclusion

In conclusion, this paper introduced ChatDesign, an innovative approach that leverages pre-trained large language models (LLMs) to generate floor plan designs from natural language descriptions. By incorporating iterative modifications based on user interaction, our approach enables precise customization of floor plans to cater to specific user requirements. Through rigorous experiments and user studies, we demonstrated the feasibility and effectiveness of our method, consistently outperforming existing approaches and producing floor plans that rival those created by human designers. Our contributions include a novel method for floor plan generation using pre-trained LLMs, an interactive scheme that empowers users to modify generated floor plans, and the implementation of multiple interaction modalities for a user-centric design process.

Acknowledgements

We gratefully thank the National Science Foundation of China (Grant Nos. 52078294 and 51825802) for funding this work. Jinmin Li and Yilu Luo are co-first authors and Shuai Lu is corresponding author.

References

Leng, S., Zhou, Y., Dupty, M. H., Lee, W. S., Joyce, S. C., & Lu, W. (2023). Tell2Design: A Dataset for Language-Guided Floor Plan Generation. arXiv preprint arXiv:2311.15941.

Wu, W., Fu, X. M., Tang, R., Wang, Y., Qi, Y. H., & Liu, L. (2019). Data-driven interior plan generation for residential buildings. ACM Transactions on Graphics (TOG), 38(6), 1-12.

Li, W., Zhang, P., Zhang, L., Huang, Q., He, X., Lyu, S., & Gao, J. (2019). Ob ject-driven text-to-image synthesis via adversarial training. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 12174-12182).

Ding, M., Yang, Z., Hong, W., Zheng, W., Zhou, C., Yin, D., ... & Tang, J. (2021). Cogview: Mastering text-to-image generation via transformers. Advances in Neural Information Processing Systems, 34, 19822-19835.

Saharia, C., Chan, W., Saxena, S., Li, L., Whang, J., Denton, E. L., ... & No rouzi, M. (2022). Photorealistic text-to-image diffusion models with deep language understanding. Advances in Neural Information Processing Systems, 35, 36479-36494.

Chen, Q., Wu, Q., Tang, R., Wang, Y., Wang, S., & Tan, M. (2020). Intelligent. home 3d: Automatic 3d-house design from linguistic descriptions only. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 12625-12634).

Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., ... & Liu, P. J. (2020). Exploring the limits of transfer learning with a unified text-to-text transformer. The Journal of Machine Learning Research, 21(1), 5485-5551.

Liu, X., Yu, H., Zhang, H., Xu, Y., Lei, X., Lai, H., ... & Tang, J. (2023). Agentbench: Evaluating llms as agents. arXiv preprint arXiv:2308.03688.

Zhang, Y., Yang, J., Yuan, Y., & Yao, A. C. C. (2023). Cumulative reasoning with large language models. arXiv preprint arXiv:2308.04371.

DEEP SPATIAL MEMORY

Quantifying Architectural Spatial Experiences through Agent-driven Simulations and Deep Learning

SHUHAN MIAO[1], WENZHE PENG[2], DANIEL TSAI[3] and TAKEHIKO NAGAKURA[4]
[1]*Harvard Graduate School of Design*
[2,3,4]*Massachusetts Institute of Technology*
[1]*shuhan_miao@gsd.harvard.edu, 0009-0006-4522-6814*
[2]*pwz@alum.mit.edu, 0009-0008-0129-7578*
[3]*dtsai@mit.edu, 0000-0002-8082-6392*
[4]*takehiko@mit.edu, 0000-0002-3219-3930*

Abstract. In architectural theory, the spatial experience is dynamic, evolving from sequences of interconnected views shaped by past encounters and future expectations. Traditional computational methods such as Isovists provide geometric insights but fall short in representing their sequential nature. To address this gap, the paper introduces a novel methodology that combines agent-driven simulation, 3D Isovist sampling, and deep learning for quantitative analysis and comparison of spatial experiences in architecture. This approach leverages the Grasshopper plugin Pedsim for simulating pedestrian paths and a self-supervised video representation learning model MemDPC for processing depth panorama sequences and extracting numerical features for each sequence. The methodology is first validated through a controlled experiment with various sequence typologies, affirming its efficacy in recognizing typological similarities. A case study is conducted comparing Louis Kahn's designs with Roman architecture, quantitatively analysing their intertwined spatial experiences. This research offers a framework for quantitatively comparing spatial experiences across buildings and interpreting the nuanced impact of historical references on modern spaces.

Keywords. Deep Learning, Artificial Intelligence, Spatial Experience, Isovist, Agent-driven Simulation, Self-Supervised Learning

1. Introduction

In the field of architectural theory, the representation of spatial experience has been discussed extensively, particularly focusing on its dynamic and selective nature. Zevi views in *Architecture as Space* that architectural spatial experience as a time-imbued continuum, best captured through film rather than static images,

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 109-118.

to capture the full essence of navigating and interacting with spaces (Zevi, 1974, p. 59). Adding to this dynamic aspect is the selectivity of spatial experience. Ching suggests in *Architecture, Form, Space, and Order* that our experience of space is inherently selective, influenced by our navigation in the space (Ching, 1979, p. 280). This selectivity in perception means that we don't absorb all elements of a space equally; rather, our focus shifts as we navigate, leading to a subjective interpretation based on our position and movement (Hershberger, 1970, p. 43). Such subjectivity presents a challenge for architects and theorists who seek to understand and compare architectural spaces in a more objective lens. Traditional methods of quantitative representation—such as plans, sections, and static images—offer a fragmented view that misses the fluid continuum of experience as described above.

Addressing this gap, the research introduces a novel methodology for quantitatively representing and comparing architectural spatial experiences. It integrates agent-driven simulation, 3D Isovist techniques, and artificial neural networks for pedestrian path generation, spatial data capture, and feature extraction. The methodology is validated through two experiments. The first employs self-supervised learning for feature extraction from spatial sequence typologies, using unsupervised clustering to validate the features' self-clustered quality, then applies a supervised classifier to translate features into legible sequential types. The second experiment extends the methodology to a real-world application, examining the influence of Roman architecture on Louis Kahn's designs. The model, fine-tuned for complex architectural spaces, categorizes sequential types to initially understand typological similarities and differences. Subsequent in-depth feature analysis through unsupervised clustering and nearest neighbour methods uncovers latent patterns within same sequential type. This methodology converts subjective spatial experiences into quantifiable data, enabling an objective comparative analysis of historical and contemporary architectural elements.

2. Related Works

Machine learning involves computer systems that analyse and deduce patterns in data through algorithms and statistical models. This approach is particularly beneficial in architectural spatial analysis, facilitating the understanding and interpretation of complex spatial data. Machine learning includes categories such as supervised and unsupervised learning. Supervised learning trains models on labelled datasets for classifying data or predicting outcomes. In contrast, unsupervised learning explores unlabelled data to uncover its inherent structure (Deng & Yu, 2014). In architectural spatial analysis, both methods offer distinct benefits. The predefined labels of supervised learning enhance interpretability and specificity but lack generalizability across diverse architectural spaces. Unsupervised learning, which discerns patterns from unlabelled data, offers broader exploration capabilities. However, the lack of human-readable labels necessitates additional analysis for interpretation.

Isovist analysis is essential in architectural studies for quantifying spatial environments. It maps visible points from a selected perspective, translating spatial geometry into perceptual experiences, capturing both phenomenological and morphological aspects of spaces (Benedikt, 1979). This approach has evolved over decades in fields such as analysing plan visibility and urban space quality (Dawes & Ostwald, 2014; Leduc et al., 2011; Turner et al., 2001). Recent years have seen a surge in the integration of Isovist methods with breakthroughs in deep learning. Peng et al. (2017) use 2D depth map images of 3D Isovists to train a deep convolutional neural network in a supervised learning manner to classify spatial typologies. Conversely, Johanes & Huang (2021) applied self-supervised learning on 2D Isovist data, using an unsupervised clustering algorithm to categorize spatial plans. Additionally, recent studies have merged 2D Isovist analysis with pedestrian trajectory simulations, analysing Isovist data along simulated paths using unsupervised clustering to uncover spatial properties (Feld et al., 2020; Sedlmeier & Feld, 2018). Building on these developments, the proposed experiments seek to integrate the strengths of both supervised and unsupervised learning methods, combining human-readable categorization of supervised learning with the flexible pattern discovery of unsupervised learning. This hybrid method is designed to create a comprehensive and adaptable framework for architectural analysis.

3. **Methods**

This study employs a comprehensive methodology combining agent-driven simulation (Section 3.1), 3D Isovist sampling (Section 3.2), and self-supervised representation learning (Section 3.3). It aims to capture the complexity of spatial experiences by simulating pedestrian behaviour, sampling spatial geometry through 3D Isovists, and extracting latent representations of these experiences using self-supervised learning techniques (Figure 1). This approach is designed to systematically encode and analyse spatial sequences in architecture.

3.1. AGENT-DRIVEN SIMULATION

To effectively capture the subjective nature of spatial sequences, the study utilizes PedSim, a Grasshopper plug-in that simulates pedestrian paths using the social force model and anticipatory collision avoidance (Riise, 2022). While actual pedestrian movement data could be sourced from real-world sensors, the controlled simulation ensures consistent input for analysis purpose. Agents in the simulation algorithm will move from designated start to end points, visiting various points of interest within their field of vision. For consistency and comparability across different spaces, each trajectory is established with a single start and end point, along with ten interest points strategically selected to represent architectural features such as columns and arches. The intentional inclusion of expert knowledge ensures a realistic simulation of pedestrian behaviour. Paths taken by these simulated agents are recorded and converted into 3D polylines at a standard eye level of 1.6 meters for 3D Isovist sampling.

By simulating various paths within a single trajectory, the approach reflects diverse individual experiences in a unified spatial sequence. This variety also serves as data augmentation, ensuring robustness in subsequent model training.

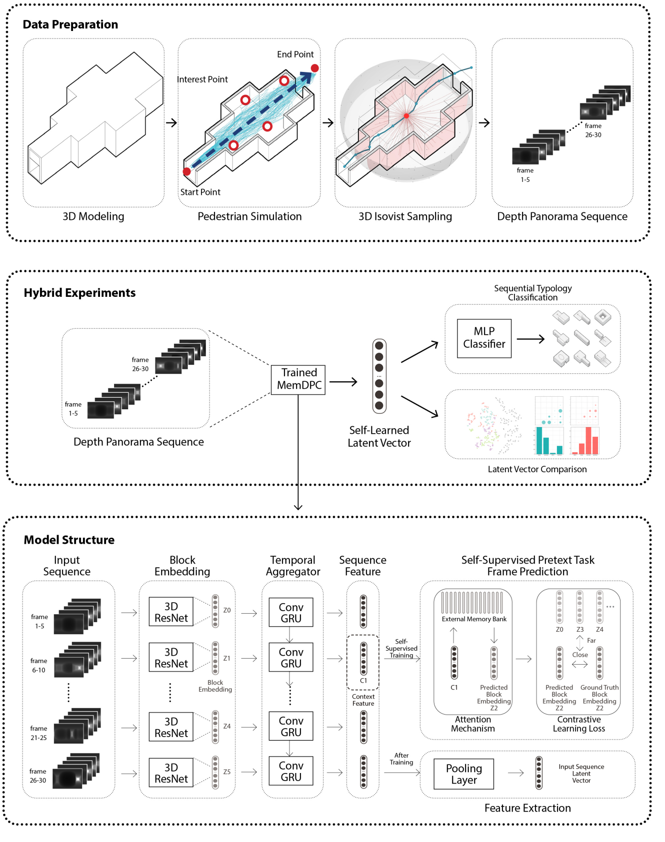


Figure 1. Illustration of Proposed Pipeline

3.2. 3D ISOVIST SAMPLING

This study's 3D Isovist sampling employs a custom Python-scripted component in Grasshopper to process simulated pedestrian paths. Each path is segmented into 30 equally spaced points. At each point, rays are projected onto a sphere's surface, and the distances at intersection points are computed and transformed into grayscale values. These values are then remapped into depth panorama images, conforming to the resolution parameters established by Peng et al. (2017). A consistent maximum sampling distance of 20 meters is set to accommodate the monumental scale of the buildings in the following case studies.

It is vital to recognize that while these depth map images effectively capture spatial geometrical boundaries facing rays, they represent a reduced version of the complete spatial experience. This study specifically focuses on spatial geometry information, presenting it as a simplified form of spatial experience. Therefore, the complexity of representing spatial experience is distilled into 30 sequential depth panorama images for each simulated path. These image sequences, formatted as videos, are compatible with existing self-supervised video representation learning models, enabling a more nuanced analysis of spatial perception.

3.3. SELF-SUPERVISED REPRESENTATION LEARNING

Self-supervised learning (SSL) is a subset of unsupervised learning. It trains models on large datasets without the need for predefined labels, thus overcoming the limitations of supervised learning (Jing & Tian, 2019). SSL employs 'pretext tasks', derived directly from unlabelled data, such as video reconstruction or frame order prediction. Solving such complex pretext tasks requires the model to have a high-level understanding of the training data to learn generalizable features (Schiappa et al., 2023). SSL in this study involves employing MemDPC, a model that uses frame prediction as its pretext task (Han et al., 2020). This model divides each video into blocks, encoding them into embeddings. These embeddings are time-aggregated using RNNs (Recurrent Neural Networks) to extract the context feature for each block. A global attention mechanism is initialized as a learnable memory bank for hypothesis formation and future block prediction. After training, an average pooling layer condenses the context features of each block into a 256-dimensional feature vector, representing the latent spatial experience of a simulated path. The memory bank approach conceptually resembles how humans accumulate spatial experiences. For humans, this includes visual impressions and understanding of spatial relationships. However, unlike human cognition, neural networks identify patterns without semantic understanding, necessitating further analysis. Subsequent experiments use sequential typologies as labels to translate these SSL-acquired features into human-readable terms and analyses these features to discern underlying similarities.

4. **Experiments**

This section details experiments that validate our framework for studying sequential spatial experiences, combining a supervised approach for identifying typological similarities and differences, and an unsupervised method for in-depth pattern discovery. The first experiment, utilizing the 'typology dataset', validates the model's capability in autonomously clustering features through self-supervised learning. A supervised classifier is then applied to these features to translate the 256-dimensional vector into nine human-readable sequential typologies. The second experiment, with the 'case study dataset', fine-tuned this model to adapt to the complexities of real-world architectural spaces. Here, the classifier initially categorizes sequential typologies, facilitating a preliminary organization of trajectories. This is followed by a detailed feature analysis using similarity measures and clustering, designed to reveal deeper patterns and nuances within architectural spaces categorized under the same typology.

4.1. EXPERIMENT I: SEQUENTIAL TYPOLOGIES

The first experiment examines the SSL model's ability to effectively extract features, using a typology dataset of 2,700 path sequences with three space types: 'room', 'passage', and 'exterior' at both start and end, creating nine distinct combinations (Figure 2). These sequence data undergo augmentation during SSL training, including adjustments in brightness, playback speed, cropping, and horizontal flips. After training, each sequence is transformed into a 256-dimensional feature vector. The model's performance in feature extraction is initially assessed by its validation accuracies in the SSL pretext task of frame prediction, achieving 75% top-1 and 99% top-5 accuracies. These results indicate a robust feature learning capability without the need for labelled data. To further evaluate the learned features, unsupervised clustering is applied. Using the K-means algorithm (Lloyd & S., 1982) to cluster these 256-dimensional feature vectors and comparing the results against the ground truth of sequential typology labels yielded a 96% accuracy. Additionally, the clustering's relevance and precision are confirmed by an Adjusted Mutual Information score (Vinh et al., 2010) of 0.95, reflecting a high degree of correspondence between the unsupervised clustering outcomes and the actual labels. This demonstrates that the SSL-extracted features inherently possess a self-clustered organization that significantly correlates with human-defined categories.

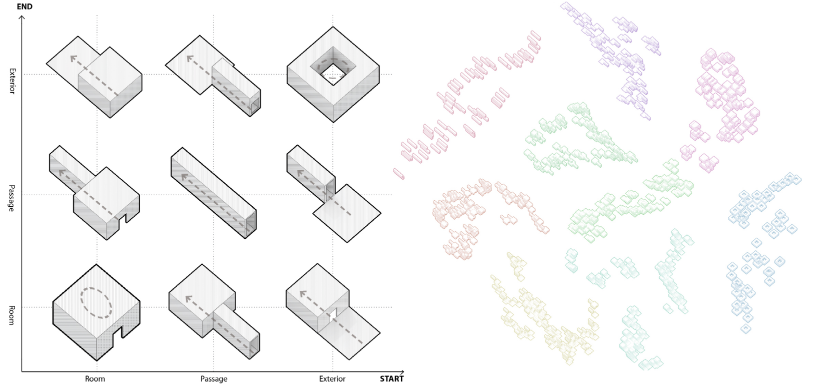


Figure 2. Sequential typologies(left) and t-SNE visualization of K-means Clustering (right)

To further interpret these self-organized features in a subsequent experiment, a Multi-Layer Perceptron (MLP) classifier was trained on these features to translate them into human-readable labels. The MLP outputs a 6-D probability vector, with the first three dimensions predicting the start type and the last three the end type. This classifier's high validation (99%) and test set (98%) accuracies, where accuracy implies correctly matching both start and end types, confirm its efficacy in accurately categorizing sequential typologies from SSL-extracted features.

4.2. EXPERIMENT II: COMPARATIVE ARCHITECTURAL CASE STUDY

The second experiment aims to explore the influence of Roman architecture on Louis Kahn's designs, fine-tuning the SSL model and MLP classifier from the first experiment. This approach mirrors human cognitive processes, where new spatial experiences are interpreted based on prior knowledge. The strength of such a framework in this comparative study lies in its ability to analyse complex spatial relationships and discern nuanced design influences, tasks typically challenging for traditional methods. By adapting a framework that already has a baseline understanding of various spatial sequences, the models are equipped to interpret intricacies of real-world architectural spaces.

The choice of case study was motivated by the well-documented influence of Roman architecture on Kahn's design philosophy. Kahn's 1950 visit to Rome as a resident architect left a profound and well-known impact on his subsequent designs. This is evidenced in his own writings, accounts from his family and colleagues, and analyses by historical theorists (Barizza, n.d.; Rabifard, n.d.; Scully, 1992). The Indian Institute of Management (IIM) was chosen for its intentional design references to Roman ruins, aiming to invoke a sense of monumentality. Similarly, the Pantheon, Trajan's Market, and Baths of Caracalla were acknowledged by Kahn as influential in his design process. The study aims to quantitatively analyse these interactions using the proposed framework, seeking insights into Kahn's design strategy and philosophy.

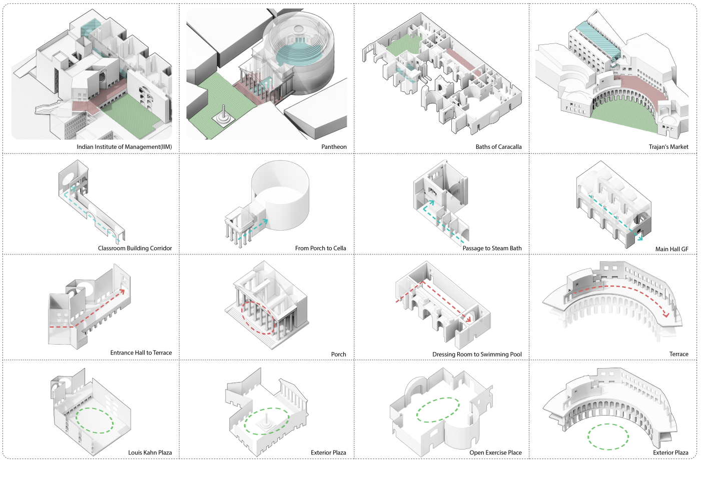


Figure 3. Sampled Trajectories in Four Case Study Buildings

The sampling strategy acknowledges the challenges of representing the universal average experience, given the subjectivity inherent in spatial perception. The criteria for selecting trajectories prioritized architectural significance, ensuring that each path chosen highlighted either unique design elements, areas frequently engaged by visitors, or spaces with notable architectural details. The final selection of 12 trajectories was carefully curated to ensure that they were not only encompassing a spectrum of experiences but also provided the most valuable data for the study's comparison objectives as shown in Figure 3. Each trajectory is simulated with 100 paths to enrich the dataset.

The SSL model and MLP classifier were fine-tuned using the combination of typology dataset and a subset of the case study dataset. The selected subset, constituting 31% of the total case study dataset, included 5 trajectories with clearly distinguishable sequential types, carefully chosen to avoid ambiguous labels that might confound the model. This fine-tuning aimed to retain the model's foundational knowledge while adapting to the architectural complexities of the case study.

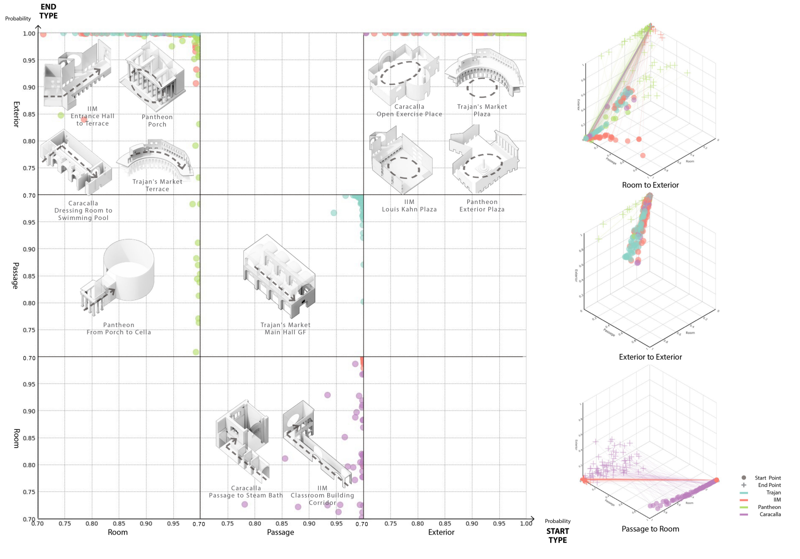


Figure 4. Visualization of Output of MLP Classifier. Left is a nested scatter plot illustrating the predicted probabilities for 12 sampled trajectories across 9 sequential types. On the right, 3D plots highlight the probability vectors for specific start and end types, connecting pairs of points for each building

The fine-tuned models, with 96% validation accuracy on the combined dataset and 94% on the labelled case study subset, effectively grouped trajectories with similar sequence typologies in the case study dataset (Figure 4). Such a typological grouping is crucial for in-depth comparative studies, as it initially sorts sequences into similar typologies for further comparison of interest. Additionally, this approach highlights the framework's efficiency in processing large-scale datasets, a typically time-consuming and challenging task for manual analysis. In the 'exterior to exterior' category, the K-means clustering and nearest neighbour similarity analysis on the feature vectors revealed a notable similarity between the IIM and Trajan's Market, while distinctly differentiated from the Pantheon, Caracalla's Baths, and a standard courtyard (Figure 5). This pattern suggests shared architectural elements such as terraces and colonnades in both buildings and the absence of arches in the Pantheon's Plaza, exemplifying the framework's ability to uncover nuanced connections not immediately apparent to human observers. By quantitatively assessing how IIM's feature vector diverges from those of Roman buildings and a standard courtyard, the framework enables a data-driven exploration of Kahn's design choices, thereby complementing traditional qualitative analyses.

Figure 5.t-SNE Visualization of K-means Clustering of 'Exterior-to-Exterior' Feature Vectors and Nearest Neighbour Similarity Analysis

5. **Conclusion**

This research showcases a robust analytical pipeline effectively merging deep learning's computational pattern recognition with the detailed interpretation required in architectural studies. This research's contributions extend from enabling immediate comparative analysis to establishing a foundation for future methodological exploration. Moving forward, the aim is to enhance the spatial perception capabilities of the model by incorporating additional sensory channels beyond grayscale, such as integrating semantic segmentation or object detection layers. Further refinement of Isovist data resolution will allow for a more detailed understanding of architectural elements like ornamentation. Such enhancements will enable the direct visualization of correlations between the model's features and specific architectural elements, thereby greatly enriching the interpretive value. In addition, expanding the dataset and utilizing clustering algorithms will enable the quantification of common qualities in spatial experiences, such as 'monumentality' in the case study experiment, thereby paving the way for a new paradigm in architectural analysis that bridges qualitative assessment with quantitative precision.

References

Barizza, E. (n.d.). *Rome and the legacy of Louis I. Kahn*. Routledge, an imprint of the Taylor & Francis Group.

Benedikt, M. (1979). To Take Hold of Space: Isovists and Isovist Fields. *Environment and Planning B: Planning and Design*, *6*(1), 47–65. https://doi.org/10.1068/b060047

Ching, F. D. K. (1979). Circulation. In *Architecture: Form, Space, and Order* (pp. 280–334). John Wiley & Sons, Inc., Hoboken, New Jersey.

Dawes, M. & Ostwald, M. (2014). Prospect-Refuge theory and the textile-block houses of Frank Lloyd Wright: An analysis of spatio-visual characteristics using isovists. *Building and Environment, 80*, 228–240. https://doi.org/10.1016/j.buildenv.2014.05.026

Deng, L. & Yu, D. (2014). Deep Learning: Methods and Applications. *Found. Trends Signal Process.*, *7*, 197–387. https://api.semanticscholar.org/CorpusID:53304118

Feld, S., Illium, S., Sedlmeier, A. & Belzner, L. (2020). Trajectory annotation using sequences of spatial perception. *ArXiv*. https://doi.org/10.48550/arxiv.2004.05383

Han, T., Xie, W. & Zisserman, A. (2020). Memory-augmented Dense Predictive Coding for Video Representation Learning. *ArXiv*. https://doi.org/10.48550/arxiv.2008.01065

Hershberger, R. G. (1970). Architecture and Meaning. *The Journal of Aesthetic Education*, *4*(The Environment and the Aesthetic Quality of Life (Oct., 1970), pp. 37–55), 37–55. https://www.jstor.org/stable/3331285

Jing, L. & Tian, Y. (2019). Self-supervised Visual Feature Learning with Deep Neural Networks: A Survey. *ArXiv*. https://doi.org/10.48550/arxiv.1902.06162

Johanes, M. & Huang, J. (2021). Deep Learning Isovist: Unsupervised Spatial Encoding in Architecture. *Proceedings of the 41st Annual Conference of the Association of Computer Aided Design in Architecture (ACADIA)*, 134–141. https://doi.org/10.52842/conf.acadia.2021.001

Leduc, Chaillou, T. and, Ouard, F. and & Thomas. (2011). *Towards a ``typification'' of the Pedestrian Surrounding Space: Analysis of the Isovist Using Digital processing Method*. 275–292. https://doi.org/10.1007/978-3-642-19789-5_14

Lloyd & S. (1982). Least squares quantization in PCM. *IEEE Transactions on Information Theory*, *28*(2), 129–137. https://doi.org/10.1109/tit.1982.1056489

Peng, W., Zhang, F. & Nagakura, T. (2017). Machines' Perception of Space. *Proceedings of the 37th Annual Conference of the Association of Computer Aided Design in Architecture (ACADIA)*, 474–481.

Rabifard, M. (n.d.). *The Integration of Form and Structure in The Work of Louis Kahn*. https://api.semanticscholar.org/CorpusID:147416394

Riise, J. (2022). *PedSim* [Grasshopper Script]. https://github.com/julianriise/pedsim

Schiappa, M. C., Rawat, Y. S. & Shah, M. (2023). Self-Supervised Learning for Videos: A Survey. *ACM Computing Surveys*, *55*(13s), 1–37. https://doi.org/10.1145/3577925

Scully, V. (1992). Louis I. Kahn and the Ruins of Rome. *MoMA (New York, N.Y.), 12*, 1–13.

Sedlmeier, A. & Feld, S. (2018). Learning indoor space perception. *Journal of Location Based Services*, *12*(3–4), 179–214. https://doi.org/10.1080/17489725.2018.1539255

Turner, A., Doxa, M., O'Sullivan, D. & Penn, A. (2001). From Isovists to Visibility Graphs: A Methodology for the Analysis of Architectural Space. *Environment and Planning B: Planning and Design*, *28*, 103–121. https://doi.org/10.1068/b2684

Vinh, N. X., Epps, J. & Bailey, J. (2010). Information Theoretic Measures for Clusterings Comparison: Variants, Properties, Normalization and Correction for Chance. *J. Mach. Learn. Res.*, *11*, 2837–2854.

Zevi, B. (1974). The Representation of Space. In B. Zevi, *Architecture as Space* (pp. 45–60). Horizon Press.

DSNL IN ARCHITECTURE– A DEEP LEARNING APPROACH TO DECIPHERING ARCHITECTURAL SKETCHES AND FACILITATING HUMAN-AI INTERACTION

WEI HU[1,2]
[1]*College of Architecture and Urban Planning, Tongji University*
[2]*Urban Computing Lab, Shanghai Artificial Intelligence Laboratory*
[1]*1990huwei@sina.com, ORCID: 0000-0001-7754-1858*

Abstract. The language of interaction between architects and machines has been evolving towards a more user-friendly paradigm. As the capabilities of machines and artificial intelligence have advanced, it has become increasingly feasible for architects to communicate with machines using their customary expressive methods. Consequently, this has led to the development of Domain-Specific Natural Language (DSNL), which, unlike traditional Domain-Specific Language (DSL), places greater emphasis on naturalness. While this naturalness enhances usability for architects, it also presents challenges in machine comprehension. To address this issue, we propose a data-driven approach that utilizes domain-specific data for model training or fine-tuning through unsupervised or weakly supervised methods. Our study, which focuses on teaching AI to learn architectural sketching from architects, demonstrates that our proposed method captures the characteristics of human architectural sketching more effectively than traditional approaches.

Keywords. Domain Specific Natural Language, Human-AI interaction, Architectural sketches, AIGC, Deep learning.

1. Introduction

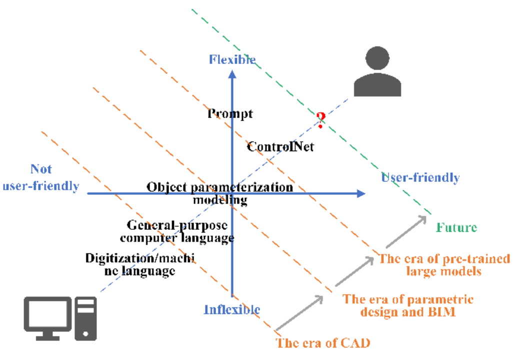


Figure 1. Trends in the Development of Interaction Languages between Architects and Machines

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 119-128.

Recent technological advancements have significantly advanced the role of Artificial Intelligence (AI) in creative design fields. Despite some algorithms impressing experts, AI still lacks the capacity to completely replace human experts in design. Consequently, a practical approach is to merge human expertise with AI, fostering effective collaboration. Effective communication stands central to this collaboration. An examination of the evolution in human-machine interaction (Hu, 2023), depicted in Figure 1, reveals a progression from basic machine languages to more complex systems like CAD, BIM, and PLM (Pre-trained Language Model). Advances in machine and AI technologies have led to more human-centric interaction patterns, signifying a shift towards user-friendly interfaces and processes. The ongoing evolution of AI and machine capabilities raises important considerations regarding the future nature of interaction between architects and machines.

Therefore, building upon DSL-based interaction methods like BIM and visual programming (Fowler, 2010), we propose that future human-machine interaction languages will more closely align with human expressive habits, offering greater flexibility and user-friendliness while better addressing professional needs. In the future, design experts could teach computers to understand 'Domain-Specific Natural Language (DSNL)', rather than adapting their own habits to computers and AI. DSNL primarily introduces the characteristic of naturalness, in addition to the linguistic aspects, domain specificity, and limited expressiveness of DSN. DSNL, in contrast to typical computer languages, mirrors human expression habits more closely. This approach makes interacting with computers more intuitive, akin to human-human communication, and reduces the cognitive load on users. While naturalness enhances user-friendliness, it also presents challenges for computer understanding, especially when large amounts of labelled data are difficult to obtain in a particular domain. We propose using a data-driven approach to train miniaturized models in unsupervised or weakly supervised ways or fine-tune pre-trained large models to solve this problem.

This paper investigates the practicality of DSNL through the novel approach of instructing an AI model to sketch buildings. Architectural sketching, a common DSNL among architects, offers insight into their design thinking and cognitive processes (Lawson, 2012). We introduce a method to impart this DSNL to AI, inspired by the Beaux-Arts educational system's master-apprentice teaching model (Chen, 2020). This method posits that, akin to the transmission of sketching techniques in architecture, humans and machines can similarly learn domain-specific natural language. Traditional algorithms struggle to teach ambiguous language to AI due to their reliance on structured rules. This study introduces architectural sketches as a training set for teaching DSNL to AI, utilizing data-driven deep learning techniques. To overcome the scarcity of labelled domain-specific samples, our algorithm efficiently trains smaller networks using more readily available unpaired datasets.

The experimental evaluations of our method reveal significant improvements in performance compared to various traditional approaches. Notably, our method excels in accurately interpreting the often ambiguous and nuanced meanings that professional architects convey through DSNL in their sketches.

2. Related works

Historical research on transforming scene images into sketches has predominantly used computer graphics algorithms for style conversion, such as Lu's method that combines image gradients and forward differences for edge extraction and style transfer (Lu, 2012). Additionally, smartphone camera filters and sketch conversion apps, such as "Photo Sketch Maker" and "Pencil Sketch," have gained popularity. However, these methods only replicate the image's "form" and lack value judgments and aesthetic choices based on the image's "meaning," which are crucial for architects and artists in their creative process.

Recent developments, including DALL-E (Ramesh, 2021; Ramesh, 2022), Stable Diffusion (Rombach, 2021), and ControlNet (Zhang, 2023), have innovated in image generation by introducing natural language prompts and image pair control. However, they struggle to capture the nuanced expression in architectural design, a task that is challenging to achieve through text alone or with precise image pairs. Even with the added edge detection techniques, such as Canny and HED (Xie, 2015), fail to replicate the perceptual richness found in architectural sketches.

3. Method

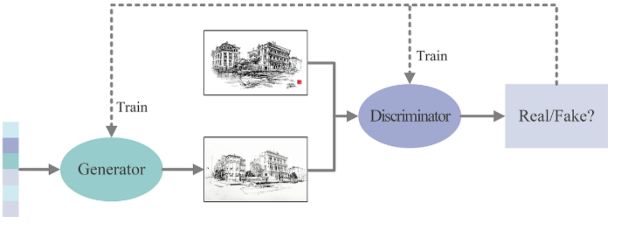


Figure 2. The schematic diagram of GAN

This study attempts to use Generative Adversarial Networks (GANs) (Goodfellow et al. 2014) to transform architectural scene images into sketches. This involves a network comprising a generator and a discriminator, functioning in a zero-sum game to enhance image generation fidelity, as depicted in Figure 2. However, a basic GAN falls short for our goal of generating architectural sketches that correspond accurately to specific scene photos. Consequently, we turn to Conditional GANs (cGANs) (Mirza, Osindero, 2014), trained with scene and sketch image pairs, to ensure the relevance between the generated sketch and the original scene image. Owing to the scarcity of matched scene-sketch pairs, we adopt the Cycle-GAN structure (Zhu et al. 2017) for image-to-image domain mapping without the necessity of matched pairs. Cycle-GANs employ two interconnected GAN networks to guarantee a one-to-one correspondence between actual and generated images, utilizing similarity-based loss functions to regulate the generated content, as illustrated in Figure 3.

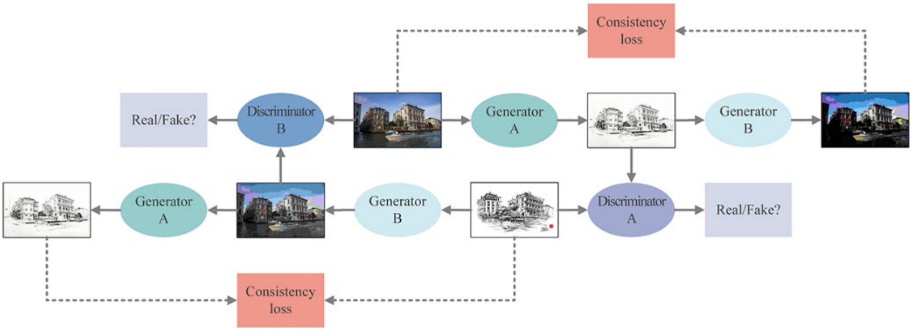


Figure 3. The schematic diagram of Cycle GAN

The generator that transforms images from the photo domain X to the sketch domain Y is denoted as Generator A or $G_A : X \rightarrow Y$. Its objective is to generate corresponding sketch y' from input photo $x \in X$. The quality of the generated sketches is evaluated by the discriminator, referred to as Discriminator A or D_A. On the other hand, the generator that transforms images from the sketch domain to the photo domain is denoted as Generator B or $G_B : Y \rightarrow X$. Its purpose is to generate corresponding photo x' from input sketch $y \in Y$. The quality of the generated photos is assessed by the discriminator, referred to as Discriminator B or D_B.

Generator G_A loss function:

$$\mathcal{L}_{G_A} = \mathcal{L}_{adv}(D_A, y') + \lambda_{cycle}\mathcal{L}_{cycle}^{A\to B}(x, x")$$

Discriminator D_A loss function:

$$\mathcal{L}_{D_A} = \mathcal{L}_{adv}(D_A, y) + \mathcal{L}_{adv}(D_A, y')$$

Generator G_B loss function:

$$\mathcal{L}_{G_B} = \mathcal{L}_{adv}(D_B, x') + \lambda_{cycle}\mathcal{L}_{cycle}^{B\to A}(y, y")$$

Discriminator D_B loss function:

$$\mathcal{L}_{D_B} = \mathcal{L}_{adv}(D_B, x) + \mathcal{L}_{adv}(D_B, x')$$

A key aspect of this model is its use of cyclic mapping between two domains to generate cyclic losses, effectively addressing the lack of one-to-one corresponding training samples. In the Photo-to-sketch-to-photo cycle, where an input photo x undergoes transformation into a generated sketch $y' = G_A(x)$ and subsequent reconstruction into a photo $x'' = G_B(y')$, the objective is to minimize the dissimilarity between the original photo x and the reconstructed photo x''. Similarly, in the Sketch-to-photo-to-sketch cycle, involving an input sketch y transformed into a generated photo $x' = G_B(y)$ and subsequently reconstructed into a sketch $y'' = G_A(x')$, the objective is to minimize the dissimilarity between the original sketch y and the reconstructed sketch y''.

The cyclic mapping is vital for establishing a bidirectional relationship between the two domains, allowing the model to preserve essential characteristics of the input data. Enforcing cycle consistency objectives enables the model to understand the structure and features of each domain, even without direct correspondence between individual training samples. This method aids in learning meaningful mappings, contributing to successful, cycle-consistent domain translation.

Cycle consistency loss functions:

$$\mathcal{L}_{\text{cycle}}^{A \to B}(x, x'') = MSE(x, x'') = |x - x''|_1$$
$$\mathcal{L}_{\text{cycle}}^{B \to A}(y, y'') = MSE(y, y'') = |y - y''|_1$$

$\mathcal{L}_{adv}$ represents the adversarial loss, in this study we apply binary cross-entropy loss, λ_{adv} and λ_{cycle} are weighting coefficients to balance different components of the generator loss. $|\cdot|_1$ denotes the L1 norm, used to measure the discrepancy between images.

In our research, we incorporate ResNet (He et al., 2016) as the foundational architecture for the generator within the Cycle-GAN network, and employ Patch-Discriminator (Isola et al., 2017) as the discriminator. To balance the generated results and compute resource cost, the input network image size is transformed, and the images are pre-processed and scaled to 512×512 pixels during training and application. The generated result images are then scaled back to their original size.

4. Dataset and Model Training

We developed a web crawler to collect architectural scene and sketch photos, using keywords such as 'architectural scene' and 'architect sketch'. We then refined this dataset by eliminating duplicates, text-laden images, book photos, low-resolution images, and those with watermarks or conspicuous text. This process resulted in a final dataset comprising about 2900 architectural scene photos and 1600 sketches, forming an unpaired image-to-image dataset, as shown in Figure 4.

We configured our experiment with a maximum of 500,000 training epochs, performing validations at every 1,000th epoch. Training parameters yielding the lowest validation loss were saved, with the entire process taking approximately 100 hours. The lowest validation loss occurred at epoch 395,000, and we selected the model parameters from this epoch for subsequent analyses, as illustrated in the photo-to-sketch conversion examples in Figure 5.

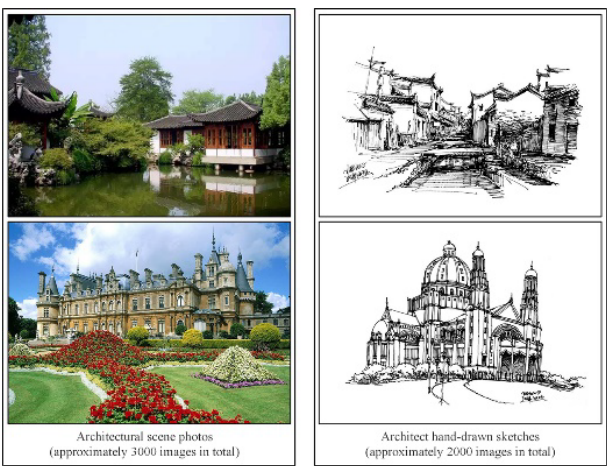


Figure 4. The data samples used in this paper

Figure 5. The architectural scene image input and the sketch image generated by the trained model

5. Experimental Results Compared with Existing Methods

This experiment focuses on training AI to interpret architects' specialized sketching language using data-driven methods. We conducted a comparative analysis, involving two distinct approaches. First, we replicated Lu's (2012) rule-based method, which uses image gradients and forward differences for edge extraction and style transfer. Second, we evaluated popular smartphone sketch generation apps, which are based on pre-training algorithms using non-domain-specific datasets.

In our replication of Lu's method using parameter sets A and B, we observed distinct results. Set A focused on outline strokes, while Set B emphasized image edges. However, both approaches predominantly concentrated on the fine details of architectural scene photos, rather than their overall content. In contrast, our method effectively mimics architect sketches in both strokes and lighting and uniquely interprets the image's meaning. For instance, in scenes with trees, lawns, or rivers, it sketches only the edges near buildings, leaving the rest blank. This demonstrates the AI's ability to process sensory semantic information in a manner akin to architects, as exemplified in Figure 6.

We also compared our method with popular sketch generation applications like 'My Sketch' and 'Photo Sketch Maker', which are widely used on iOS and Android platforms. These apps transform scene photos by applying computer graphics principles to aspects like colour, texture, and edges, with 'My Sketch' incorporating an additional simulated sketch stroke mask. Despite their popularity, these applications lack the capability to process semantic information. Conversely, our method not only captures stroke textures but also semantically condenses scenes, simplifying non-architectural elements such as backgrounds, and accentuating architectural details in a distinct manner, as demonstrated in Figure 7.

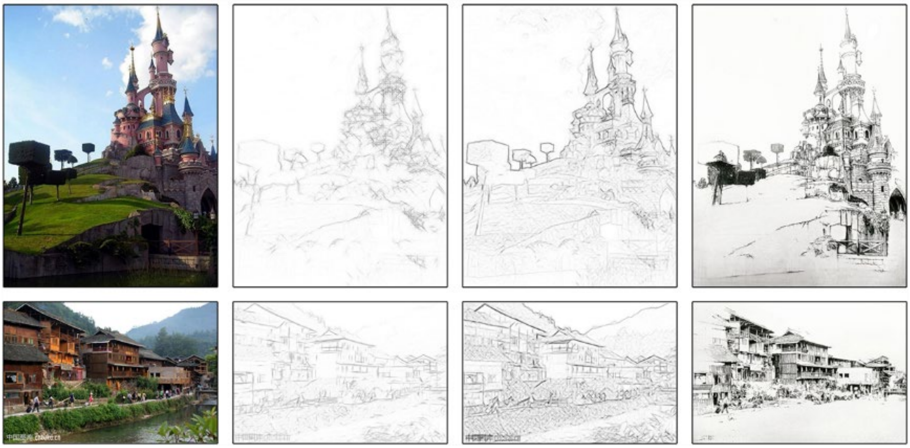


Figure 6. Comparison between Our Method and the Rule-based Method (Left 1: Original Image; Left 2: Parameter A; Left 3: Parameter B; Left 4: Our Method)

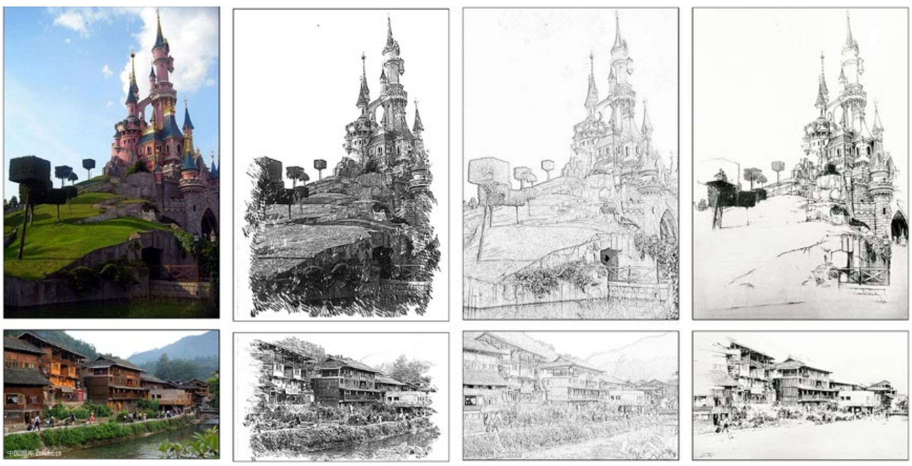


Figure 7. Comparison between Our Method and Apps (Left 1: Original Image; Left 2: My Sketch; Left 3: Photo Sketch Maker; Left 4: Our Method)

Comparative results, shown in Figure 8, reveal that our method more effectively replicates human architects' semantic processing than the compared apps. Our method intricately details key buildings, streamlines distant structures, and often leaves areas such as water surfaces blank. However, due to limited data volume and computational power, a discernible gap remains between our AI-generated results and genuine architect sketches. For example, our model tends to oversimplify elements like yachts, typically detailed in architects' sketches, due to their absence in our training dataset. We anticipate these limitations will lessen as the training dataset expands.

To substantiate the analysis of various methods, we conducted a quantitative survey with a matrix scale questionnaire. This questionnaire was tailored to assess factors architects deem crucial in sketches, like expressive strokes, emphasis variance, vividness, aesthetics, and artistic conception. Participants first viewed architectural scene photos, followed by images generated by My Sketch, the rule-based method, our method, and Photo Sketch Maker. They evaluated six aspects: Visual Appeal, Closeness to Human Stroke, Emphasis Variance, Vividness, and Artistic Conception, on a scale from 'very dissatisfied' (1) to 'very satisfied' (5).

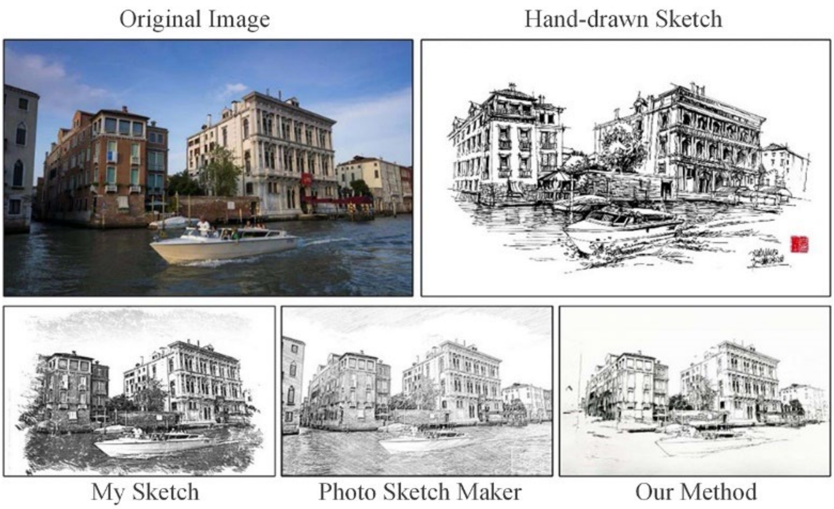


Figure 8. Comparison between Architect Sketch and Machine-generated Sketch

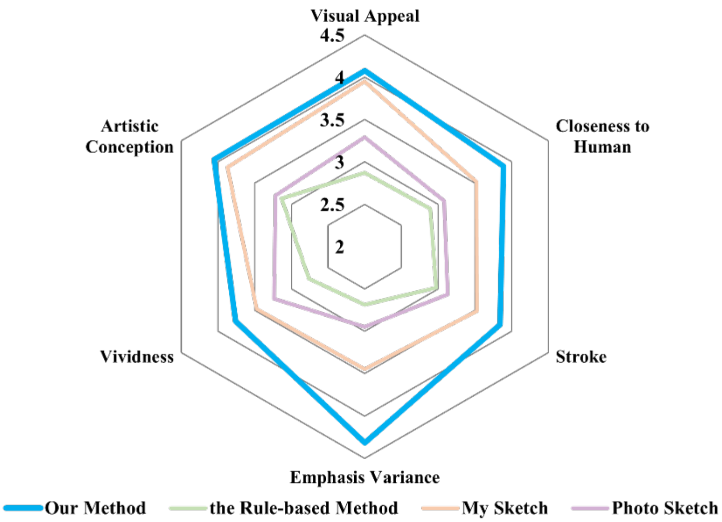


Figure 9. Rating Scores of Four Generated Results by Participants with Professional Background

The survey collected 50 valid matrix scale questionnaires, with 37 respondents from architecture/design backgrounds and 13 from non-architecture/design/fine arts. The evaluation results of participant-rated sketches, as shown in Figure 9, reveal that our proposed sketch generation method achieved the highest average scores across all metrics, especially in Emphasis Variance, among professionals, significantly surpassing existing methods. This highlights its ability to reflect architects' professional knowledge and skills.

Figure 10 displays the average ratings for each method. Participants with architecture or fine arts backgrounds favoured our method over existing ones. Conversely, non-professionals showed a preference for the visually polished sketches from 'My Sketch', with our method as a close second. This indicates that our method effectively captures the subtleties of professional architectural sketches while maintaining broad visual appeal.

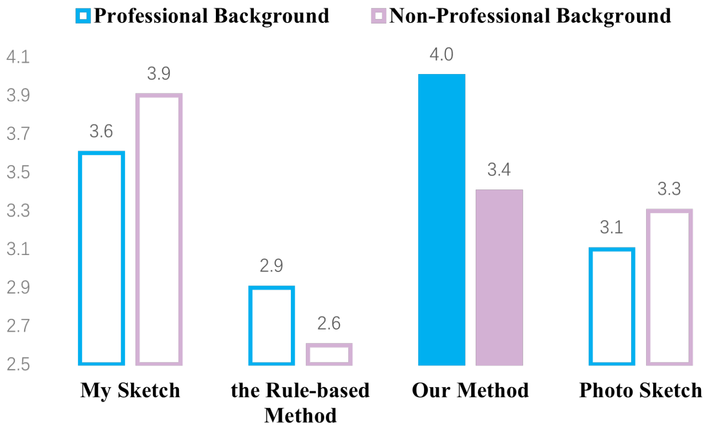


Figure 10. Mean Rating Scores of Four Results by Participants with Different Backgrounds

6. Conclusions

Architectural sketching, a natural language used by field experts, exhibits the proposed DSNL characteristics. We use data-driven deep learning methods to solve problems that computers have difficulty understanding naturalized domain-specific expressions. This involves collecting unpaired domain-specific datasets and employing relatively small neural networks, optimized for training with such datasets.

In this case study, we trained an AI algorithm to learn the sketching language used by architects by using a collected dataset of architectural sketches. The trained algorithm was then used to generate results that resemble hand-drawn sketches by architects. Compared to the results obtained by traditional methods, the results obtained by our proposed method include not only more visual information but also a selective process of semantic information, similar to how architects use their own judgment and aesthetic preferences to selectively depict architectural scenes in sketches, emphasizing certain aspects and differentiating between primary and secondary elements. The findings from the matrix scale survey also support the analysis of the experimental results.

Through this experiment, we have demonstrated that AI algorithms can be used to enable machines to use language that is more free-form, flexible, and inclusive of more ambiguity and generality, similar to that used by human domain experts, thus expanding the possibility space in meaning communication. This freedom and flexibility are crucial in the early stages of creative work, particularly in fields that require the use of DSNL.

References

Hu, W., Wang, X., Wang, D., Yao, S., Mao, Z., Li, L., Wang, F.-Y., & Lin, Y. (2023). Ir design for application-specific natural language: A case study on traffic data. arXiv preprint arXiv:2307.06983.

Fowler, M. (2010). Domain-specific languages. Addison-Wesley Professional.

Lawson, B. (2012). What designers know. Routledge.

CHEN Chi, L. C. (2020). Return with success: The influence of "collegiate" tradition on chinese academic tradition and architectural practice. Journal of World Architecture, No.356, 125–128. https://doi.org/10.16414/j.wa.2020.02.024

Lu, C., Xu, L., & Jia, J. (2012). Combining sketch and tone for pencil drawing production. Proceedings of the symposium on non-photorealistic animation and rendering, 65–73.

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

Ramesh, A., Dhariwal, P., Nichol, A., Chu, C., & Chen, M. (2022). Hierarchical text-conditional image generation with clip latents. arXiv preprint arXiv:2204.06125.

Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2021). High-resolution image synthesis with latent diffusion models.

Zhang, L., & Agrawala, M. (2023). Adding conditional control to text-to-image diffusion models. arXiv preprint arXiv:2302.05543.

Xie, S., & Tu, Z. (2015). Holistically-nested edge detection. Proceedings of the IEEE international conference on computer vision, 1395–1403.

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, 27.

Mirza, M., & Osindero, S. (2014). Conditional generative adversarial nets. arXiv preprint arXiv:1411.1784.

Zhu, J. -Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. Proceedings of the IEEE international conference on computer vision, 2223–2232.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition, 770–778.

Isola, P., Zhu, J.-Y., Zhou, T., & Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. Proceedings of the IEEE conference on computer vision and pattern recognition, 1125–1134.

FROM TOPOLOGY TO SPATIAL INFORMATION: A COMPUTATIONAL APPROACH FOR GENERATING RESIDENTIAL FLOORPLANS

MOHAMED EL MESAWY[1], NAWAL ZAHER[2] and AHMED EL ANTABLY[3]
[1,2,3]*Arab Academy for Science, Technology and Maritime Transport.*
[1]*m.elmesawy50@student.aast.edu, 0009-0005-2268-7928*
[2]*nawalzaher@aast.edu, 0000-0001-7984-8060*
[3]*ahmed.antably@aast.edu, 0000-0002-0151-0809*

Abstract. Multimodal models that combine different media like text, image, audio, and graph have revolutionised the architectural design process, which could provide automated solutions to assist the architects during the early design stages. Recent studies use Graph Neural Networks (GNNs) to learn topological information and Convolution Neural Networks (CNNs) to learn spatial information from floorplans. This paper proposes a deep learning multimodal model incorporating GNNs and the Stable Diffusion model to learn the floorplan's topological and spatial information. The authors trained a Stable Diffusion model on samples from the RPLAN dataset. They used graph embedding for conditional generation and experimented with three approaches to whole-graph embedding techniques. The proposed Stable Diffusion model maps the user input, a graph representing the room types and their relationships, to the output, the predicted floorplans, as a raster image. The Graph2Vec and contrastive learning methods generate superior representational capabilities and yield good and comparable results in both computationally derived scores and evaluations conducted by human assessors, compared to the Graph Encoder-CNN Decoder.

Keywords. Floorplan Generation, Deep Generative Models, Multimodal Machine Learning, Graph Neural Networks [GNNs], Representation Learning.

1. Introduction

In recent years, advancements in artificial intelligence (AI) have brought renewed attention to the role of Computer-Aided Architectural Design (CAAD). The use of AI in architectural design has proliferated from the early days of CAAD to the current visions of man-machine symbiosis. This growth relies on computer hardware and software advances and requires a shift in architects' design thinking paradigms. It necessitates collaboration among professionals from multiple domains, including architects, computer scientists, data scientists, and machine learning engineers. This

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 129-138.

paper proposes a deep-learning multimodal model incorporating GNNs and diffusion models to learn topological and spatial information from architectural floorplans. The aim is to develop a system that generates high-quality architectural floorplans to assist architects during the early design stages.

The significance of this paper lies in several key aspects. Firstly, the authors employ the whole graph embedding technique to capture and represent a floorplan's intricate structure and topology. This approach enables a comprehensive understanding of the floorplan spatial relationships and design elements. Secondly, the proposed use of the diffusion model for floorplan generation demonstrates its superior performance compared to Generative Adversarial Networks (GANs) regarding image generation quality. Lastly, this study encompasses multiple approaches to obtain whole-graph embeddings as an agnostic task. These embeddings have broader implications beyond floorplan generation, as they can be utilised in various downstream tasks such as graph classification and conditional generation. This versatility enhances our research's applicability and potential impact in diverse architectural and computational design domains.

2. AI and Floorplan Generation

The recent growth in the use of Artificial Neural Networks (ANNs) in computational design reflects the fast advancement in research in generative models (Dhondse A et al., 2020), the increase in computational power, and the availability of training datasets (Hodas & Stinis, 2018) such as RPLAN, CubiCasa5K, and CubiGraph5K. Moreover, using many computational algorithms and variants of neural networks based on graphs, such as Graph Convolutional Networks (GCN) (Carta, 2021), has led to considerable advancements in graph processing and generation. In general, algorithms based on graph theory result in quite an effective manipulation of data in spatial configurations. Designers and those trained in spatial abstraction usually find topological approaches intuitive, for they have data structures like bubble diagrams and direct applications to spatial organization. Wassim Jabi's work innovatively employs topological graphs using the Topologic plugin, presenting a notable advancement for building classification tasks (Alymani et al., 2023).

The literature may be divided into three categories of floorplan conditional generative methods based on the input type: pixel-based, language-based, and graph-based approaches. Moreover, generative deep learning models typically comprise two main components: the encoder and the decoder. Combining different modalities within the encoder and decoder exposes different generative models utilising different media types, each with advantages and disadvantages for floorplan generation tasks. Some approaches may employ a pre-processing step to convert the input from one type to another before feeding it into the model or even post-processing the output.

First, following a pixel-based approach, some researchers developed their model using the Pix2Pix model, which is a version of Generative Adversarial Networks (GANs) that takes an image representing the floorplan footprint as input and a floorplan raster image as output (Figure 1). The model consists of convolutional parts in both the encoder and decoder. This approach's advantage is utilising convolutional layers in the model decoder. CNNs are critical in preserving and leveraging spatial information during the model's training and prediction phases. They extract spatial information from an image by performing localised computations using convolution kernels, which generate floorplans that look like real images and make much sense from an architectural perspective. The limitation of this approach is that the user has no control over the floorplan program as it is solely conditioned on the floorplan footprint and the door location. Recently, Veloso et al. tried to address this limitation by converting the input graph with the room areas to a bubble diagram image, then dealing with the input bubble diagram image and output floorplan image as Pix2Pix (Veloso et al., 2022). However, while the input in the graph format has the floorplan topological information, the conversion of the graph into a pixelated image does not retain this information explicitly as in the graph format. As a result, the model may easily generate unrealistic designs (Figure 1).

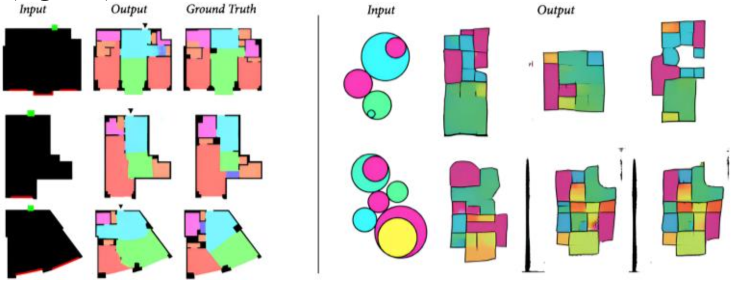


Figure 1. Left: Success example of using Pix2Pix in architectural floorplans (source: Chaillou, 2020). Right: Examples of overlapping discs after converting the graph and zones to an image (source: Veloso et al., 2022).

Second, in the graph-based approach, the user's input is a graph containing the room types and their edge relationships, fed to GNNs in the encoder part (Nauata et al., 2020, 2021). The generated output of this approach is a series of bounding boxes representing each room, which are later combined to generate the final floorplan during the post-processing phase. The advantage of this model is rooted in the GNN's ability to learn the underlying topology of the input graph. This model leverages the graph structure as a source of non-spatial information that captures the connectivity patterns between the nodes (rooms). However, this model's limitation lies in its decoder component, which treats the output as a regression task by predicting the bounding boxes of each room separately. Consequently, there is a risk of some rooms overlapping, resulting in generated floorplans that do not make architectural sense regarding rooms' dimensions and spatial qualities, as shown in the failure cases in Figure 2.

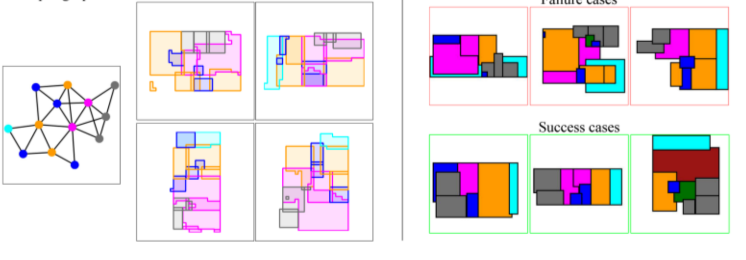


Figure 2. Left: Overlapping bounding boxes of rooms in House GAN. Right: Failure and success examples of the generation (source: Nauata et al., 2020).

Third, in the language-based approach, the language model encoder learns the mapping between the textual input and the corresponding floorplan generated by the decoder. The decoder may also predict bounding boxes (Galanos et al., 2023) (Figure 3) or generate images from the text as done in generative models such as Stable Diffusion, GAN text, Dall-2, and Mid-Journey. Nevertheless, the models that generate the room output as bounding boxes may lack spatial information.

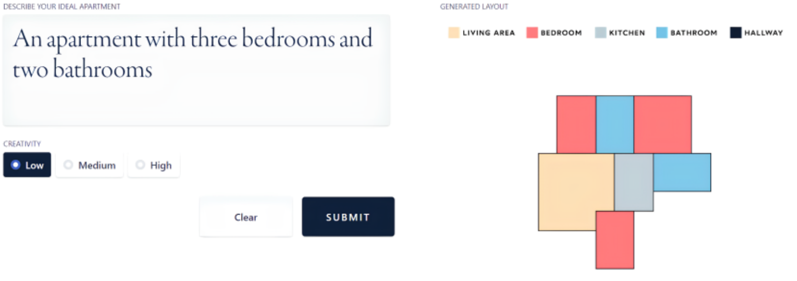


Figure 3. Architex user interface and its raw prediction (source: Ramesh et al., 2022).

To this end, this paper proposes a deep-learning multimodal model that combines the strengths of GNN and CNN in Stable Diffusion to generate a high-quality raster image of the floorplan. Specifically, it employs GNN in the encoder component to capture the graph input's topological non-spatial information and the decoder's convolutional layers to preserve the spatial characteristics of the generated floorplan. In other words, the paper replaces the text encoder part in the original Stable Diffusion model with a graph encoder pre-trained as whole-graph embedding.

We have approached the graph encoder as an architectural language model by training it as an agnostic task to learn the whole-graph embedding, which holds the floorplan semantics. The learned embeddings can then be utilised in downstream tasks such as graph classification or floorplan generation.

3. Methods

The paper uses a three-stage method: data preparation, generative model, and finally, post-processing.

3.1. DATASET PREPARATION AND PRE-PROCESSING

The RPLAN dataset (Wu et al., 2019) is a collection of real-world residential buildings containing over 80k floorplans with 13 room types (Hu et al., 2020). We split 1000 floorplans from the RPLAN dataset into 800 and 200 samples for the train and test sets, respectively. We adopted the same colouring labels and room mappings from the pre-processing stage (Rodrigues et al., 2021), as the colours have high contrast, which helps in the room's segmentation. We mapped the dataset room types to only six: public area, room, storage, kitchen, bathroom, and balcony. In the pre-processing step, we extracted the graph and its corresponding floorplan image from the raw multichannel floorplan file provided in the dataset. Figure 4 shows an example of the extracted geomatic information from the structured floorplan, which is stored in a serializable format to generate a graph that describes the room relation in a floorplan.

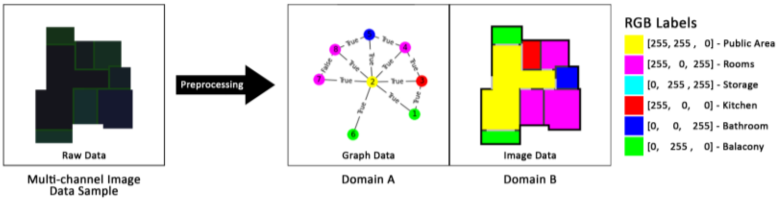


Figure 4. The pre-processing stage: extracting the graph and the floorplan image-RPLAN Dataset.

3.2. MODEL TRAINING

First, we trained the graph encoder as an agnostic task to get the whole-graph embeddings. Then, we used these embeddings in the downstream task of floorplan generation. The whole-graph embeddings represent graph structures in a lower-dimensional space (Figure 5). The aim is to encode the graphs (reduced bubble diagram) such that the similarity in the embedding space approximates the similarity in the original network. After separately training the graph encoder, we replaced the text encoder in the original Stable Diffusion model with the pre-trained graph encoder. Finally, we tested three implementations for the graph encoder in conjunction with the Stable Diffusion model, where we trained the diffusion model from scratch with each graph encoder. We discuss the graph encoders in the following sections.

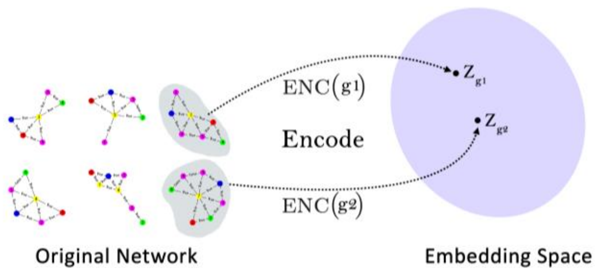


Figure 5. Representation learning on networks: whole graph embedding.

The first graph encoding method is Graph2vec (Narayanan et al., 2017) for whole-graph embedding. This method has received attention recently due to its ability to capture graph structure information and generate embeddings usable in various downstream tasks. The method uses a skip-gram model to learn vector representations of subgraphs from a given graph. These vector representations are combined to generate an embedding for the entire graph (Figure 5). We used the off-the-shelf Graph2vec algorithm implemented in Karate Club (Rozemberczki et al., 2020).

The second approach is borrowed from the Contrastive Language-Image Pre-training (CLIP) model (Radford et al., 2021). One of the key features of CLIP is its use of contrastive learning, a technique that allows the model to learn by contrasting similar and dissimilar pairs of data, text, and images. Using a contrastive loss function, CLIP utilises a transformer-based model pre-trained on a large corpus of images and text. The model may then be finetuned for specific downstream tasks, such as image classification, object detection, and image captioning. For this paper, we applied the concept of contrastive learning by replacing the text encoder with a graph encoder and used a pre-trained image encoder, which we trained first within a convolutional autoencoder model Figure 6. In the next training step, only the graph encoder is learnable by contrastive learning between the graph and its corresponding floorplan raster image, as shown in Figure 7.

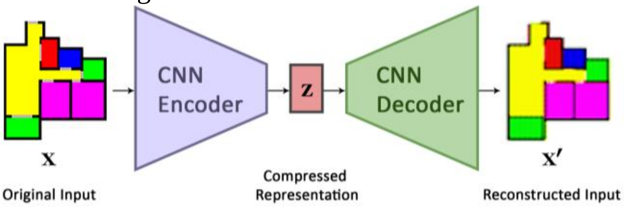


Figure 6. Learning phase of the CNN encoder and decoder components using a convolutional autoencoder model.

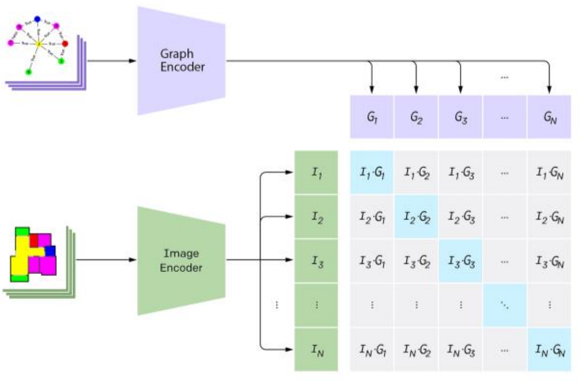


Figure 7. Contrastive learning between graph and images (adapted from: Radford et al., 2021)

The third method for training the graph encoder is the Graph Encoder-CNN Decoder. This method replaces only the convolution encoder part from the convolutional autoencoder model (Figure 6) with a graph encoder. The original CNN decoder is trained with a graph encoder consisting of stacked GNN layers, particularly Graph Sage (Hamilton et al., 2018) (Figure 8). The model learns the representation of the graph input data by compressing it into a lower-dimensional latent space and then constructing the corresponding floorplan image from this compressed representation. The model is trained by minimising the difference between input and constructed data. This model could be used as an end-to-end approach to generate the floorplan image from an input graph. However, we only used the trained graph encoder part in the downstream floorplan generation task.

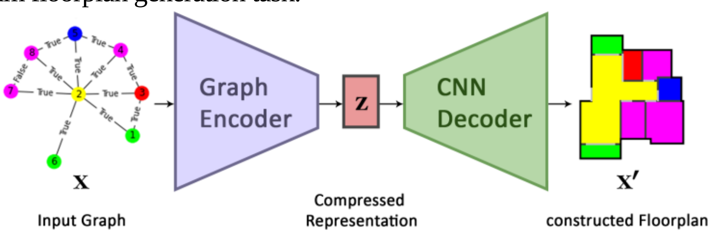


Figure 8. Graph encoder/CNN decoder.

3.3. VISUALISING THE GRAPH EMBEDDINGS

After training the graph encoder separately to learn the embeddings of the graphs corresponding to the floorplans in the training set, we visualised the output graph embeddings with the corresponding floorplan images using the Embedding Projector in Tensorboard. To visually assess the effectiveness of the learned embeddings, it is possible to randomly select a sample of floorplans and evaluate the K-nearest similar plans according to the graph embeddings. An effective graph encoder model should embed similar graphs with similar embeddings to be closer and dissimilar graphs farther apart in the latent space.

3.4. DIFFUSION MODEL TRAINING

We trained the Stable Diffusion model on a single NVidia P100 GPU using PyTorch implementation. Our Stable Diffusion model consists of the pre-trained graph encoder and the image generator component (Figure 9). The training parameters of the Stable Diffusion model are 1000 noising steps, an image size of 64x64 pixels for faster training, training epochs of 1500 epochs with a batch size of 8, and a learning rate of 0.0003. Each of the three trained models took 12 hours to train with the parameters mentioned above.

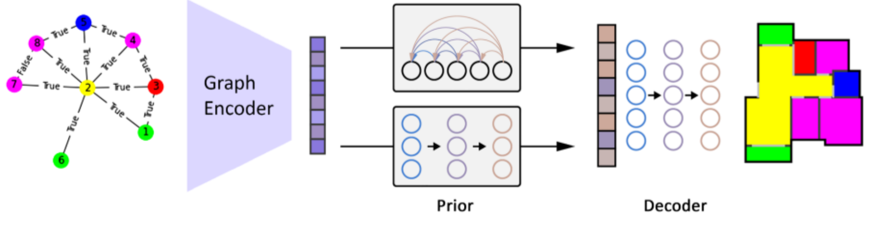


Figure 9. Proposed Stable Diffusion: graph-image mapping (adapted from: Ramesh et al., 2022).

3.5. POST-PROCESSING

We employed Real-ESRGAN (Wang et al., 2021), a super-resolution model, to enhance the generated lower-resolution images by up-sampling the output four times from 64x64 pixels to 256x256 pixels. This step allowed us to obtain higher-quality floorplan images without having to perform training on larger images in the Stable Diffusion model (Figure 10).

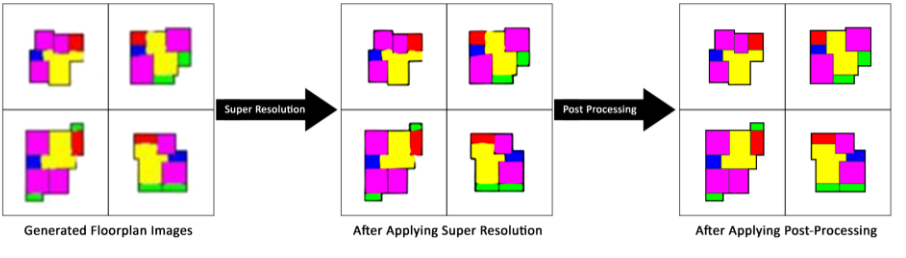


Figure 10. The post-processing step using the super-resolution model to enhance the quality of the generated image, then applying wall alignment.

4. Results and Discussion

We evaluated the three versions of the generative models, denoted as Model A, Model B, and Model C, using three metrics focusing on image generation performance: Fréchet Inception Distance (FID), Intersection over Union (IoU) and survey questionnaire.

FID is a metric used to evaluate the quality of generated images by comparing them to a set of real images. It is a widely used measure in the field of GANs and generative models in general (Heusel et al., 2018). A lower FID score indicates that the generated images are of high quality and diversity. However, an FID score for architectural floorplan generation does not consider essential architectural criteria, such as the conditional program of the input. Hence, we developed a method for calculating the Intersection over Union (IoU) between the program of the output-generated floorplan and the ground truth. The analysis of variance (ANOVA) test, applied to the IoU metric, indicated a significant difference among the models ($F=51.95$, $p<0.001$). Subsequently, we used a post-hoc analysis to explore specific group differences. The post-hoc analysis, with all p-values less than 0.001, revealed significant differences between all pairs of models, reinforcing the ANOVA findings.

Besides the above computational methods, we developed a survey questionnaire to assess the quality of the generated floorplans from a human perspective. We used a form containing 15 samples of generated floorplans, five from each model in random order, populating 60 random forms, each with 15 different floorplans selected from the test set. Sixty architectural students responded to the survey. Employing repeated measures ANOVA to analyse students' responses indicated a statistically significant ($F=67.73$, $p<0.001$) difference between the models. A subsequent post-hoc analysis ($p<0.001$) showed a significant difference between Models A and C, as well as B and C, but none between A and B.

Table 1: compares the FID, IoU and the survey questionnaire for each model on the test set.

Model	Graph Encoder	Graph Enc. Training Time	SD Training Time	FID	IoU		Human Eval.	
					mean	std	mean	std
A	Graph2Vec	1 min.	12 hours	66.9	0.88	0.12	3.60	1.19
B	Contrastive Learning	8 hours	12 hours	72.1	0.8	0.12	3.70	1.18
C	Graph Enc. - CNN Dec.	12 hours	12 hours	134.5	0.69	0.28	2.35	1.3

From a qualitative point of view, the Stable Diffusion model generates better floorplans in terms of the rooms' architectural qualities when conditioned with a pre-trained whole-graph embedding (Figure 11). This paper's Graph2Vec and contrastive learning methods generate superior representational capabilities compared to alternative graph encoders and have close scores across the FID, IoU, and human evaluation. On the other hand, the Graph Encoder-CNN Decoder produced the least favourable results in terms of FID, IoU, and human visual inspection. Interestingly, the computational scores are close to human evaluation in terms of generation quality.

Generally, three factors influence the quality of the generated output: the quality of the pre-trained graph embeddings, the length of the embedding vectors, and the generative model itself. First, the quality of the pre-trained graph embeddings is a limitation of the Stable Diffusion generation. In this aspect, the Graph2Vec and the graph encoder based on contrastive learning outperform the third method, hence generating floorplans from the Stable Diffusion model that are highly similar to the ground truth. The second limitation is the length of the embedding vectors. Generally, higher dimensional latent space can capture more information to some extent. Last, the generative model itself, especially the design of the loss function, penalises the model for generating outputs that lack architectural coherence. Future investigations may consider diversifying the types of datasets used for training to enhance the model's adaptability and exploring the model's applicability in non-residential or diverse cultural contexts.

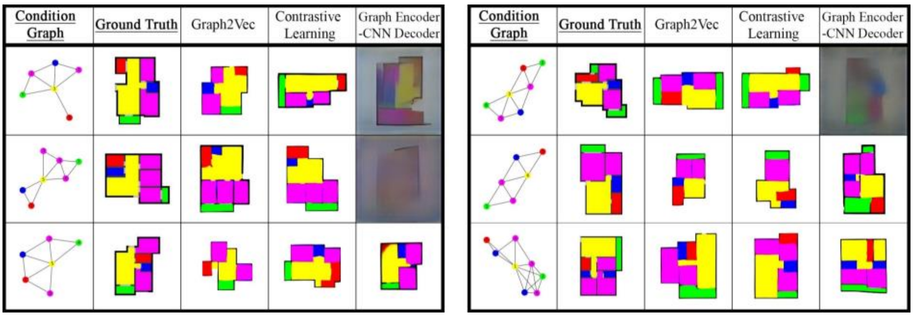


Figure 11. A comparison between the generated floorplans and the ground truth for the same input graph using the three proposed methods.

5. Conclusion

This research proposes a deep-learning multimodal architecture that uses GNN and the Stable Diffusion model to learn the topological and spatial information of architectural floorplans. Eight hundred samples from the RPLAN dataset were used for training and 200 for testing. Three different evaluation metrics show that our proposed method, which replaces the text encoder with a pre-trained graph encoder, generates high-quality architectural floorplans. However, the quality of the graph embeddings, the length of the embedding vectors, and the architecture of the generative model influence the quality of the generated output. The Graph2Vec and the graph encoder based on contrastive learning methods produced superior representations compared to the third graph encoder. Overall, the proposed method has significant potential to assist architects in the early design stages, enabling them to select the most suitable floorplan for their design needs.

References

Alymani, A., Jabi, W., & Corcoran, P. (2023). Graph machine learning classification using architectural 3D topological models. SIMULATION, 99(11), 1117–1131. https://doi.org/10.1177/00375497221105894

Carta, S. (2021). Self-Organizing Floor Plans. Harvard Data Science Review. https://doi.org/10.1162/99608f92.e5f9a0c7

Chaillou, S. (2020). ArchiGAN: Artificial Intelligence x Architecture. In P. F. Yuan, M. Xie, N. Leach, J. Yao, & X. Wang (Eds.), Architectural Intelligence (pp. 117–127). Springer Nature Singapore. https://doi.org/10.1007/978-981-15-6568-7_8

Dhondse A, Kulkarni S, Khadilkar K, Kane I, Chavan S, Barhate R, & 3rd International Conference on Data Management, A. and I., ICDMAI 2019. (2020). Generative Adversarial Networks as an Advancement in 2D to 3D Reconstruction Techniques. Adv. Intell. Sys. Comput. Advances in Intelligent Systems and Computing, 1016, 343–364.

Galanos, T., Liapis, A., & Yannakakis, G. N. (2023). Architext: Language-Driven Generative Architecture Design (arXiv:2303.07519). arXiv. http://arxiv.org/abs/2303.07519

Hamilton, W. L., Ying, R., & Leskovec, J. (2018). Inductive Representation Learning on Large Graphs (arXiv:1706.02216). arXiv. http://arxiv.org/abs/1706.02216

Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., & Hochreiter, S. (2018). GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium (arXiv:1706.08500). arXiv. http://arxiv.org/abs/1706.08500

Hodas, N. O., & Stinis, P. (2018). Doing the Impossible: Why Neural Networks Can Be Trained at All. Frontiers in Psychology, 9, 1185. https://doi.org/10.3389/fpsyg.2018.01185

Hu, R., Huang, Z., Tang, Y., Van Kaick, O., Zhang, H., & Huang, H. (2020). Graph2Plan: Learning floorplan generation from layout graphs. ACM Transactions on Graphics, 39(4). https://doi.org/10.1145/3386569.3392391

Narayanan, A., Chandramohan, M., Venkatesan, R., Chen, L., Liu, Y., & Jaiswal, S. (2017). graph2vec: Learning Distributed Representations of Graphs (arXiv:1707.05005). arXiv. http://arxiv.org/abs/1707.05005

Nauata, N., Chang, K.-H., Cheng, C.-Y., Mori, G., & Furukawa, Y. (2020). House-GAN: Relational Generative Adversarial Networks for Graph-constrained House Layout Generation. arXiv:2003.06988 [Cs]. http://arxiv.org/abs/2003.06988

Nauata, N., Hosseini, S., Chang, K.-H., Chu, H., Cheng, C.-Y., & Furukawa, Y. (2021). 02_House-GAN++: Generative Adversarial Layout Refinement Networks. arXiv:2103.02574 [Cs]. http://arxiv.org/abs/2103.02574

Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, G., & Sutskever, I. (2021). Learning Transferable Visual Models From Natural Language Supervision (arXiv:2103.00020). arXiv. http://arxiv.org/abs/2103.00020

Ramesh, A., Dhariwal, P., Nichol, A., Chu, C., & Chen, M. (2022). Hierarchical Text-Conditional Image Generation with CLIP Latents (arXiv:2204.06125). arXiv. http://arxiv.org/abs/2204.06125

Rodrigues, R. C., Imagawa, M. K., Koga, R. R., & Duarte, R. B. (2021). Big Data vs Smart Data on the Generation of Floor Plans with Deep Learning. Blucher Design Proceedings, 217–228. https://doi.org/10.5151/sigradi2021-114

Rozemberczki, B., Kiss, O., & Sarkar, R. (2020). Karate Club: An API Oriented Open-source Python Framework for Unsupervised Learning on Graphs (arXiv:2003.04819). arXiv. http://arxiv.org/abs/2003.04819

Veloso, P., Rhee, J., Bidgoli, A., & Ladron de Guevara, M. (2022). Bubble2Floor: A Pedagogical Experience With Deep Learning for Floor Plan Generation. 373–382. https://doi.org/10.52842/conf.caadria.2022.1.373

Wang, X., Xie, L., Dong, C., & Shan, Y. (2021). Real-ESRGAN: Training Real-World Blind Super-Resolution with Pure Synthetic Data (arXiv:2107.10833). arXiv. http://arxiv.org/abs/2107.10833

Wu, W., Fu, X.-M., Tang, R., Wang, Y., Qi, Y.-H., & Liu, L. (2019). Data-driven interior plan generation for residential buildings. ACM Transactions on Graphics, 38(6), 1–12. https://doi.org/10.1145/3355089.3356556

GRAPH2PIX:

A Generative Model for Converting Room Adjacency Relationships into Layout Images

ZHEN HAN[1], XIAOQIAN LI[2], YE YUAN[3] and RUDI STOUFFS[4]
[1,2,3]*School of Architecture, Tianjin University*
[1,2,4]*Department of Architecture, National University of Singapore*
[1]*zhen.han@u.nus.edu, 0000-0002-9547-9063*
[2]*lixiaoqian_95@tju.edu.cn, 0009-0009-0160-6817*
[3]*yyarch@tju.edu.cn, 0000-0002-2277-5806*
[4]*stouffs@nus.edu.sg, 0000-0002-4200-5833*

Abstract. With the advancement of mathematics and computer science, deep learning-based generative design for floorplans has increasingly garnered attention among researchers. This study proposes a graph-based deep learning model, graph2pix (G2P) to synthesize floorplans guided by user-defined constraints. By incorporating room area and type information into the nodes of the graph, G2P can generate floorplans tailored to specific user requirements. It contains three sub-models: the Translator, Generator, and Discriminator. The Translator serves as the foundational layer, interpreting and mapping room information into a coherent building boundary. Following this, the Generator takes the helm, synthesizing this information to formulate a preliminary floorplan layout. This layout is further refined and evaluated by the Discriminator, ensuring that the final output maintains a high degree of fidelity both to the user's constraints and to architectural feasibility. Our empirical investigation, utilizing metrics such as "area error" and "adjacency error", underscores the model's exceptional ability to achieve user tasks with a high degree of accuracy and efficiency. These findings underscore the potential of G2P as a transformative tool in the domain of automated architectural design.

Keywords. Floor Plan Generative Design, Graph to Image Generative, Graph Neural Network, Room Information Addition, Conditional Generative Adversarial Network.

1. Introduction

The architectural design process is an intricate and labour-intensive endeavour, governed by a detailed design brief that provides essential information. This necessitates architects to engage in planning of various architectural elements including room layouts, floor plans, and façade designs. Equipped with foundational data such as the types and areas of rooms, architects are tasked with the precise delineation of the size, shape, and spatial relationships of each room. These elements are strategically arranged to create a layout that is both coherent and functional, ensuring the alignment of architectural design with the envisaged utility and aesthetic goals. In the early stages of architectural design, architects are often required to generate multiple initial proposals based on the requirements of a detailed design brief. Subsequently, through a process of meticulous refinement and modification, a selection of these proposals is chosen for further detailed design.

In the realm of generative design assisted by computational methods, researchers typically encode established norms and principles into algorithmic directives, enabling Computer-Aided Design (CAD) systems to generate architectural layouts that conform to recognized best practices. Utilizing computer's advanced computational power enables rapid provision of numerous design options for architects, significantly boosting design efficiency. In recent years, with the advancement of artificial intelligence, and more specifically, AI-generated content, there has been a growing interest in using deep learning algorithms for the generation of architectural floorplans. Researchers have employed Generative Adversarial Networks (GANs) to create building floorplans based on architectural outlines (Wu et al., 2019), (Hu et al., 2020), (Sun et al., 2022). Although this approach can produce highly detailed models, it diverges from the traditional design practices and workflows followed by architects. Architects typically arrange and combine rooms based on their area and interconnections, rather than first delineating an outline, and then filling in different rooms accordingly. Therefore, deep learning models for floorplan generation should ideally only incorporate the most fundamental requirement information as input: number of rooms, room types, room areas, etc. Models that use architecturally processed information like building outlines as input may not effectively alleviate the design burden of architects.

To address these challenges, some researchers have shifted to using different deep learning models, employing more primitive information such as architectural function bubble diagrams or adjacency graphs, rather than architectural outlines, for floorplan generation. Figure 1 shows the input and output of the deep learning models built by different researchers.

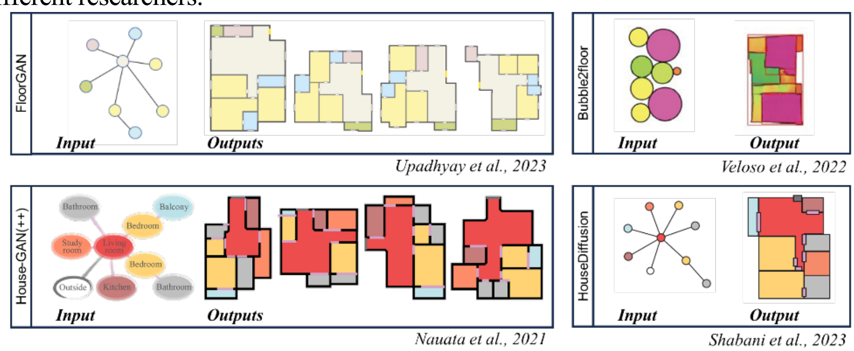


Figure 1. Inputs and outputs of deep learning models used in related literature

As shown in Figure 1, Veloso et al. (2022) continued to utilize the pix2pix framework, transforming architectural function bubble diagrams into images to serve as model inputs for generating floorplans. These diagrams represent the intended function and approximate size of each room, allowing architects to provide input without having to draw precise outlines; they only need to simply sketch approximate room locations to obtain feasible floor layouts. Nauata et al. (2020 & 2021) represented rooms adjacency information as graphs and employed Graph Neural Networks (GNNs) to generate architectural floorplans (vector floorplan) that comply with the constraints of these adjacencies, further reducing the workload of architects and enabling the generated floorplans to reflect the connectivity between rooms. Upadhyay et al. (2023) combined GNNs with GANs to develop a deep learning model that also takes rooms adjacency information as inputs and produces architectural floorplan images (in image format). Shabani et al. (2023) proposed a method for generating vector floor diagrams through diffusion models, utilizing a transformer architecture. This architecture controls attention masks based on input graphic constraints and generates architectural floorplans (vector floor diagrams) directly through a process of discrete and continuous denoising.

This study builds upon the foundation of room adjacency information based on graphs by embedding additional information such as room area and room type into the nodes of the graph. Compared to existing studies, this research introduces a method that ensures generated floor plans simultaneously meet adjacency relationships and area requirements, addressing the issue of existing models' inability to quantitatively control the generated room areas. Furthermore, this approach obviates the need for architects to input building outlines, better aligning with the design process. Consequently, this enhancement significantly improves the applicability of the model. The resulting floorplans are more controllable and more closely aligned with the needs of architects, reducing the workload associated with design processes.

2. Problem Formulation

In the standard architectural design process, clients provide architects with a brief containing specific design parameters, such as total building area, types and numbers of rooms, and the area of each room. Initially, architects create function bubble diagrams based on the client's requirements. Subsequently, each 'bubble' is assigned area information and transformed into various shapes. Finally, through a process of design exploration and iteration, these are converted into architectural floorplans. To comprehensively incorporate the requirements detailed in the design brief into the deep learning model, enabling controlled generation of floorplans with respect to area and room adjacency relationships, this study amalgamates the concepts of GNNs and GANs. A novel deep learning architecture is proposed, encompassing Translator, Generator, and Discriminator components, aimed at accomplishing production-level tasks.

INPUT: The input to the model is a building bubble diagram, which is represented as a graph where a node encodes a room with its room type and room area. Room adjacency is depicted by edges connecting the respective nodes, providing a graph-based representation of the floorplan's layout. The building's total area and the building ratio, which is the footprint area as a proportion of the total lot area, are also provided to inform the generative model of the overall scale and density of the structure (shown in Figure 2). The model input is formalized as three matrices: a feature matrix X encoding room areas and types, a sparse adjacency matrix A representing room connectivity, and a global matrix G detailing the building's area and aspect ratio. These matrices provide a structured representation for the generative model to produce building floorplans.

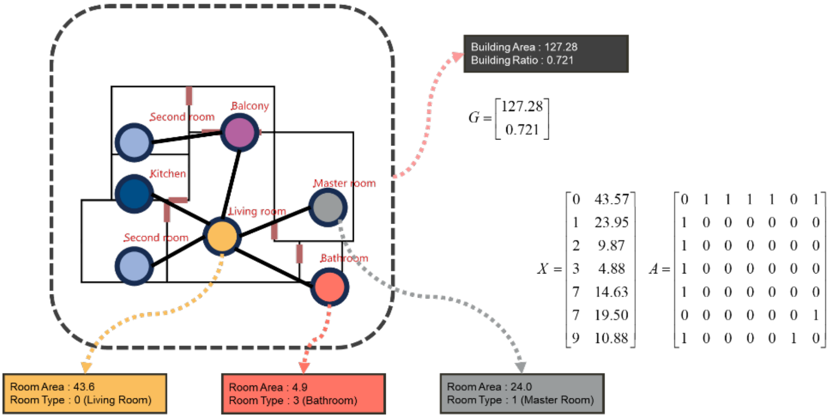


Figure 2. The information contained in the input and its expression

OUTPUT: The output of the proposed model is a detailed architectural floorplan, rendered as a 256x256 pixel image. In this discretized spatial representation, distinct colours correspond to different room types, facilitating immediate visual discrimination of the various functional areas within the building. This color-coding scheme is predefined (shown in Figure 3) and consistent across all generated floorplan images, ensuring uniformity and ease of interpretation. The resolution of the output image has been chosen to balance the need for detail against computational efficiency, providing sufficient granularity to discern room boundaries and configurations while maintaining a manageable image size for model processing.

Figure 3. Color-coding scheme (dataset from RPLAN: Wu et al., 2019)

3. Technical Innovations

In this study, the input contains room and building information, whereas the output is a pixelated image representation. Given the variability in room count per building, the input matrices are of variable dimensions, a format not inherently accommodated by traditional GANs. As shown in Figure 4, to address this, this study introduces a Graph Neural Network-GAN (GNN-GAN) architecture that facilitates the mapping from graphs of arbitrary size to structured images. Also, it is noteworthy that the input and output exhibit substantial differences in scale, a factor that could degrade the quality of the generated images due to insufficient feature extraction, potentially resulting in outputs that do not meet to the input constraints.

To address these challenges, this study introduces a novel network architecture, named "Translator", designed to encode room and building information and predict the corner points of the layout boundary. Subsequently, these predicted boundaries, along with the encoded information, are fed as conditional inputs to both the Generator and Discriminator of a GAN framework. Experimental results indicate that this hierarchical, phased output approach effectively mitigates issues stemming from the significant disparity in data scale between inputs and outputs.

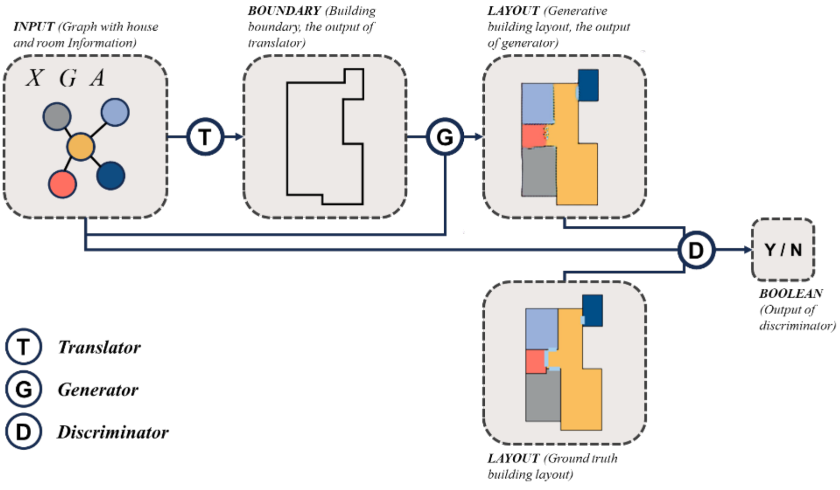


Figure 4. Overall framework of the proposed GNN-GAN model

3.1. FRAMEWORK OF THE TRANSLATOR

As shown in Figure 5, the Translator (T) is composed of an encoder and multiple decoders. The encoder, leveraging a graph attention mechanism coupled with Long Short-Term Memory networks (LSTM), processes the input room and building information (X, A, G) to generate a coded information. This information, along with the LSTM's hidden state, is then fed into the decoders, as illustrated in the left part of Figure 6. The decoding process is iterative and recursive, with each decoder outputting the position of a single vertex of the building contour until the contour is complete. A standard decoder unit, as shown in the lower right of Figure 6, includes an LSTM layer that receives the encoded information and hidden state from the encoder, as well as the prediction from the preceding decoder (with the initial decoder receiving a zero value).

The point set outputted by the Translator is utilized to render a 256x256x1 binary image, with pixel values of 0 or 1, delineating the architectural contour. Concurrently, an Intersection over Union (IoU) Loss, calculated between the generated contour and the ground truth, is employed as the loss function for the Translator, enabling the optimization of the network towards precise contour delineation.

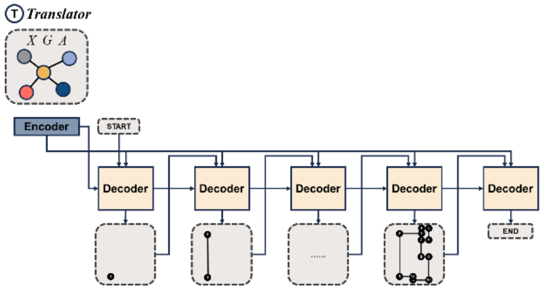


Figure 5. Framework of the Translator

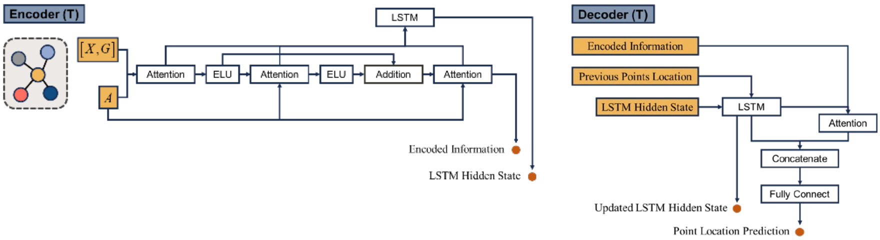


Figure 6. Framework of the encoder and decoder in the Translator

3.2. FRAMEWORK OF THE GENERATOR AND DISCRIMINATOR

Figure 7 illustrates the framework of the Generator (G) and Discriminator (D) within our proposed GAN model. Both G and D are equipped with an encoder-decoder structure. For the Generator, its encoder processes the input room and building information (X, A, G), encoding them into a compressed latent space. The encoded information is then passed through to the decoder, which will also receive the 256*256*1 building boundary image. In the case of the G, the decoder's task is to construct a plausible building layout, which takes the form of a 256*256*3 RGN image representing the floorplan. For the D, the decoder assesses whether the generated layout is authentic (Y) or not (N), based on the learned representations. This adjudication is predicated on discriminative features learned during training, enabling the Discriminator to effectively differentiate between genuine and synthesized designs.

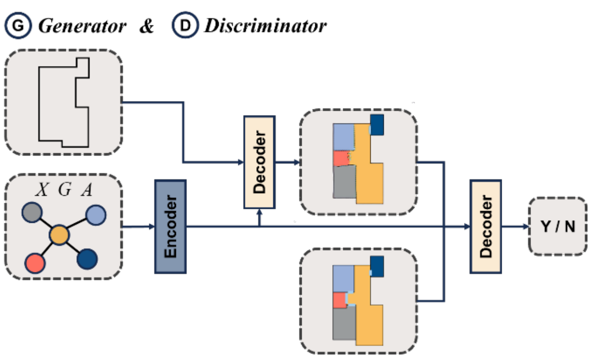


Figure 7. Framework of the Generator and Discriminator

It is important to note that both the G and D leverage a shared parameterization within their respective encoders. This parameter sharing is strategic, ensuring that both G and D utilize identical feature representations derived from the input data. By employing this architecture, we enhance the Discriminator's capacity to discern between genuine and synthesized designs. The shared encoder framework not only facilitates a more efficient learning process by reducing the number of free parameters within the system but also encourages a more nuanced feature extraction, leading to improved generalization of the Discriminator. This shared parameterization underscores the synergetic learning dynamics between G and D, integral to the adversarial training regimen that underpins the GAN's ability to generate compelling architectural layouts.

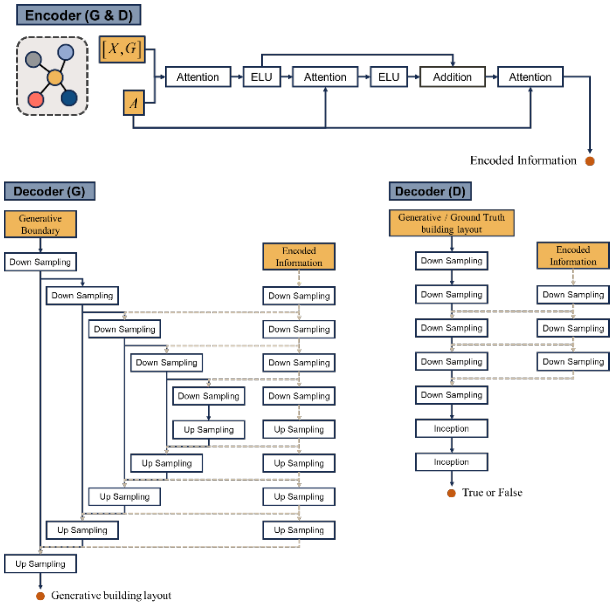


Figure 8. Framework of the encoder and decoder in Generator and Discriminator

Analogous to the Translator (T) network, the encoders within both the G and D incorporate a graph attention-based network structure. This design enables the encoding of room and building information into a latent space, effectively capturing the complex interdependencies and features present in the input data. However, distinct from T, the encoding processes in G and D do not involve sequential data output, and therefore, their architectures omit the LSTM layers. The exclusion of LSTM layers from G and D's encoders is aligned with the nature of their decoding tasks, which do not require processing temporal or ordered data, thus streamlining the architecture for the specific demands of generating and discriminating static images (shown in Figure 8 upper part).

Figure 8 also showcases the distinct frameworks of the decoders for the G and D, positioned in the lower left and lower right of the figure, respectively. The decoder of G is structured as a multimodal U-net, which utilizes the building boundary generated by the Translator (T) along with the encoded information as inputs to synthesize the floorplan. On the other hand, the decoder of D adopts a multimodal Patchnet design, which processes the encoded information in conjunction with either the generated floorplan from G or the ground truth data to execute its adjudication function. This configuration allows D to determine the veracity of the floorplans by examining localized patches, providing a granular assessment of authenticity.

4. Experimental Result

This study uses MATLAB for implementation and a workstation with Xeon CPUs and Nvidia Tesla V100-32GB. The GNN-GAN model uses ADAM optimizer and is trained for 120k iterations. The learning rates of the translator, generator and discriminator are 0.0001, 0.00001, 0.00001, respectively. The batch size is 10. And the model uses leaky-ReLUs ($\alpha = 1$) for all activate functions.

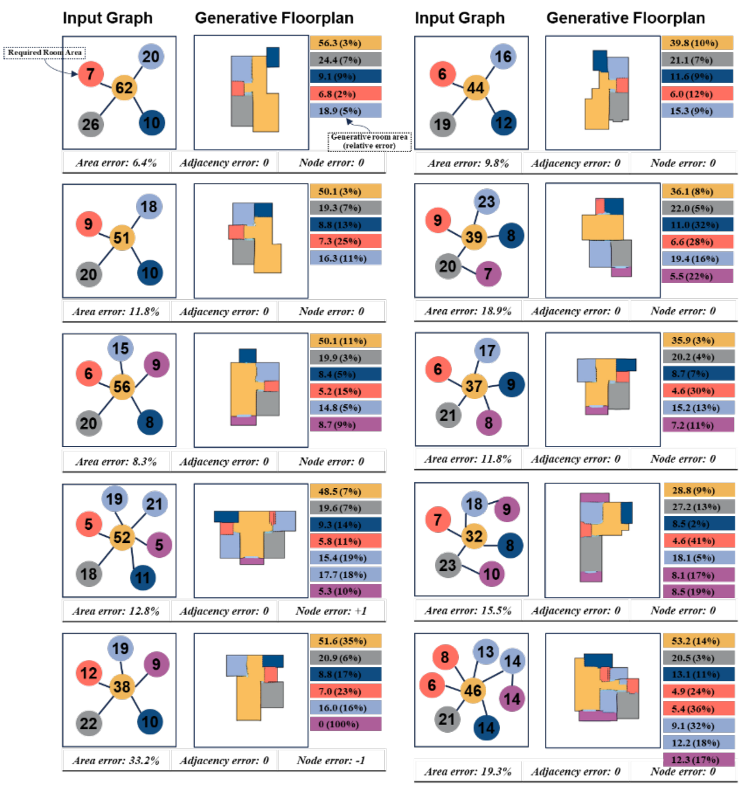


Figure 9. Framework of the encoder and decoder in Generator and Discriminator

"Area error", "Adjacency error" and "Node error" are used to evaluate the performance of the proposed model. The "Area error" represents the relative error between the input demanded area and the generated area. It serves as a measure of the model's precision in conforming to the spatial dimensions dictated by the design criteria:

$$Area\ error = \sum_{i=1}^{n} \left(\left| \hat{R}_i - R_i \right| / R_i \right) / n \tag{1}$$

where n is the number of rooms, R_i is the area of room i, $\hat{R}_i$ is the area of generative room i. The "Adjacency error" and the "Node error", respectively, count the number of rooms in the produced floorplan that fail to comply with the adjacency matrix constraints and the number of missing or redundant rooms, reflecting the model's fidelity to the prescribed room-to-room spatial relationships.

In figure 9, the model demonstrates the capability to generate diverse floorplans tailored to varying requirements. It not only adheres to the specified adjacency relationships among rooms but also responds adaptively to the area demands of individual rooms. Furthermore, the model exhibits a nuanced approach to floorplan generation; for inputs with identical adjacency relationships but differing room areas, it produces distinct floorplans with unique configurations, rather than resorting to mere scaling of room dimensions. This indicates the model's sophisticated understanding of spatial constraints and its ability to generate architecturally viable layouts beyond simple geometric transformations. The ability to produce varied layouts from the same topological constraints underscores the model's potential utility in architectural design, offering a tool that can accommodate a broad spectrum of design scenarios and site-specific conditions.

5. Conclusion

In this study, we propose a novel deep learning model for floorplan generation that takes room adjacency, room area, building area, building ratio as input variables. Departing from traditional pix2pix frameworks that rely on building contours as input, we establish an innovative Graph Neural Network-Generative Adversarial Network (GNN-GAN) architecture. We introduce the "Translator", a component designed for generating building boundaries, effectively addressing the issues of image distortion and non-compliance with input constraints due to the significant magnitude discrepancy between input (building information) and output (floorplan).

Our model advances the state of the art by embedding information such as room area and type into the graph nodes, which permits precise control over room dimensions, aligning the generated floorplan more closely with architects' specifications. Furthermore, by incorporating building area and aspect ratio into the global information, architects can manipulate the overall shape of the generated floorplan to better suit site-specific requirements. Experimental validation confirms the efficacy of our proposed model, demonstrating its ability to fulfil professional demands in architectural design with heightened accuracy and relevance.

Despite the advancements introduced by the proposed model, there remain certain limitations that merit attention. Notably, the generated floorplans occasionally lack detail, particularly in the articulation of door placements, which are not rendered with sufficient precision. Additionally, there are instances where the quantity and types of rooms produced do not align with the specified input constraints. These discrepancies highlight areas for further development. Future research endeavours will focus on refining the granularity of the output, enhancing the sophistication of the model's ability to represent minute architectural features. Moreover, a deeper extraction of features from the input data will be crucial to ensure that the generated floorplans adhere more strictly to the given constraints. Addressing these challenges is essential for bridging the gap between automated floorplan generation and the nuanced requirements of architectural design.

Acknowledgements

This study is supported by the National Natural Science Foundation of China (Grant No. 5220080610) and China Scholarship Council grant (No. 202206250074). Part of the research was conducted at the Singapore-ETH Centre, which was established collaboratively between ETH Zürich and the National Research Foundation Singapore. This research is supported by the National Research Foundation, Prime Minister's Office, Singapore under its Campus for Research Excellence and Technological Enterprise (CREATE) program. We also would like to acknowledge that computational work involved in this research work is partially supported by National University of Singapore IT's Research Computing group.

References

Hu, R., Huang, Z., Tang, Y., Van Kaick, O., Zhang, H., & Huang, H. (2020). Graph2plan: Learning floorplan generation from layout graphs. *ACM Transactions on Graphics (TOG)*, 39(4), 118-1. https://doi.org/10.1145/3386569.3392391.

Nauata, N., Chang, K. H., Cheng, C. Y., Mori, G., & Furukawa, Y. (2020). House-gan: Relational generative adversarial networks for graph-constrained house layout generation. *In Computer Vision–ECCV 2020: 16th European Conference,* Glasgow, UK, August 23–28, 2020, Proceedings, Part I 16 (pp. 162-177). Springer International Publishing. https://doi.org/10.1007/978-3-030-58452-8_10.

Nauata, N., Hosseini, S., Chang, K. H., Chu, H., Cheng, C. Y., & Furukawa, Y. (2021). House-GAN++: Generative adversarial layout refinement network towards intelligent computational agent for professional architects. *In 2021 IEEE/CVF conference on computer vision and pattern recognition (CVPR)*. Nashville, TN, USA (IEEE). 2021: 13627-13636. In CVF Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 13627-13636). https://doi.org/10.1109/CVPR46437.2021.01342.

Shabani, M. A., Hosseini, S., & Furukawa, Y. (2023). Housediffusion: Vector floorplan generation via a diffusion model with discrete and continuous denoising. *In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 5466-5475). https://doi.org/10.1109/CVPR52729.2023.00529.

Sun, J., Wu, W., Liu, L., Min, W., Zhang, G., & Zheng, L. (2022). Wallplan: synthesizing floorplans by learning to generate wall graphs. *ACM Transactions on Graphics (TOG)*, 41(4), 1-14. https://doi.org/10.1145/3528223.3530135.

Upadhyay, A., Dubey, A., Mani Kuriakose, S., & Agarawal, S. (2023, January). FloorGAN: Generative Network for Automated Floor Layout Generation. *In Proceedings of the 6th Joint International Conference on Data Science & Management of Data (10th ACM IKDD CODS and 28th COMAD)* (pp. 140-148). https://doi.org/10.1145/3570991.3571057

Veloso, P., Rhee, J., Bidgoli, A., & Ladron, G. (2022). Bubble2Floor: A Pedagogical Experience With Deep Learning for Floor Plan Generation. In *27th International Conference on Computer-Aided Architectural Design Research in Asia: Intelligent and Informed, CAADRIA 2022* (pp. 373-382). The Association for Computer-Aided Architectural Design Research in Asia (CAADRIA). https://doi.org/10.52842/conf.caadria.2022.1.373.

Wu, W., Fu, X. M., Tang, R., Wang, Y., Qi, Y. H., & Liu, L. (2019). Data-driven interior plan generation for residential buildings. *ACM Transactions on Graphics (TOG)*, 38(6), 1-12. https://doi.org/10.1145/3355089.3356556.

LATENT PETROGLYPHS: PATTERN EXTRACTION FROM PREHISTORIC ROCK ART THROUGH GENERATIVE WORKFLOWS FOR A DESIGN PROJECT IN GREECE

SMARO KATSANGELOU[1], PARKER WILMOTH[2], DIMITRIS A. PADOS[3] and EMMANOUIL VERMISSO[4]
[1,4] School of Architecture, Florida Atlantic University
[2,3] Department of Electrical Engineering & Computer Science, Florida Atlantic University
1skatsangelou2023@fau.edu 0000-0003-1135-4667
2pwilmoth2023@fau.edu, 0000-0003-1731-4690
3dpados@fau.edu, 0000-0001-8959-6450
[4]evermiss@fau.edu, 0000-0001-9116-8877

Abstract. This paper regards the translation of indigenous rock art (petroglyphs) into training data for deep learning algorithms. Vis-à-vis the recent popularity of pre-trained AI models, the authors examine the potential of domain-specific search procedures to inform the process for a building design in Greece. Petroglyphs are a primitive form of artistic expression which has survived through the ages due to the medium upon which it was engraved. The practical aspect of this art was navigating through nature. The "Rock Art Center" aims to exhibit the narratives and culture behind rock art scattered in the mountains. Considering the adoption of generative adversarial networks (GANs) in the architectural workflow, the landscape and local prehistoric graffiti are viewed as datasets for tackling different design decisions, formally and conceptually interrogating the project's scope. The existing rock art sites serve as the primary dataset to explore the building's form, by accessing the 'latent' space of prehistoric rock art and its interpolation with the demands of the project. A number of algorithms and digital tools is employed to interpret the data in question.

Keywords. Artificial Intelligence, Deep Learning, Architectural Ideation, Design Workflow, Image Generation.

1. Introduction

The intersection of prehistoric art, architecture and deep learning (DL) is explored to navigate a design method for the creation of a building for a 'Rock Art Center' in Northern Greece. The authors draw inspiration by leveraging the gestural simplicity present in prehistoric carvings and the robustness of deep learning to extract features from large amounts of data. Architecture is often viewed as system, which comes as response to an existing set of conditions (Ching, 2007) In addition to this, "Architecture

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 149-158.

is a gesture" according to (Wittgenstein, 1980); in that sense, the satisfaction of functional requirements in a building is inadequate to render this "Architecture". True architectural achievement relies on communicating a notion, regardless -or in addition to- other (utilitarian) criteria. The need to communicate ideas and experiences in a non-verbal manner is inscribed in humanity's course, transcending time. Gesture in architecture and visual arts plays a crucial role in conveying ideas in a non-verbal manner. When "gesture" as a movement comes into play with the specific demands of professional architectural projects, a design negotiation occurs between the functional and aesthetic aspects of design.

One of the first, primitive efforts of man to create artistic gestures survives to this day in the form of petroglyphs, geoglyphs, rock, and cave paintings. The archaic is one of the greatest inventions of the 20th century (Davenport, 1981). The clarity and simplicity of prehistoric art has been the basis of many 20th-century artists' work such as Picasso and Modigliani. Romanticizing and misinterpreting prehistoric or indigenous heritage has been at times criticized as cultural appropriation of "primitive" art forms (Lock, 1994). Petroglyphs can only be experienced in-situ, "en plein air". The point of their existence is not to be carried and exhibited in the neutrality of a museum space and maybe not even to be viewed by the greater public (Anati, 1994).

Rock etchings are found across a span of 20km in the mountainside of Mount Paggaion, telling a story of non-verbal, prehistoric communication shared amongst groups of shepherds to mark and navigate through the natural landscape (Chatzilazaridis, 2000). This graffiti, dating from the Early Bronze Age up until the Middle Ages, depicts deer, horse riders, archers, hunters, standing figures, tools, and boats. Due to the inaccessibility of the locations where they were discovered (Moutsopoulos, 1969) they survive to this day. Some of the positions are impossible to access, so the creation of a building, which narrates the story of prehistoric art in Greece was necessary. The “Rock Art Center” Project is developed to inform the public about prehistoric civilization in the area, drawing visitors and attention to this yet unexplored and unknown part of Greek cultural heritage and history.

Prior related work exploring petroglyphs through generative design methods includes artificial petroglyphs generation through Hough Transform (Zhu, 2009). In this instance, the authors considered the use of deep neural networks (DNNs), in order to broaden the search in the ideation process drawing inspiration from these gestural depictions. Reinterpreting and exploring the design solution ("latent") space of the petroglyphs becomes a driving force for designing features of the building. These "latent" design opportunities explored in the proposed experimental workflow spark a debate about design agency and the idiosyncrasy of design as 'layered' process. Architectural design is a complex, multi-step process, which cannot be replicated entirely by AI, but can rather be infused with artificially generated outputs. Singular AI models cannot adequately capture a global configuration of complexities and multimodality of architectural production (Bolojan, Vermisso & Yousif, 2022). This has typically been addressed by multiple models, a sort of 'ecosystem' of algorithms designed to tackle different architectural tasks using the Deep Himmelblau network (Bolojan, 2022). Along these lines, the authors used custom-made Deep Learning algorithms to address the multimodality of architecture when considering the workflow for this design project.

2. Design Stages and Relevant Datasets

The case of using a piece of art or an everyday object as a starting point for lining out the shape of a building has unravelled manifold in the course of architectural history. From Claude-Nicolas Ledoux's Oikema to modern interpretations of mimetic architecture, the case for drawing inspiration from everyday life and embedding it into the formal configuration of buildings has been explored in various manners.

This project tries to intuitively link the gestural and artistic qualities of primitive carvings to those of a contemporary building attributes (façade; massing) through the designer's agency and personal design sensibilities. Our proposed workflow examined an AI-assisted method for addressing two architectural layers (1) Building Façade and (2) Building Massing, featuring two deep neural networks (GAN1; GAN2 respectively) which are informed by features from the petroglyphs data.

2.1. DEEP NEURAL NETWORK 1: PETROGLYPH CREATOR (GAN1)

The use of two-dimensional characters as design space driver has been previously explored in projects like Zaha Hadid's seminal design for Monsoon restaurant in Sapporo, Japan (Pavitt & Howe, 2023), where the Arabic writing of "Zaha" and "Monsoon", Hadid were used to kick-start the design process.

Inspired by the architecture of a StyleGAN, this network (DNN1) creates 'deep-fake' petroglyphs. The artificially generated petroglyphs are used as a reference to create a metal clad façade for the building, while also serving as serve as synthetic dataset for running a second network for massing. The objective is to explore the latent space of prehistoric rock art; by interpolating the different forms encountered within the dataset, new, artificial, yet untapped formations are navigated. Retrieved from archaeological documentation studies, sketches depicting the different prehistoric etchings found in the Mountain are used as a second datasets involved in the design of this building. This dataset initially included 122 images of 1024x1024 pixels depicting animals, bows, chariots, deer, hunters, people, ships, and miscellaneous objects.

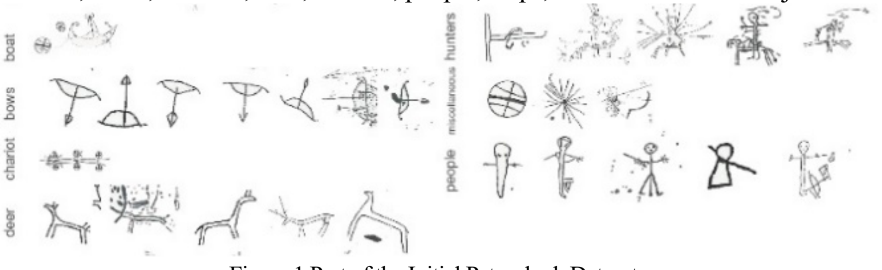


Figure 1 Part of the Initial Petroglyph Dataset.

As this number of instances in inadequate for training a DNN, the dataset was augmented tenfold. During augmentation, there were no restrictions on orientation, scaling, or other distortion of the images, as their semantic meaning was not altered in any way. These 8 classes of instances are inequilaterally represented in the initial dataset, as some appear in larger numbers than others. The augmentation was done in a manner that does not distort the quota of each class in the overall dataset. The results of this DNN trained on petroglyphs extracts features encountered in all eight classes.

Harun Farocki concocted the term "operative images" to describe images which are not meant for human viewing, but rather to be interpreted by machine vision (Farocki, 2010). In that sense, even though the images were initially cropped and formatted into 1024x1024 pixels, the team found that the semantic information could be conveyed in far lesser pixels, reducing the trainable parameters of the network and hence the training time. Consequently, the dataset comprises of 10k images of petroglyphs with a resolution of 64x64 pixels. The idea of operative image consideration therefore interrogates the threshold of size, aspect ratio needed to find optimize our computation needs for this experiment. The objective is to pinpoint the minimum resolution threshold, through which semantically meaningful results, which useful to the progression of the project can be produced.

2.2. DEEP NEURAL NETWORK 2: BUILDING PROGRAM (GAN2)

The use of artificial intelligence in architecture has been criticized for the lack of domain-specific datasets and agency in the design process (Bernstein, 2022). In this project, the creation of a dataset, which caters to a specific need for the design process, is used to inform design decisions regarding the spread of the square feet within the plot and the building program. Exploring the formal configuration of the project was of particular importance and interest, due to the lack of adjacent buildings in the plot. Found in a mountainous area, with no built but rather only natural environment surrounding it, this plot measures 5000 m^2, whereas one can construct up to 800 m^2.

A holistic approach to designing both the shell of the building, the exhibition and figuring out the building program was followed. The building program is structured around four fundamental pillars linked to the function of the spaces. Firstly, a robotic arm petroglyph carver is set as the main focus of the entrance of the building. Secondly displays of a DNN, which extracts and compares features encountered in petroglyphs from different regions of the world structures the space. Thirdly, a space for physical models either of the mountain, the rocks or the building itself and 3D printed real size petroglyph rocks, produced through photogrammetry models, are showcased. Lastly, a space where a shop, café and a gathering space is envisioned.

Using a Pix2Pix network architecture, corresponding massing iterations are created by the designer, based off 100 different petroglyphs which were created using DNN1, as explained in section 2.1. Even though the correlation between the provided petroglyphs and the formal configurations is at times visible, a direct correlation is avoided. In that sense, the objective is not to generate a building, which emulates the shape of a deer or a hunter, but rather to transcribe the fundamental rules, which connect the two systems. To encode the four different modules of the building program and corresponding square footage, four colours -Green, Blue, Peach and Grey- are used to distinguish the function of the different building volumes. The order of the system components come together to make up a whole based on the principles of hierarchy, rhythm, datum and transformation (Ching, 2007). From the resulting 3D iterations, a top view snapshot is captured to describe the building's shape. To work with the DNN, the produced dataset is subsequently augmented to 2000 images of building massings, initially formatted to 1024x1024 pixels. To make this experimental workflow semantically meaningful and useful to the design process, every module of the floorplan for each iteration has specific square footage, representing the exact space, which can be built on the plot. Split between 250m^2 for the petroglyph carver, 250m^2 for the showcase of petroglyphs from around the world, 200m^2 for the physical models and 100m^2 for the café and shop, for each iteration, the massings coalesce into a cohesive ensemble, whose geometry could potentially evolve into the final building.

Even though most GANs are trained on 8 GPUs, at a resolution of 1024x1024 pixels, this magnitude of computational power is not attainable for most designers and architects. To make this experimental workflow easily replicable by other architects, the team chose to run it on a single GPU, rather than multiple ones. For time-efficiency during training, images were converted to Grayscale, decreasing parameters by 66.6% and consequently, the necessary training time and computational power. The democratization and proliferation of an architectural workflow which includes custom, domain-specific datasets were driving forces for designing the specific network. Within architectural workflows, most AI tools commonly used are either pre-trained and readily available online, trained on a general dataset such as ImageNet (which is not catered to designers' needs) or developed with specific foci and large-scale computational resources, which are hard to replicate by small practices or teams of designers. To this end, bridging the gap between small-scale architectural production and existing AI generative models was considered paramount in this study.

3. Workflow - Methodology

Design is viewed as the accumulation of data and information from various sources and their attribution to inform the multimodality of architectural design (Ching, 2007). Before computational tools were invented, this process was carried out through extensive amounts of travelling by young, aspiring architects and artists, who roamed through Europe to experience and observe antiquities (Frampton, 1980). In this process, the petroglyphs found in the vicinity of the site are used as the conceptual basis for tackling two distinct tasks interwoven with the design of the building. Architectural

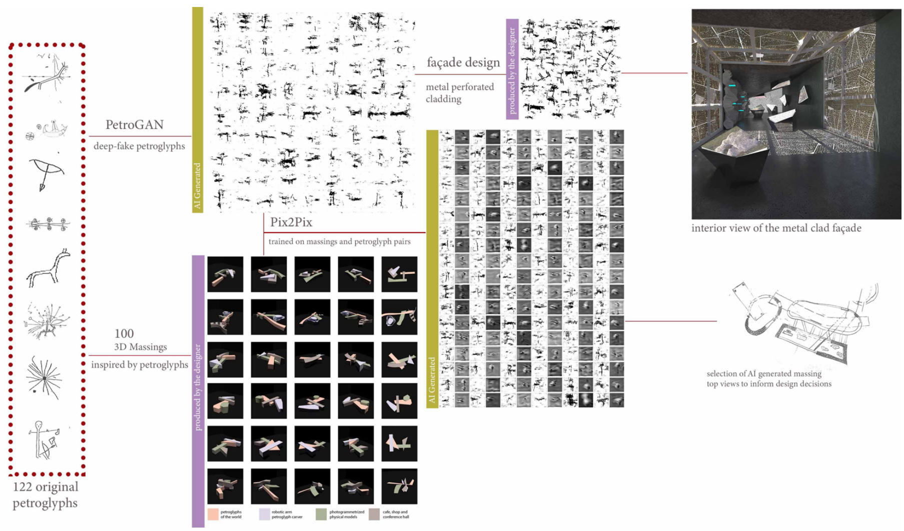


Figure 2 Workflow diagram

objects embody a thought process, which is repeated mentally by each person experiencing the space (Heinrich de Fries). Composing a building is often a negotiation between utilitarian and aesthetic considerations (Pehnt, 1973). The two realms can at times be conflicting, walking the line between what 'could' and what 'would' be. The term 'Expressionist Architecture' was coined in the brink of the 20th century to describe a style of architecture, which creates a certain mood or atmosphere (Behne,1915).

3.1. TECHNICAL DETAILS: GENERATIVE MODELS' ARCHITECTURE

This project used two DNNs (GAN1; GAN2) run on a virtual environment using Jupyter Notebook with Tensorflow, due to the obsolescence in many libraries used to train GANs. GANs gradually turn initial Gaussian noise into artificially generated images (Karras et al., 2021) through a minimax game, where a generator and a discriminator network compete against each other. The activation function used for both networks is LeakyReLU. The training progression of the first network showcases that more training time outputs decreasingly semantically meaningful results, whereas the second network moves from abstract interpretations of fake petroglyphs towards more defined geometries until it overfits, making a direct translation of the petroglyphs.

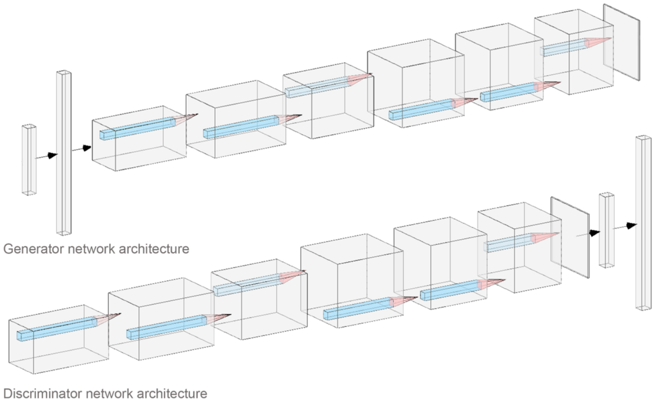


Figure 3 Employed GAN Network Architecture for GAN1

GAN1, a StyleGAN-type network, was trained on original petroglyphs, and comprises of a discriminator which inlcudes a LeakyReLU activation function, a dropout rate to regularize the network as well as a final dense layer with sigmoid activation for one-hot encoding. The discriminator has a 64x64x1 pixel image as input and outputs a value from 0.00 to 1.00. The Generator includes multiple convolutional layers, batch normalization and LeakyReLU activation. The output layer uses linear activation function to cater to a wider range of value. The generator initially takes random Gaussian noise vector of size 100 and outputs an image of 64x64 pixels.

GAN2, a Pix2Pix type network is trained to map images from one domain (petroglyphs) to another (massings). Like GAN1, its architecture consists of a "generator" and a "discriminator". The generator tackles transforming the petroglyphs from 2D representations to 3D top-down massing iterations. A series of convolutional layers extract features from the input (fake petroglyphs). The transposition of convolutional layers (UpSampling2D) up-samples the features of the images to gradually reconstruct the target image through Gaussian Noise. Finally, the last layer uses linear activation to produce the predicted massing. The discriminator distinguishes between fake and real using convolutional layers to process the input and extract features.

Other experiments conducted by the team include the use of different angles of the representations of the 3D iterations to train a third GAN. The network did not seem to produce semantically meaningful results to inform the ideation process of the architectural design. Consequently, this network was not used further in our proposed workflow. Potentially, the use of more high-resolution images of the iterations could yield more useful results. Nonetheless, the team chose to keep the resolution low to reduce the training time and required computational power.

4. Results

GAN1 trained on 10k images of Petroglyphs, solely from the area of Paggaion. The network trained for more than 1000 epochs. The results of the network can be identified as the initial different classes. Hunters and animals occur, as well as standing figures. Moreover, hybridized forms, which interpolate between the forms of people and animals. When we attempt a one-on-one comparison of the provided original petroglyphs and the results emulate the structure and compositional values of the training set. The line weight, as well as the clarity of the lines is not of the same consistency as the training dataset. The outputs of GAN1 were used to create a new composition, which turned to be a laser-cut metal clad façade for the 'Rock-Art Center'. Anthropomorph and animal-like figures can be distinguished from the outputs. While training, the team decided to omit some of the samples belonging to the 'Miscellaneous' class, as their unstructured composition seemed to affect the network negatively.

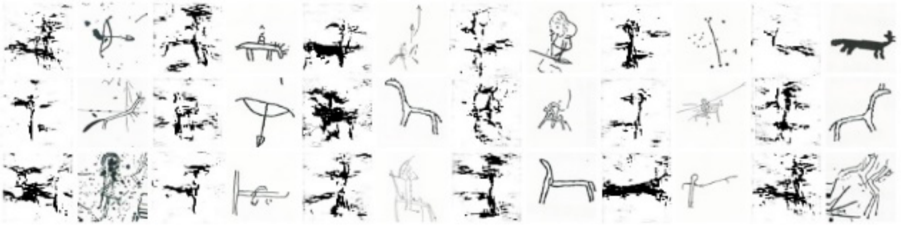


Figure 2 Fake and Real Petroglyph Comparison

The second network (GAN2) for this experimental workflow was based on the Pix2Pix architecture, combining the top view of the massings created by the designer as well as the original augmented petroglyphs. The network was trained on original petroglyphs as well as the original massings created by the team. For the final use of the artificially produced massings, the input of the network was 'deep-fake' petroglyphs. In that sense, the concept of exploring the latent space of prehistoric forms was expanded in the overall configuration of the building.

From a total of 1000 massings based on fake petroglyphs, the team chose 10 which were considered for the process of finalizing the formal aspect of the building. The network, as it is pretrained on the provided original petroglyphs as well as the massings, can potentially output new massings provided any two dimensional black and white drawing. The designer's agency and personal style is passed onto the network through the manual creation of the initial massing dataset. This network caters to a very specific design problem, bridging the gap between function and aesthetics. The ambiguity of the results is beneficial to the development of the architectural project, as it leaves room for further interpretation and curation based on the designer's intuition and aesthetics.

It is observed that most of the produced massing diagrams include two, three or four predominant volumes, which are correlated with the initial dataset upon which the network was trained. The initial conception for the dataset was to be structured around four color-coded modules which make up a whole. So, it is accessed that some of the principles followed to create the dataset were learned successfully by the network, even with a resolution as low as 64x64. There are various instances in the training progression of the second network where the results create stimulating results for the aforementioned purpose. It is nonetheless obvious when the network starts to overfit, as the outputs start to mimic the petroglyphs as a one-to-one translation rather than a loose interpretation. Even when subtracting color and resolution from the dataset, the outputs still show variation with 3 or 4 differentiated volumes, which follows the rationale of the dataset of massings produced by the designer. This shows that the network has learnt some of the compositional characteristics of the input, whilst definitely lacking definition. What's more, the output images take up similar square footage as the one's provided as training material.

5. Conclusions and Future Work.

In the 20th century, the use of material from the deeper past of human civilization is equivalent to the use of design and aesthetic ideals from Hellenistic Rome during the European Renaissance (Davenport, 1981). Prehistoric art has been explored and interpreted in the context of modernity, lacking at times context and deeper understanding of the circumstances upon which these creations came to be. The ideation process as well as the final output of the building is informed through generative algorithms. Interpreting and reiterating the petroglyphs of Paggaion gave ground to the conception of new ideas, as well as formal configurations for the building. Moreover, the accessibility of this method, rather than using multiple GPUs, has potential of broader use either commercially or pedagogically.

The intersection of Architecture and Artificial Intelligence has been explored in multiple ways spanning from GANs to diffusion models. This year, the debate has been monopolized by the rise of powerful text-to-image models which heuristically kick-start the design process, serving as a tool to boost brainstorming. In our study, a domain-specific dataset, created manually is used to aid in exploring the latent space of Petroglyph forms and building configurations. The designer's agency or copyrights to the produced building is not compromised as the network serves as a tool for specific sub-modalities of what makes up architecture.

Based off selected produced massings, the designer can use the synthetic output as a basis for tracing the floorplan of the building. Going back and forth between analogue, digital and artificially generated content is proposed as an alternative workflow, which could be easily, and (time-wise) efficiently replicated by small scale design teams of 5 to 10 people. The team is working towards higher resolution outcomes, which can capture the expressiveness of BIM modelled building massings.

For these experiments, the dataset's resolution was reduced to make this workflow easily accessible and replicable for potential educational purposes. The training time for each model is estimated at 2 minutes with a GForce RTX 3070 GPU. The reduced size of the dataset results in reduced training time and computational power. The network was trained locally, rather than on GoogleCollab, making it more resilient to updates and easily accessible to architects and students. This study acts as an early version of an architectural workflow ran on a very small initial dataset. This experiment is to be used as a proof of concept for the team's workflow.

In future experiments, the authors will focus on using three-dimensional geometry to produce artificially generated massings in 3D. The use of actual three-dimensional geometry, rather than two-dimensional representations of it, is envisioned as a basis for the design process in a more sculptural manner. Experiments of this sort have been conducted by NVIDIA Lab. The generation of high-quality 3D textured shapes has been achieved through training a network with 442 shapes of animals, 563 houses of shapes and 500 shapes of people (Gao, 2022). The non domain-specific nature of the datasets and their random collection from Turbosquid may yield high-quality results but cannot be utilized for particular design needs, so the authors intend to curate further bespoke datasets for training the fine-tuned current models (GAN1; GAN2).

Acknowledgements

The project was commissioned and funded by the Kavala Municipality in Northern Greece. The team also acknowledges the contribution of the resources and know-how of the "Center for Connected Autonomy & Artificial Intelligence" at FAU.

References

Anati, E. (1994). Archetypes, constants, and universal paradigms in prehistoric art. *Semiotica 100* (2-4):125-140. https://doi.org/10.1515/semi.1994.100.2-4.125

Bolojan, D., Vermisso, E., & Yousif, S. (2022). Is Language All We Need? A Query Into Architectural Semantics Using a Multimodal Generative Workflow. *CAADRIA 2022: Post-Carbon*. Sydney: CAADRIA.

Behne, A, (1915). *Biologie und Kubismus. Der Sturm* 5(11/12), 68–71.

Bernstein, P. (2022). *Machine Learning: Architecture in the age of Artificial Intelligence*. RIBA publishing, London.

Bolojan, D.: Creative AI: Augmenting design potency. *Archit. Des*. 92(3), 22–27 (2022). https:/ /doi.org/10.1002/ad.2809

Chatzilazaridis, L. (2000). *Prehistoric rock engravings of Northern Greece*. Doctoral Dissertation, Aristotle University of Thessaloniki. Thessaloniki, Greece. https://thesis.ekt.gr/thesisBookReader/id/17938#page/1/mode/2up

Ching, F. (2007). *Architecture: form, space and order.* John Wiley & Sons, Inc. New Jersey. https://rsd2-alert-durden-reading-room.weebly.com/uploads/6/7/1/6/6716949/ching-architecture_form_space_order.pdf

Davenport, G. (1974). The symbol of the archaic. *The Georgia Review,* 28(4), 642–657. http://www.jstor.org/stable/41397160

Davenport, G. (1981). *The Geography of the Imagination: Forty Essays.* North Point Press, San Franscisco. https://issuu.com/polinecia/docs/the_geography_of_the_imagination.

Farocki, H. (2010). *Installations. 2010. Re-pouring. 2010*. Retrieved November 17, 2023, from http://www. harunfarocki.de/installations/2010s/2010/re-pouring.html.
Frampton, K. (1980). *Modern Architecture: A Critical History.* Thames & Hudson, London. DOI:10.1080/10331867.2021.1989898
Gao J., Shen T., Wang Z., Chen W., Yin K., Li D., Litany O., Gojcic Z., and Fidler S. (2022). GET3D: A Generative Model of High Quality 3D Textured Shapes Learned from Images. In *Advances in Neural Information Processing Systems.* Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2022/file/cebbd24f1e50bcb63d015611fe0fe767-Paper-Conference.pdf
GET3D: A Generative Model of High Quality 3D Textured Shapes Learned from Images https://research.nvidia.com/labs/toronto-ai/GET3D/assets/paper.pdf
Harding, J. (2016). Dimensionality reduction for parametric design exploration. In: *Advances in Architectural Geometry 2016*, September 9th - 13th, 2016, Zurich, Switzerland. DOI:10.3218/3778-4_19
Karras T., Aittala M., Laine S., Harkonen E, Hellsten J., Lehtinen J., and Aila T. (2021). Alias-free generative adversarial networks. *Advances in NeurIPS*, 34, 2021.
Moutsopoulos, N. (1969). *Ta akidografimata tou Paggaiou (The graffiti of Paggaion).* Aristotle University of Thessaloniki Publications. Thessaloniki, Greece. http://apothesis.teicm.gr/xmlui/handle/123456789/3031
Pavitt, J. and Howe, C. (2023) Zaha Hadid Foundation: Interdisciplinarity and Intersectionality. In *annual conference of the Jaap Bakema Study Centre Architecture Archives of the Future,* TU Delft – Nieuwe Instituut in Rotterdam, November 2023.
Pehnt, W. (1973). *Expressionist Architecture.* Thames & Hudson, London. https://pdfcoffee.com/expressionist-architecture-2-pdf-free.html
Zhu Q., Wang X., Keogh E., and Lee S. (2009). Augmenting the generalized hough transform to enable the mining of petroglyphs. In *Proceedings of the 15th ACM SIGKDD International conference on Knowledge discovery and data mining (KDD '09).* Association for Computing Machinery, New York, 1057–1066. https://doi.org/10.1145/1557019.1557133

LEARNING AND GENERATING SPATIAL CONCEPTS OF MODERNIST ARCHITECTURE VIA GRAPH MACHINE LEARNING

ERIK BAUSCHER[1], ANNI DAI[2], DIELLZA ELSHANI[3], THOMAS WORTMANN[4]
[1,2,3,4]*Chair for Computing in Architecture, Institute for Computational Design and Construction, University of Stuttgart*
[1]*erikbauscher5@gmail.com, 0009-0007-6980-0873*
[2,3,4]*{anni.dai|diellza.elshani|thomas.wortmann}@icd.uni-stuttgart.com, 0009-0002-8521-5045|0000-0003-2902-341X|0000-0002-5604-1624*

Abstract. This project showcases a use case away from most other research in the field of generative AI in architecture. We present a workflow to generate new, three-dimensional spatial configurations of buildings by sampling the latent space of a graph auto-encoder. Graph representations of three-dimensional buildings can store more data and hence reduce the loss of information from building to machine learning model compared to image- and voxel-based representations. Graphs do not only represent information about elements (nodes/pixels/etc.) but also the relationships between elements (edges). This is specifically helpful in architecture where we define space as an assemblage of physical elements which are all somehow connected (i.e., wall touches floor). Our method generates valuable, logical and original geometries that represent the architectural style chosen in the training data. These geometries are highly different from any image-based generation process and justify the importance of graph-based 3D geometry generation of architecture via machine learning. The method also introduces a novel conversion process from architecture to graph, an adapted decoder architecture, and a physical prototype to control the generation process, all making generative machine learning more applicable to a real-world scenario of designing a building.

Keywords. generative 3D architecture, generative graph machine learning, graph-based architecture, human-computer interaction, graph autoencoder, latentwalk

1. Introduction

The main player, image-based generative AI systems lack seriousness and logic when applied in a real-world design scenario, mainly due to their focus on the visual. Image Generation tools like Dall-E, Midjourney and Stable Diffusion produce images from text prompts. One or multiple generated images are used to extract information about architectural concepts, geometry and materials to inform building designs. Here the image as the medium cannot offer logical or coherent information due to its pixel-based nature, and human architects need to perform a manual conversion into a working building design. This makes AI image generation great as an inspiration in the very early stage of design, but hard to apply to the following design process.

In addition, photographs of buildings, floor plans, sections, etc. contain only compressed and partial information about building designs. In the world of computers this step of compressing is redundant (Carpo, 2017) and rather unhelpful for neural networks (NNs) to understand complex content.

This work wants to suggest a different approach, showing the benefits of graphs and three-dimensional representations of buildings for generating new architecture. Graphs can contain more information from different domains, because they are not grid-based, and do not only represent information about elements (nodes, pixels, etc.) but also the relationships between nodes (Hamilton, 2020).

Here, rather than focusing on visual information, the definition and generation of space are in the foreground, intending to give the architects a tool they can use hand in hand with their current design workflow. Compared with existing methods for architectural AI generation, the presented method generates three-dimensional spatial arrangements that are more complex than those allowed by voxel grids.

1.1. SUBSYMBOLIC AI

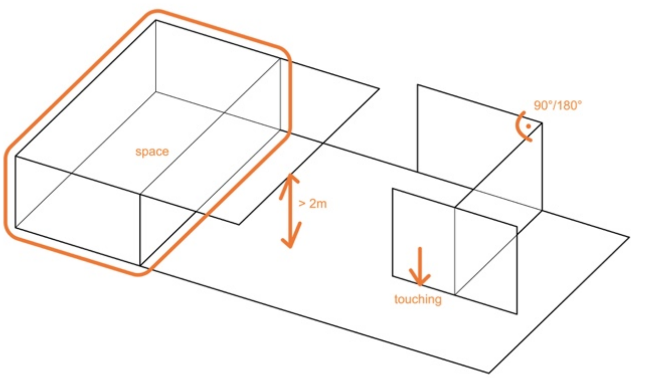


Figure 1: Subsymbolic AI System finds high-level features in 3D buildings

Symbolic AI focuses on teaching an AI system the rules of the game of chess. Subsymbolic AI only lets the NN play chess making the model learn the rules by its own (Mitchell, 2020).

By using unlabelled data for training a generative AI system in architecture, the second described approach is applied. We try to make a deep neural network understand very complex, three-dimensional buildings, without giving it i.e., information about gravity (Fig. 1). Without this information, the NN must figure out that walls are always touching horizontal elements by itself, and then use this knowledge when generating. As buildings are extensively more complex than images, this task of understanding the so-called high-level features becomes more difficult for the NN. In the following, we try to lower the complexity as much as possible while still maintaining the original idea of teaching an NN to generate three-dimensional buildings from scratch.

1.2. STATE OF THE ART

While there already is architectural research about graphs and graph neural networks in architecture (Alymani et al., 2022), it mostly focuses on classification tasks rather than generative ones. Alymani et al. use the dual graph representation for classification tasks with graph machine learning and already show promising results. The trend can also be seen in other research on generative architecture lately. Zhong et al. (2023) use graph representation learning in combination with a recursive neural network to generate 3D massing models of architecture, while other research uses layout graphs to generate 2D floorplans (Hu et al., 2020). In both cases, the results are far from perfect and not completely generalizable. Yet being based on graph representations, the models seem to be more capable of generating logical architectures than models built upon pixel- or voxel data (del Campo, 2022; Koh, 2020), where although the generated architectures are 3D and partially look stunning, they seem to still need manual or parametric postprocessing to represent a real-world building design.

1.3. GRAPHS

Buildings and graphs are deeply related because they both have the property of being non-discursive, meaning that they cannot be fully described by words or rules, but rather by their patterns and relations (Hillier, 1996). Early research on converting architectural ideas to graphs was already done in the 1970s inter alia by Christopher Alexander (1977) in his book "A Pattern Language", where he provides rules, patterns and grammar to any possible situation in buildings. There are different types of graphs, including undirected (topology graphs) and directed graphs. In the last two decades, directed graphs using Semantic Web Standards have been visible in the building industry often in combination with Building Information Modelling standards, focusing mainly on ontologies (McGlinn & Pauwels, 2022). Semantic Web Standards can include languages such as the Web Ontology Language, which is used for enriching the data with semantics in the form of ontologies. Ontologies represent knowledge about things and the relations between them (Elshani et al., 2023). Work on directed graphs using Semantic Web Standards in the building industry includes the translation of the IFC schema to an OWL language, ifcOWL (Beetz et al., 2009), BOT ontology (Rasmussen et al., 2020), the Building and Habitats object Model Ontology and workflow (Elshani et al., 2022), etc.

Due to the construction of a custom dataset, in this project, the graph structure was simplified as much as possible resulting in an undirected graph holding mostly geometrical information in the node features and no edge features. In the future, the same method could be applied to much more detailed graphs as introduced here, resulting in more detailed, generated building designs.

2. Methodology

2.1. DATASET

Due to the lack of accessible 3D building datasets, a custom dataset needed to be constructed. One possibility would be to use one of many 2D floor plan datasets available (de las Heras et al., 2015; Kalervo et al., 2019; Wu et al., 2019) and extend the plans parametrically in the third dimension. The main advantage here is the great number of plans available to use for training the AI model. Conceptually however, it is hard to argue for the use of mostly unspecified floor plans which are only loosely connected to any architectural idea and style. In addition, a 2D to 3D conversion might defy any argument for a 3D-based AI model, when even the training data originally was 2D. The decision was taken to construct a simple dataset from scratch, where the biggest hurdle was to achieve a certain number of training data. Four well-known buildings from the modernist era were chosen to be remodelled and augmented each a hundred times. The augmentation was done parametrically by varying the position, rotation and dimensions of the individual building elements (wall, floor, ceiling) to a degree where the original spatial concept of the building would not be lost. The final dataset consists of 400 individual 3D buildings to train and test the model on. The four original houses are:

- Mies van der Rohe's Barcelona Pavilion (1929)
- Ray and Charles Eames' Eames House (1949)
- Mies van der Rohe's Farnsworth House (1951)
- Pierre Koenig's Stahl House (1960)

Modernist buildings were chosen for their simplicity in the definition of space and their good documentation. In modernism, structural elements were separated from space-defining ones, giving the architects more freedom in planning. This paired with the consistent use of right angles and geometrically simple volumes make the modernist style in architecture at least geometrically easy to understand, simplified and remodelled.

2.2. CONVERSION FROM BUILDING TO GRAPH

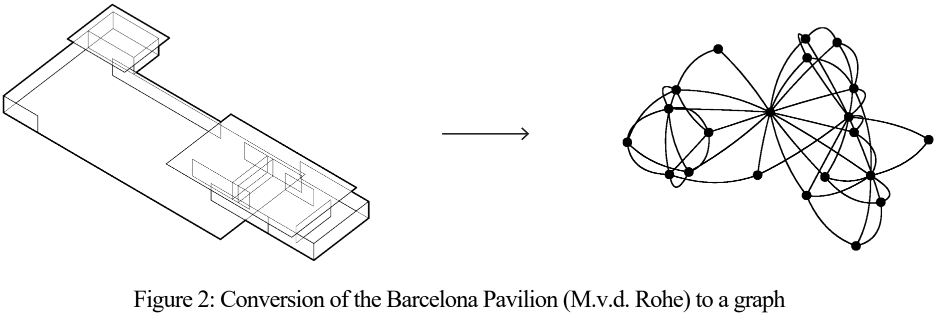


Figure 2: Conversion of the Barcelona Pavilion (M.v.d. Rohe) to a graph

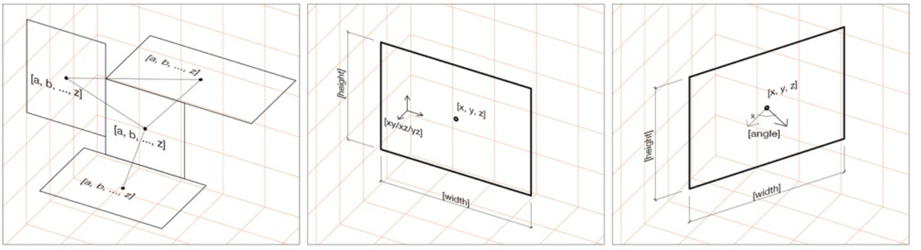


Figure 3: Left: Definition of a simple graph; Centre: Representation (A); Right: Representation (B)

To make a 3D building machine-readable, it is converted into a graph. Preparing the data as a graph highly influences the output of the generative model, as the conversion is the first and the last step performed in training and generation. This process is done parametrically and sets the geometric limitations for the output.

Each node in the graph contains geometric and material information about its respective physical element in the 3D model. Here, the advantage is made from the modernist buildings, which mostly allow being remodelled as 2D surfaces arranged in space. To keep the complexity low, the depth of the elements is not considered. The material of each element is described as either solid or transparent. Doors are not modelled but the space is left empty. Each element is described as a flat rectangle as visible in Figure 3. If the elements touch along any edge, they are linked in the graph by an edge (Fig. 2). The edges do not contain any features but are only used for message passing in the encoder. For the features of the nodes (elements), two versions were used to produce different results (Fig. 3), where (A) will be the default representation for the explanation of the model in the rest of the paper. Representation (A) represents one rectangular element by the X-, Y-, and Z- coordinates of its centre, the width and length of the rectangle and its orientation in space as a choice between the three world planes in the coordinate system. Representation (B) also includes the centre, width and height but uses the rotation of the surface's normal vector around World Z as information about orientation in space. (A) only allows three orientations of each element thus producing very static results. (B) gives the model more freedom but also enables more outcomes of the original style of the dataset. A representation with full freedom was also tried only using information from the four corner points, which resulted in too much sloping and uninterpretable results.

2.3. AUTOENCODER MODEL

A graph autoencoder as described first by Kipf and Welling (2016) was used to learn the features of the input architecture. The training data is unlabelled, thus while training the model is optimized only for matching the newly generated graph with the input graph. While the encoder consists of standard message passing layers (Hamilton et al., 2018), the decoder does not make use of the graph structure but decodes the latent vector through linear layers. A graph-based message-passing decoder was not chosen because of the difficulty of reconstructing a full graph from just the flattened one-dimensional latent vector (Guo & Zhao, 2022).

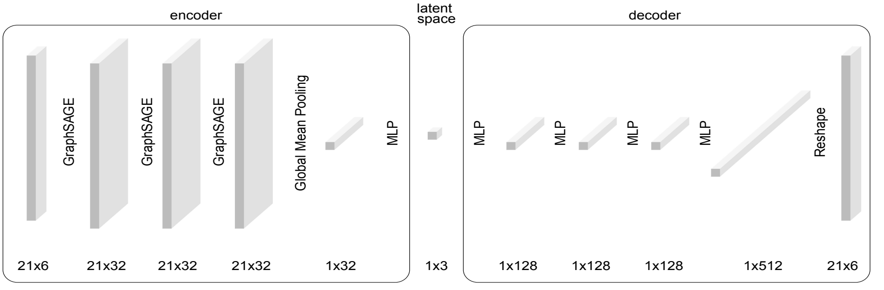


Figure 4: Definition of the used Graph Autoencoder Model

Using flattened data in the decoder brings the following limitation for the model: All graphs (hence all different buildings in the training data) need to have the same number of nodes. If just message passing was used, only the number of node features would need to be the same, but here due to the decoder architecture, it is the case for both the number of node features and number of nodes. Of course, this highly limits the range in which the model can generate new architecture as the buildings always will have the same number of elements. This loss was accepted due to the gained simplicity in training the model and using it for generation. The model was implemented using the Python machine learning framework PyTorch (Paszke et al., 2019) and its library PyTorch Geometric (Fey & Lenssen, 2019), which is specially written for graph machine learning. The final model used the following variables:

- Epochs: 1000
- Learning Rate: 0.001
- Optimiser: Adam
- Loss function: Mean Squared Error
- Activation function: ReLu
- Train, Validate, Test: 80%, 10%, 10%

2.4. GENERATION

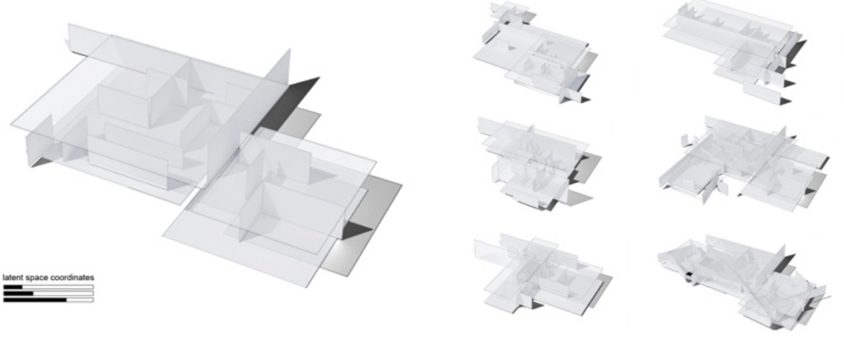


Figure 5: Newly generated, three-dimensional buildings

We generate new buildings by sampling the latent space of the trained autoencoder. When sampling unknown data points of the latent space, we expect the model to apply its learned logic and generate a new design that combines, reinterprets and varies features of buildings from known data points in the same area.

3. Results

In Figure 5 you can see the resulting generated geometry which was sampled in one session with the goal of finding new, never seen and buildable 3D arrangements. In the following, three geometries will be further described to make the model as well as its results more understandable.

3.1. EXAMPLES

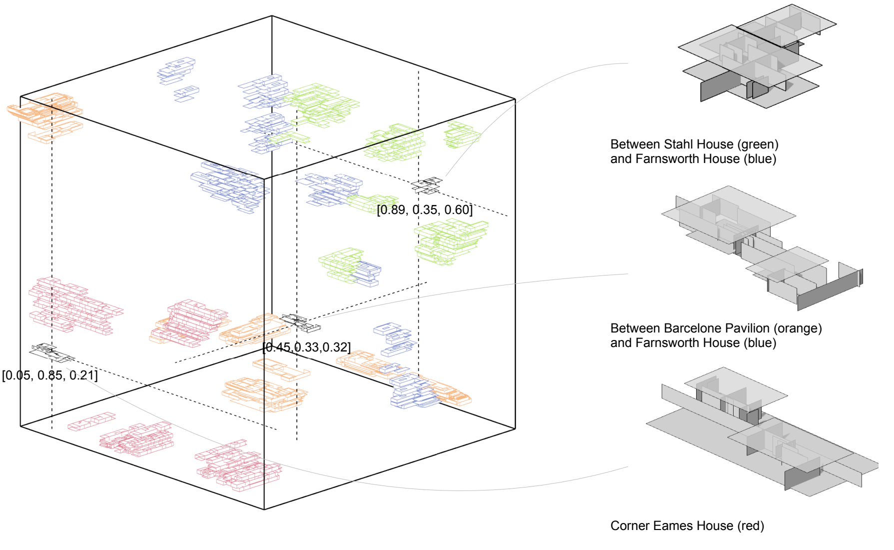


Figure 6: Left: Latent Space of the trained model; Right: Sampled new buildings

On the left-hand side, Figure 6 shows the 3D latent space of the trained model. All 400 buildings are mapped to their respective location in the latent space, showing that the model organised them in smaller groups mainly influenced by their affiliation to one of the original four modernist houses (colour) and their general orientation.

3.1.1. Between Stahl and Farnsworth House (Fig. 6, top)

While the slab alignments and sizes are still in the style of the Farnsworth House, the vertical elements are already arranged in the "L"- pattern as known from the Stahl House. The confused arrangement of the vertical elements results from the principle of "geometric travel" between known data points in the latent space. As usual in latent walks, when travelling between known geometry, the Autoencoder needs to negotiate the known geometries around the unknown point and consider any logic and rules learned when training. This sometimes results in the displayed semi-logical layouts of the building.

3.1.2. Between Barcelona Pavilion and Farnsworth House (Fig. 6, centre)

Here the model negotiates the concept of directionality from the Barcelona Pavilion with the unidirectional floor plan of the Farnsworth House. In general, the Autoencoder finds specific arrangements of groups of elements in the training data and tries to reproduce this arrangement in the bigger context of the full, newly generated building.

3.1.3. Corner Eames House (Fig. 6, bottom)

The sampled data point here is in proximity to where the model mapped most of the augmented Eames Houses in the latent space. The model discovered the two separated living areas along the long wall. In comparison to the original, one area is mirrored along the wall, creating a new layout, while maintaining most of the other space elements one can find in the original.

3.2. APPLICATION

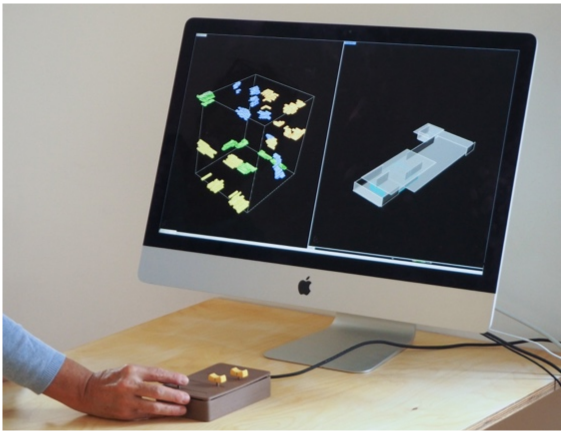


Figure 7: Photograph of the physical setup for using the model

To make the generative model approachable and usable for designers and architects, a physical prototype was built and connected with the Rhino/Grasshopper interface. The prototype consists of three sliders representing the three latent variables of the AI model. As shown in Figure 7, an architect can use those sliders to adjust the generated output in real-time. On the screen, one sees information about the location in the latent space as well as the generated 3D model in real-time.

This very intuitive way of generating new architecture makes the concept of the latent space very understandable even to non-experienced architects. It also emphasises the importance of the tool itself, which is needed to make research in architecture accessible to all designers and architects.

4. Discussion

Compared with pixel- and voxel-based generation, the presented method has the following advantages: (1) It works with three-dimensional geometries that do not need to follow any grid structure. (2) The graph representations contain more information that the model can learn from. (3) It presents a very intuitive, approachable and understandable generation tool for architects.

However, the current version suffers from the following disadvantages: (1) The qualities and availabilities of 3D datasets to use for training. (2) The low amount of already existing knowledge and research done in the field of graph-based, 3D generative AI systems in architecture. (3) The absence of an objective evaluation method for the generated results.

The presented method does not yet make use of any evaluation method other than the designer's eye. Since technically in training the only evaluation of the decoded latent space is its geometrical deviation from the originally inputted architecture, nothing else is needed for a working generation process. By using unlabelled data, we expect the model to learn the essential features of the training set by itself. Here the training set is defined as good architecture, and based on the idea of subsymbolic AI, the model can then generate new, good architecture, even from unknown data points in the latent space. Of course, this statement is insufficient, and one could achieve more control as well as better results by labelling the training data. These labels could include data from programmatic, environmental or structural simulations which can be produced with the already existing three-dimensional data. In a more industrial setting, it could also be values about costs, statistics on efficiency, used materials, etc.

To make the results more usable and realistic, further work can include labels for training and generating as well as extending the dataset with non-planar geometries. The model itself can also be redesigned to be a neurosymbolic AI system, combining a graph-based NN with a rule-based NN to enable an even deeper understanding of buildings (Sheth et al., 2023).

Acknowledgements

Supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany´s Excellence Strategy – EXC 2120/1 – 390831618.

References:

Alexander, C. (1977). A Pattern Language. Oxford University Press.

Alymani, A., Jabi, W., & Corcoran, P. (2022). Graph machine learning classification using architectural 3D topological models. SIMULATION: Transactions of The Society for Modeling and Simulation International, Online, 1–15. https://doi.org/10.1177/00375497221105894

Beetz, J., Leeuwen, J. van, & Vries, B. de. (2009). IfcOWL: A case of transforming EXPRESS schemas into ontologies. AI EDAM, 23(1), 89–101. https://doi.org/10.1017/S0890060409000122

Carpo, M. (2017, October 20). The Second Digital Turn. MIT Press. https://mitpress.mit.edu/9780262534024/the-second-digital-turn/

de las Heras, L.-P., Terrades, O. R., Robles, S., & Sánchez, G. (2015). CVC-FP and SGT: A new database for structural floor plan analysis and its groundtruthing tool. International Journal on Document Analysis and Recognition (IJDAR), 18(1), 15–30. https://doi.org/10.1007/s10032-014-0236-5

del Campo, M. (2022). Deep House—Datasets, estrangement, and the problem of the new. Architectural Intelligence, 1(1), 12. https://doi.org/10.1007/s44223-022-00013-w

Elshani, D., Hernandez, D., Lombardi, A., Siriwardena, L., Schwinn, T., Fisher, A., Staab, S., Menges, A., & Wortmann, T. (2023). Building Information Validation and Reasoning Using Semantic Web Technologies. In M. Turrin, C. Andriotis, & A. Rafiee (Eds.), Computer-Aided Architectural Design. INTERCONNECTIONS: Co-computing Beyond Boundaries (pp. 470–484). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-37189-9_31

Elshani, D., Lombardi, A., Fisher, A., Hernandez, D., Staab, S., & Wortmann, T. (2022, May). Knowledge Graphs for Multidisciplinary Co-Design: Introducing RDF to BHoM. ESWC - LDAC 2022.

Fey, M., & Lenssen, J. E. (2019). Fast Graph Representation Learning with PyTorch Geometric (arXiv:1903.02428). arXiv. https://doi.org/10.48550/arXiv.1903.02428

Guo, X., & Zhao, L. (2022). A Systematic Survey on Deep Generative Models for Graph Generation (arXiv:2007.06686). arXiv. https://doi.org/10.48550/arXiv.2007.06686

Hamilton, W. L. (2020). Graph Representation Learning.

Hamilton, W. L., Ying, R., & Leskovec, J. (2018). Inductive Representation Learning on Large Graphs (arXiv:1706.02216). arXiv. https://doi.org/10.48550/arXiv.1706.02216

Hillier, B. (1996). Space Is The Machine: A Configurational Theory Of Architecture.

Hu, R., Huang, Z., Tang, Y., van Kaick, O., Zhang, H., & Huang, H. (2020). Graph2Plan: Learning Floorplan Generation from Layout Graphs. ACM Transactions on Graphics, 39(4). https://doi.org/10.1145/3386569.3392391

Kalervo, A., Ylioinas, J., Häikiö, M., Karhu, A., & Kannala, J. (2019). CubiCasa5K: A Dataset and an Improved Multi-Task Model for Floorplan Image Analysis (arXiv:1904.01920). arXiv. https://doi.org/10.48550/arXiv.1904.01920

Kipf, T. N., & Welling, M. (2016). Variational Graph Auto-Encoders (arXiv:1611.07308). arXiv. https://doi.org/10.48550/arXiv.1611.07308

Koh, I. (2020). Voxel Synthesis for Architectural Design.

McGlinn, K., & Pauwels, P. (Eds.). (2022). Buildings and Semantics: Data Models and Web Technologies for the Built Environment. CRC Press. https://doi.org/10.1201/9781003204381

Mitchell, M. (2020). Artificial intelligence—A guide for thinking humans. Penguin Books.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Köpf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., … Chintala, S. (2019). PyTorch: An Imperative Style, High-Performance Deep Learning Library (arXiv:1912.01703). arXiv. https://doi.org/10.48550/arXiv.1912.01703

Rasmussen, M. H., Lefrançois, M., Schneider, G., & Pauwels, P. (2020). BOT: The Building Topology Ontology of the W3C Linked Building Data Group. Semantic Web. https://doi.org/10.3233/SW-200385

Sheth, A., Roy, K., & Gaur, M. (2023). Neurosymbolic AI -- Why, What, and How (arXiv:2305.00813). arXiv. https://doi.org/10.48550/arXiv.2305.00813

Wu, W., Xiao-Ming, F., Tang, R., Wang, Y., Qi, Y.-H., & Liu, L. (2019). Data-driven interior plan generation for residential buildings. ACM Transactions on Graphics, 38, 234:1-234:12. https://doi.org/10.1145/3355089.3356556

Zhong, X., Koh, I., & Fricker, P. D. P. (2023). Building-GNN: Exploring a co-design framework for generating controllable 3D building prototypes by graph and recurrent neural networks. Digital Design Reconsidered: Proceedings of the 41st Conference on Education and Research in Computer Aided Architectural Design in Europe (eCAADe 2023), 431–440. https://doi.org/10.52842/conf.ecaade.2023.2.431

DRAG2BUILD: INTERACTIVE POINT-BASED MANIPULATION OF 3D ARCHITECTURAL POINT CLOUDS GENERATED FROM A SINGLE IMAGE

JUN YIN[1], PENGJIAN XU[2], WEN GAO[3], PENGYU ZENG[4] and SHUAI LU[5]
[1,4,5]Tsinghua University, [2]Zoomtech Engineering Co., Ltd, [3]Beijing University of Technology
[1]yinj22@mails.tsinghua.edu.cn, 0009-0005-7530-533X

Abstract. At present, 3D reconstruction from images has made notable advancements in simple, small-scale scenes, but faces significant challenges in intricate, expansive architectural scenes. Focusing on the early stage of design stage, we present Drag2Build, a tool for converting images into point clouds for 3D reconstruction and modification in detailed architectural contexts. Our first step involved the creation of ArchiNet, a specialized 3D reconstruction dataset dedicated to elaborate architectural scenes. Next, we developed a 3D reconstruction approach using a conditional denoising diffusion model, enhanced by incorporating a model for segmenting objects, thereby improving segmentation and identification in complex scenes. Additionally, our system features an interactive component that allows for immediate modification of 2D images via an easy drag-and-drop action, synchronously updating 3D architectural point clouds. The performance of Drag2Build in 3D reconstruction precision was assessed and benchmarked against mainstream methods using ArchiNet. The experiments showed that our approach is capable of producing high-quality 3D point clouds, facilitating swift editing and efficient handling of intricate backgrounds.

Keywords. 3D building Generation, Diffusion Model, Single Image Reconstruction, DragDiffusion.

1. Introduction

The recent progress in transforming 2D images into 3D models significantly benefits spatial data capture and is crucial in fields like GIS, and City Modelling(Melas-Kyriazi et al., 2015). It's important in Architecture, Engineering, and Construction (AEC), with uses from Building Modeling to design and interior projects. However,this field needs further research, particularly during AEC's design stage.

Image-based 3D reconstruction provides an efficient design feedback method in early AEC stages. Yet, current studies often use uniform 3D objects such as hydrants, lacking diverse architectural scenes(Chang et al., 2015). This limits precision and adaptability in reconstructing complex architectural scenes.

Architects need a precise, affordable, fast 3D reconstruction method for early-stage complex architectural designs. It should allow easy 3D model edits for preview. We propose Drag2Build, a point cloud diffusion model for single-view architectural images, utilizing a unique single-image reconstruction method and an advanced segmentation model. It updates 3D models in real-time with user inputs, enhancing accuracy and editability.Our contributions are outlined as follows:

- We have collected ArchiNet, a specialized architectural dataset featuring comprehensive images (architectural line drawings, shaded diagrams, and renderings) alongside corresponding 3D architectural point cloud data, complete with precise camera parameters.
- We propose Drag2Build, a novel interactive 3D reconstruction framework, based on a 3D point cloud diffusion model, capable of editing through drag-and-drop actions. This method generates sparse yet precise 3D point clouds from single images lacking depth information, improving segmentation and recognition with the Segment-Anything object segmentation model. Architects can conveniently adjust and refine the generated 3D models.
- Our extensive experimental evaluations and comparisons with existing baselines establish that our algorithm excels in producing high-quality 3D point clouds, enabling rapid editing and effective handling of complex background scenes.

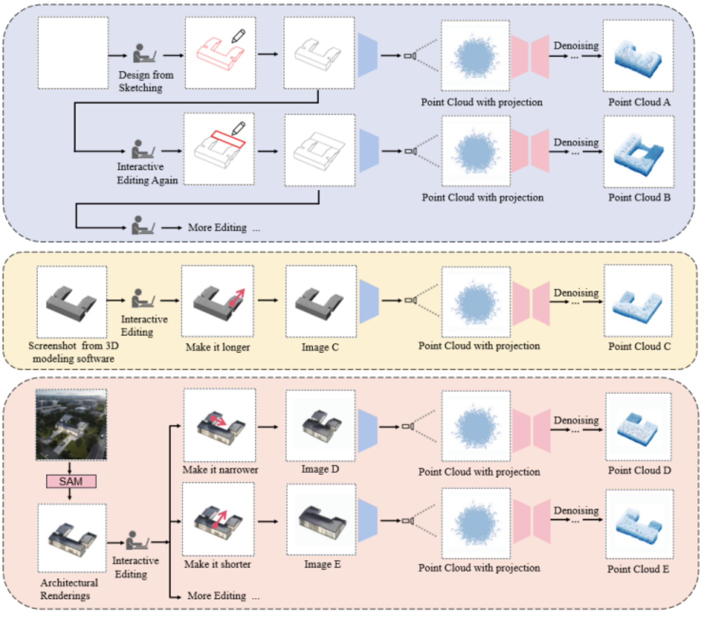


Figure 1. Utilizing Drag2Build, architects are now able to accomplish single-view 3D reconstructions within the design workflow by transforming sketches into three-dimensional models. This platform allows for immediate alterations to 2D illustrations, with concurrent updates being mirrored in the 3D architectural point cloud via straightforward drag-and-drop action.

2. Related Work

3D BUILDING GENERATION

The integration of computer technology and architecture has advanced significantly with the automated generation of 3D models(Wei et al., 2023). Traditional methods relied on fixed rules, limiting variety and increasing architects' workload (Li et al., 2021). Deep learning has brought new opportunities, merging 3D model generation with advanced algorithms (Alidoost. et al., 2019).However, these methods often face challenges in complex scenes and lack flexibility in modifications. Prior research, like Wei's diffusion point cloud model, had limitations in datasets and camera parameters, leading to less accurate reconstructions(Wei et al., 2023).Our approach innovates in 3D architectural reconstruction using a single-view, conditioned projection point-cloud diffusion model. This method enhances point cloud accuracy and allows real-time image-based modifications through user-friendly drag-and-drop interactions, contributing meaningfully to the field of 3D architectural modeling.

SINGLE IMAGE RECONSTRUTION

3D reconstruction from Single-image, blending computer vision and graphics, aims to create 3D scenes from a single 2D image. Traditional methods used 2D and 3D convolutional networks for this transformation, but often fell short in results. With deep learning, the Neural Radiance Field (NeRF) approach, requiring multiple images from different angles, has gained traction (Wang et al., 202) . However, its single-view performance remains subpar.

The rise of diffusion models has led to new single-image 3D reconstruction techniques. 3DFuse enhances diffusion models for better 3D consistency, while Zero123 uses them to render new perspectives based on camera position. OpenAI's Point-E(Nichol et al., 2022) and Shap-E (Jun et al., 2023) use internal models and latent function parameters for point clouds and textured meshes. However, these methods largely rely on common object datasets like Shapenet, with limited resource in specialized architectural datasets, and struggle with complex backgrounds.

POINT-BASED EDITING

To enhance detailed editing, a variety of point-based editing techniques have been introduced. Among these, DragGAN showcased remarkable capabilities for drag-based adjustments, utilizing two key elements: (1) the optimization of latent variable codes to guide the processing point towards its intended position, and (2) a tracking system to monitor the processing point's movement(Pan et al., 2023). Expanding upon this, Mou et al. adapted DragGAN's editing approach for diffusion models, proving its adaptability in various fields. However, despite these methods offering flexible editing between 2D images, they fall short in directly converting 2D edits into 3D structures, a crucial need in architectural design workflows.

3. METHOD

In this part, we explore the methodologies utilized in our innovative framework, Drag2Build. Our platform effectively combines SAM(Kirillovet al., 2022) with methods for converting single images to point clouds, relying on 3D diffusion models. The primary objective is to tackle the intricate task of creating editable 3D architectural reconstructions from single image.Our approach introduces an intuitive, interactive editing feature, capitalizing on the distinct advantages of each integrated component. Through straightforward drag-and-drop actions, users can edit 2D images, and these modifications are immediately reflected in the corresponding 3D point cloud. This integration significantly enhances both the convenience and efficiency of the architectural design.

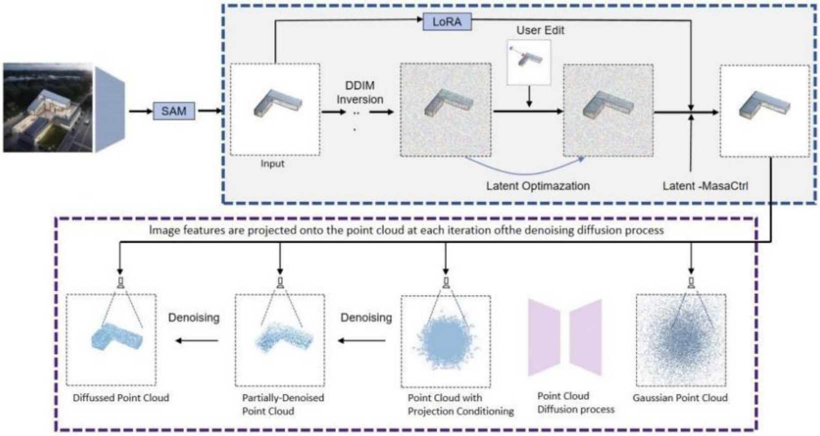


Figure 2. Drag2Build reconstructs architectural point clouds from a single input image and its camera pose in three steps. First, the SAM precisely identifies and segments buildings. Second, the point-based image editing model uses the UNet portion of the LoRA fine-tuned diffusion model to make interactive modifications. Third, the 3D point-cloud diffusion model incrementally generates a point cloud that matches the input image.

3.1. SAM

The SAM is introduced to extract buildings within the given image. To be specific, SAM processes an image, denoted as I, along with a prompt p. an image I and prompt p. SAM then produces a accurate segmentation mask M that effectively masks the background. The process involves the application of an affine transformation to produce localized image segments I' that encompass the bounding box, where the symbol ⊙ denotes element-wise multiplication using the image mask as follows:

$$\mathrm{M} = \mathrm{SAM}(\mathrm{I}, \mathrm{p});\ \mathrm{I}' = \mathrm{Affine}(\mathrm{M} \odot \mathrm{I}; \mathrm{b})\#(1)$$

SAM is a state-of-the-art object segmentation model capable of effectively handling multiple object categories. It addresses the challenge of extracting the principal architectural elements from images with complex scenes.

3.2. POINT-BASED EDITING

Inspired by the impressive success of DragGAN and DragDiffusion, we integrated a similar approach into our model to address the challenges of interactive 2D and 3D modifications in the architectural design process. Initially, we fine-tune the UNet of the diffusion model using LoRA (Augustin et al., 2016) to enhance its accuracy in encoding the features of the input image. Next, we optimize the latent of the diffusion based on user instructions, such as the positions of the handles and target points, along with an optional mask to specify the editable area. This step is crucial to achieve the desired point-based interactive editing. Considering the Markov chain of the diffusion model's latent representation, this step focuses on optimizing the latent representation of single-step diffusion to enhance efficiency and efficacy during the editing process. After completing the optimization of the latent representations, the final edited result is obtained through a denoising step. Similar to DragDiffusion, we employ the Latent-MasaCtrl mechanism to improve the consistency between the original image and edited outcome.As for the differences from DragDiffusion, we have changed the base model and employed ProbSparse Attention, which reduces interaction time and prevents overfitting.

3.3. PROJECTION CONDITONAL POINT CLOUD DIFFUSION MODELING

In machine learning, diffusion denoising probabilistic models are famous for their ability to produce images and data of high quality .These models utilize a progressive approach of noise integration, incorporating disturbances into a sample, referred to as , rooted from the targeted data distribution .Each stage's noise intensity is regulated according to the variance values in , as described below:

$$q(X_t|X_{t-1}) = \mathcal{N}\left(X_t; \sqrt{1-\beta_t}X_{t-1}, \beta_t I\right)\#(2)$$

Every $q(X_t|X_{t-1})$ follows a Gaussian distribution, determined through the widely employed method of reparameterization, as follows:

$$q(X_t|X_0) = \sqrt{\overline{\alpha}_t}X_0 + \epsilon\sqrt{1-\overline{\alpha}_t}\#(3),$$

where $\alpha_t = 1 - \beta_t$, $\overline{\alpha}_t = \prod_{s=0}^{t} \alpha_s$, and $\epsilon \sim \mathcal{N}(0, I)$.

In the development of a generative model, the reverse process is important. It begins with a noisy distribution $q(X_T)$, and reduces noise in the data points, ultimately producing samples that closely mimic the intended distribution $q(X_0)$ .In a 3D point cloud having N points, which is considered as an entity with 3N- dimensions, we train the diffusion model $s_\theta : \mathbb{R}^{3N} \rightarrow \mathbb{R}^{3N}$. This framework is to enhance the accuracy of point positions within a spherical Gaussian domain by converting them into recognizable shapes. The model, at each phase, estimates the positional variance of each point relative to its current location, and repeatedly applies this methodology to generate a sample aligned with the target distribution$q(X_0)$.

This network is specifically trained to anticipate the noise $\epsilon \in \mathbb{R}^{3N}$ injected in the preceding step. It is made by employing the L_2 loss, which calculates the disparity between the groundtruths and those noise values as below:

$$\mathcal{L} = E_{\epsilon \sim \mathcal{N}(0,I)}[||\epsilon - s_\theta(X_t, t)||_2^2]\#(4).$$

In the prediction stage, a arbitrary point cloud $X_T \sim N(0, I_{3N})$ is extracted from a

Gaussian distribution encompassing 3N dimensions. A backward process is then initiated to generate create sample instances X_0. In each stage of this sequence, the average value of the approximation of $q(X_{t-1}|X_t)$ is derived from predictions$_\theta(X_t, t)$, which is subsequently utilized as a foundational element to sample $q(X_{t-1}|X_t)$.

In this framework, 3D reconstruction is viewed as a generative process with condition, where the distribution $q(X_0|I, V)$ is dependent on the provided input image I and its camera view V. To tackle the challenge of geometric consistency, our method involves mapping the image to partially denoised points during each phase of the diffusion sequence. At first, the image undergoes transformation into a dense feature volume, accomplished through the application of a standard 2D image model, such as a convolutional neural network (CNN) or a Vision Transformer(ViT). Before every diffusion step, we map the aforementioned features onto a point cloud.

4. EXPERIMENTS AND RESULTS

4.1. DATASET

Collecting a 3D architectural model database is complex, time-consuming, and resource-intensive, involving data collection, modeling adjustments, and annotations. Traditional 3D point-cloud datasets often face incompleteness, inaccuracies, and limited styles (Li et al., 2021). To evaluate our method, we created ArchiNet, a diverse dataset with 4,688 SketchUp models from 132 architecture students (with a mix of 95 undergraduates and 37 postgraduates) and 675,180 images, each with detailed camera settings. We initially processed these models in Grasshopper to transform them into standardized point-cloud files (PLY format), and subsequently utilized Matplotlib to export images of each PLY file from 36 different angles, ensuring a comprehensive view. Our collection includes three visual types: architectural line drawings (AL), shading diagrams (AS), and renderings (AR), each with point cloud data and individual masks. The dataset is divided into training, testing, and validation sets with a ratio of 92:21:21, facilitating a balanced evaluation of our models across diverse architectural styles and complexities.

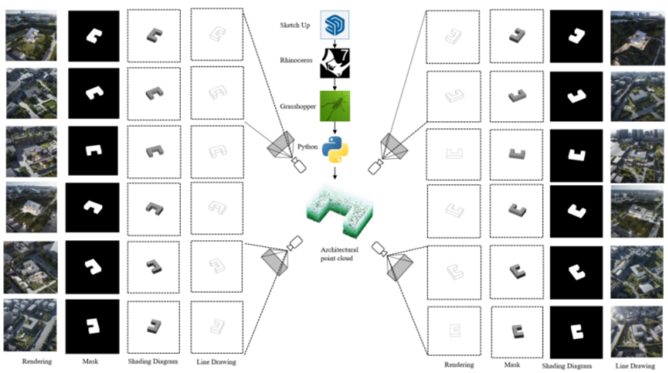


Figure 3. Dataset Collection Process Diagram

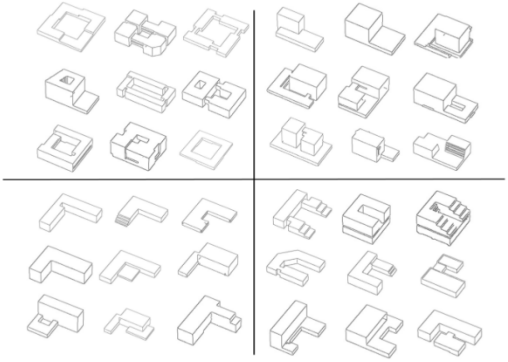


Figure 4. Example instances from our ArchiNet dataset.ArchiNet 3D assets are semantically diverse, highquality, and paired with camera parameters

4.2. EXPERIMENTS DETAILS

In our research, we utilized a high-performance workstation equipped with two NVIDIA A6000 GPUs, each boasting 48 GB of RAM, and powered by an Intel(R) Xeon(R) Gold 6326 CPU. The training phase leveraged our ArchiNet dataset, employing an Adam optimizer. The parameters set for this phase included a learning rate of 0.0001, a batch size of 24, and a cap of 100,000 steps for training. For uniformity and consistency in our analysis, each 3D point-cloud model was standardized into a OBJ file format, containing precisely 8192 points.

4.3. BENCHMARKS AND METRICS

In our study, we compared our model's point clouds with others, focusing on intricate background handling and point-cloud fidelity due to our task's interactive nature. We retrained 3D-R2N2 using our dataset, sticking to its original settings, to create voxel files from images, later converted to point clouds. For Point-E(Nichol et al., 2022), we used official code and models, following their single-image 3D reconstruction examples. These predictions required no extra scaling. With Shap-E(Jun et al., 2023), we maintained hyperparameters and followed their single-image examples, converting meshes to point clouds for consistency in distribution and count.Our evaluation was based on two key metrics:

Fréchet Inception Distance (FID): This metric assesses the disparity between the generated and actual point clouds by comparing the covariance of point cloud features. Lower FID values suggest a closer resemblance to real point clouds.

Chamfer Distance (CD): CD is a prevalent metric for gauging the likeness between two point cloud sets, calculating the mean nearest-point distance between them. Smaller CD values indicate higher similarity.

4.4. QUANTITATIVE RESULTS

Tables 1-2 provide a comprehensive summary of the performance evaluations conducted on the ArchiNet dataset. The data conclusively shows that our model surpasses competing models in all four evaluated metrics. When comparing the mean scores across the three datasets (AL, AS, and AR), our model shows a substantial reduction in Chamfer Distance (CD) results by 42.3%, 43.2%, and 64.2% compared to Shap-E(Jun et al., 2023), Point-E(Nichol et al., 2022), and 3D-R2N2. Concurrently, it demonstrates a remarkable increase in F-Scores by 25%, 104%, and 112%. This superiority of our model in terms of reconstruction accuracy and detail capture is evident. Additionally, the results highlight the variability in the adaptability of different models to distinct types of images. For example, Shap-E exhibits subpar performance with line drawing datasets as opposed to other types, while Point-E struggles more with rendered image data. In contrast, our methodology maintains consistent and stable performance across all three data categories.

Table 1.CD results on ArchiNet. The best results are indicated in bold.

Method	AL	AS	AR
Ours	**0.0828**	**0.0800**	**0.0799**
Shap-E	0.1703	0.1240	0.1263
Point-E	0.1512	0.1505	0.1262
3D-R2N2	0.2276	0.2241	0.2240

Table 2. F-Score results on ArchiNet. The best results are indicated in bold .

Method	AL	AS	AR
Ours	**0.6688**	**0.6894**	**0.6848**
Shap-E	0.4900	0.5771	0.5781
Point-E	0.3159	0.3066	0.3685
3D-R2N2	0.3377	0.3144	0.3152

4.5. QUALITATIVE RESULTS

Beyond the quantitative evaluation, qualitative results are also shown for the ArchiNet dataset (Figures 5-8). These visual examples illustrate how our technique not only facilitates immediate, user-friendly interaction with 2D images via straightforward drag-and-drop actions, directly reflecting changes in 3D architectural point clouds, but also precisely captures intricate architectural elements like platforms, extended volumes, and exterior staircases.

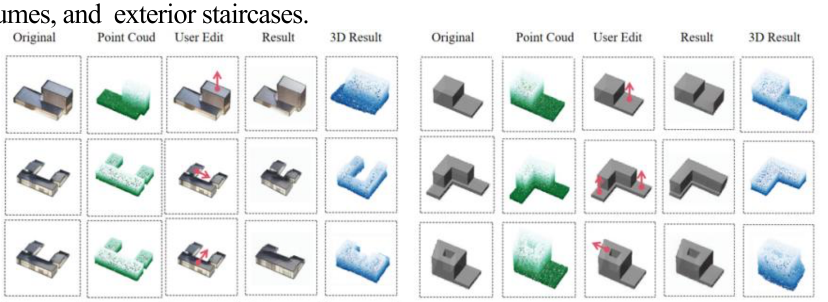


Figure 5: Example results from different methods on the ArchiNet dataset. The image demonstrates our model's ability to modify 3d shapes by editing images

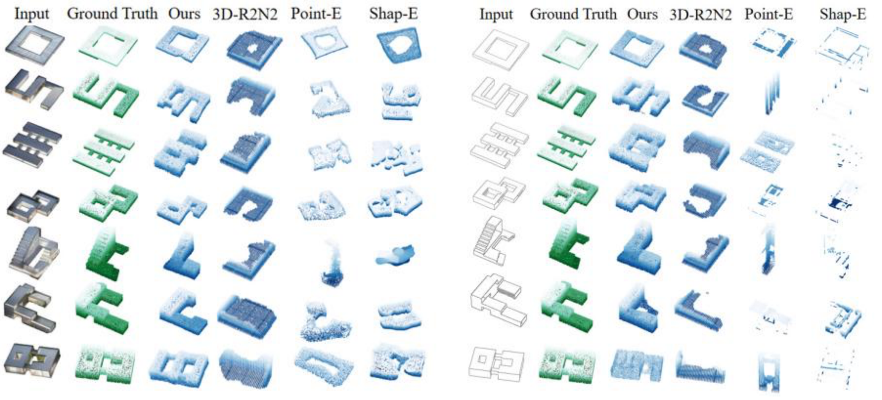


Figure 6.Example results from different methods on the ArchiNet AR dataset and the ArchiNet AL dataset. The first column shows the input image, and the second column displays the ground truth. The subsequent four columns each showcase the point cloud reconstruction results of different models.

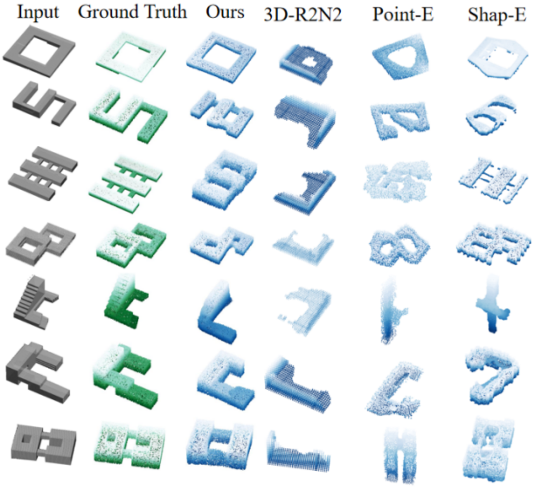


Figure 7. Example results from diverse methodologies applied to the ArchiNet AS dataset are shown.

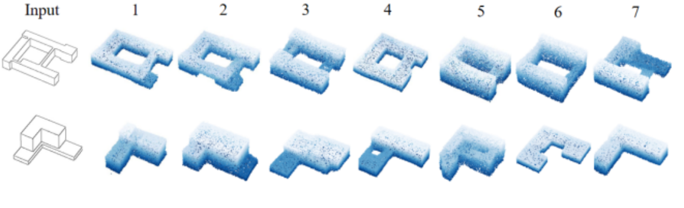


Figure 8.Examples of diversity in the results. The far left column presents a sample image selected from the architectural line drawing collection, notable for its notably ambiguous form. Subsequent images illustrate outputs generated by our model using seven distinct random seeds. Demonstrating

its proficiency, our model consistently generates considerable shape variations, all the while aligning accurately with the perspective of the input image from the reference vantage point.

5. CONCLUSION AND FUTURE WORK

The Drag2Build framework marks a meaningful progress in 3D building generation and interactive architectural design. Looking ahead, our objectives are set towards enhancing and expanding our model in several key areas:

Synergy with Digital Modeling Platforms: Our goal is to craft plugins or tools that seamlessly blend our framework with established digital modeling applications, harnessing the combined strengths of both systems.

Enhanced Precision in Editing: We are dedicated to refining the accuracy of our drag-and-drop features, focusing on the meticulous modification of small-scale elements and intricate details within 3D point clouds.

Advanced Point Cloud Processing: We plan to evolve point cloud processing into formats like meshes, making editing simpler for architects, thus enhancing practicality and utility in real-world architectural applications.

References

Alidoost, F., Arefi, H., & Tombari, F. (2019). 2D image-to-3D model: Knowledge-based 3D building reconstruction (3DBR) using single aerial images and convolutional neural networks (CNNs). Remote Sensing, 11(19), 2219.

Augustin, A., Yi, J., Clausen, T., & Townsley, W. M. (2016). A study of LoRa: Long range & low power networks for the internet of things. Sensors, 16(9), 1466.

Chang, A. X., Funkhouser, T., Guibas, L., Hanrahan, P., Huang, Q., Li, Z., ... & Yu, F. (2015). Shapenet: An information-rich 3d model repository. arXiv preprint arXiv:1512.03012.

Jun, H., & Nichol, A. (2023). Shap-e: Generating conditional 3d implicit functions. arXiv preprint arXiv:2305.02463.

Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., ... & Girshick, R. (2023). Segment anything. arXiv preprint arXiv:2304.02643.

Li, W., Meng, L., Wang, J., He, C., Xia, G. S., & Lin, D. (2021). 3D building reconstruction from monocular remote sensing images. In Proceedings of the IEEE/CVF International Conference on Computer Vision (pp. 12548-12557).

Melas-Kyriazi, L., Laina, I., Rupprecht, C., & Vedaldi, A. (2023). Realfusion: 360deg reconstruction of any object from a single image. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 8446-8455).

Nichol, A., Jun, H., Dhariwal, P., Mishkin, P., & Chen, M. (2022). Point-e: A system for generating 3d point clouds from complex prompts. arXiv preprint arXiv:2212.08751.

Pan, X., Tewari, A., Leimkühler, T., Liu, L., Meka, A., & Theobalt, C. (2023, July). Drag your gan: Interactive point-based manipulation on the generative image manifold. In ACM SIGGRAPH 2023 Conference Proceedings (pp. 1-11).

Wang, Z., Wu, S., Xie, W., Chen, M., & Prisacariu, V. A. (2021). NeRF--: Neural radiance fields without known camera parameters. arXiv preprint arXiv:2102.07064.

Wei, Y., Vosselman, G., & Yang, M. Y. (2023). BuilDiff: 3D Building Shape Generation using Single-Image Conditional Point Cloud Diffusion Models. In Proceedings of the IEEE/CVF International Conference on Computer Vision (pp. 2910-2919).

PSEUDO-CROSS-MODAL TRANSLATION

Bridging Architectural Plan and Perspective through a pix2pix Network

YUJUN MAO[1], WENZHE PENG[2] and TAKEHIKO NAGAKURA[3]
[1]*Researcher Computational Design.*
[2,3]*MIT Department of Architecture.*
[1]*arch.yjmao@gmail.com, 0009-0008-8762-3388*
[2]*pwz@alum.mit.edu*
[3]*takehiko@mit.edu*

Abstract. Architectural pedagogy often segments designs into diverse representation forms like plans and renderings. With AI's growing influence on early design through GANs, Midjourney, and Stable Diffusion, there remains a gap in translating between diverse architectural representations, a phenomenon we term 'Pseudo-cross-modal Translation', indicating the indirect transformation between non-analogous architectural representations. Addressing this, our research hypotheses a practical need and actionable possibility to translate architectural plans into perspective renderings via neural networks, exploiting the information differences between them. We navigate this intricate translation utilising a pix2pix network of which the dataset encompasses plans with designated view cones and corresponding rendered perspectives. The training data are sampled from the model of Mies van der Rohe's Barcelona Pavilion and its variations. Evaluations through perceptual surveys, which incorporate modifications in information complexity of plans, illuminate the neural networks' nuanced capability to bridge plans and perspectives under various conditions. Our results not only validate this translation but also spotlight the computational statistics' latent potential in deciphering unseen spatial features from the variance between plans and perspectives. This work unveils a novel method for generating architectural imagery, promoting a holistic spatial understanding.

Keywords. Plan, Perspective, pix2pix, Architectural Representation.

1. Introduction

The comprehension and development of spatial designs through conventional architectural pedagogy, which employs various representational forms like plans and perspectives, can be challenging especially for novices. Research, including Nagakura's experiments (Nagakura and Sung, 2014), indicates the intricacy of constructing a mental spatial image from isolated representations and the necessity of

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 179-188.

establishing a back-forth process of understanding between various representations conveying information with distinct characteristics.

Prior to the modern computer-aided design tools, translation between various architectural representations throughout history is articulated either through laying out them in attempts to establish simple spatial associations or by the techniques of perspective to represent three-dimensional objects in two dimensions. Computer graphics technology has enabled instant perspective projection from a three-dimensional geometric model. Modern computer-aided design tools, such as Building Information Modelling (BIM), further provide convenience for the aforementioned issue by integrating different media into a unified symbolic data structure (Coons, 1963). However, the construction of such integration on the one hand is time consuming requiring professional skills and on the other hand, it can potentially compromise design creativity due to its deterministic nature - In actual design thinking, a single design representation can be linked to various possibilities in another representation. There remains a gap in translating between different architectural representations in a non-symbolic, analogous manner which we termed "Pseudo-cross-modal Translation', such that facilitating design creativity. And with the rising impact of AI on early architectural design stage, particularly through Generative Adversarial Networks (GANs) to the more recent Midjourney and Stable Diffusion recognized for their intensive iterations of prompt to image translation, it is possible to integrate various architectural representations into a computational statistical structure.

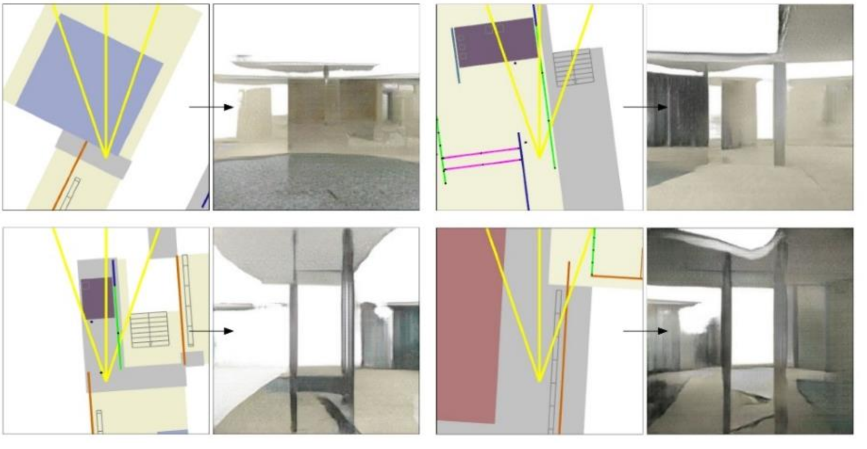


Figure 1. Examples of Pseudo-cross-modal Translation

In response, our research postulates both a pragmatic necessity and a feasible potential for translating architectural representations through neural networks, exploiting the disparities of information between them. Specifically, this research project focused on transforming architectural plans into perspective renderings using a modified pix2pix neural network, addressing the challenge of 'Pseudo-cross-modal Translation'. (Figure 1). Through training on pairs of notational plans with corresponding perspective renderings, the network not only converts one representation to another, and further provides the possibility to go beyond the rigid correspondence of symbolic data structure in the current computer-aided design tools allowing flexibility and ambiguity to occur. This research aims to foster creativity catering to a more nuanced spatial understanding through moving from a symbolic to an analogous mode of representation translation and to benefit architects by democratising the design evolution process to various stakeholders.

2. Related Works

Translation between various representations to better depict a three-dimensional world in two-dimensional space has a long history. From pre-renaissance cartography art that associates views from two or three different angles on the same flat surface to the rediscovery of perspective technique, by Filippo Brunelleschi at the beginning of the 15th century, such desire of depiction throughout history is motivated by socio-technical development at different stages. On the one hand, the technique of lacking a unified viewpoint is inherited in laying out different representations to establish simple spatial associations, for instance Palladio's aligned architectural drawings to document Temple of Mars the Avenger (Palladio, 1570) and ‘Analytique’ drawings in Beaux arts tradition (d’Espouy, 1905). On the other hand, the perspective method that directly reflects human visual experience is later developed into more advanced descriptive geometry, which employed by Ferdinando Galli-Bibiena to depict a stage design for a scena par angolo translating a floor plan to an interior perspective drawing (Vesely, 2004).

After entering the era of computer graphics and with the development of rendering algorithms, it became an easy task to produce photorealistic images using three-dimensional models. However, such translation between representations is time-consuming and the requirement for drafting techniques have changed to the skills of modelling technique and the understanding of rendering software. To respond to this, Takehiko Nagakura’s research projects such as ‘Space Barcoder’ (Nagakura, 1998), ‘Digitarama’ (Nagakura, 1997), ‘Deskrama’ (Nagakura and Oishi, 2006) and the more recent ‘Ramalytique’ (Nagakura and Sung, 2014) utilising contemporary technology explore the translation between various representations through the spatial arrangement approach, which is efficient in synchronisation sense and makes it user-friendly even for novices.

With the advance of artificial intelligence, Generative Adversarial Network (GAN) (Goodfellow et al., 2014) has gained significant attention in the arts and design for its remarkable ability to learn and generate creative products. Specifically, pix2pix as a conditional GAN (cGAN), provides general-purpose solutions to image-to-image translation problems (Isola et al., 2018). Previous works relevant to the research purpose of translation between representations includes floor plan generation (Huang and Zheng, 2018) of which the focus was on the conversion from a land plot to a furnished plan, structural analysis during the iterative design phase (Rossi and Nicholas, 2018) by substituting analytical finite element analysis (FEA) modelling with cGAN predictions, and exploration of cGAN’s capacity in generalising different perspective of the urban street views through translation from depth map (Steinfield, 2019).

3. Methodology

In this study, we focus on a significant and impactful issue of Pseudo-cross-modal Translation: transforming plans into perspective renderings. The Pseudo-cross-modal Translation from architectural plans into corresponding perspective renderings is formulated as a task which fine-tunes and tests a customised version of pix2pix neural network using plan-rendering image pairs (Figure 2). Here, we describe the stages of data preparation and neural network training.

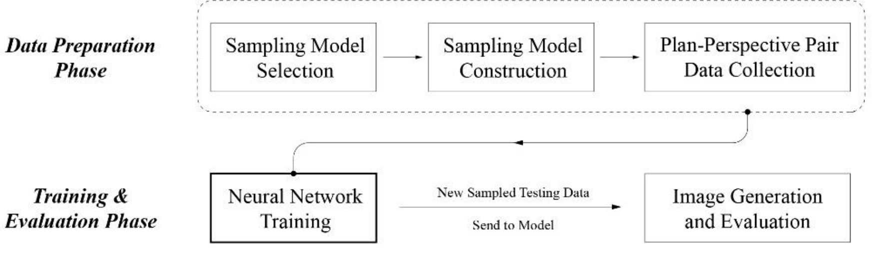


Figure 2. The workflow diagram of Pseudo-cross-modal Translation

In the data preparation stage, we first select Mies van der Rohe's Barcelona Pavilion as the base of data sampling models and create its six variations through a discretized and randomly recombined approach. Then, 1,500 pairs of plan-perspective are sampled from each variation based on a labelling principle. Finally, 10,500 pairs of samples are processed to serve as the training data for a pix2pix neural network.

In the neural network training stage, we employ the collected data pairs to train a pix2pix neural network capable of generating perspective rendering images. The particular implementations of the pix2pix model are discussed in detail below.

3.1. DATA PREPARATION

The majority of the data preparation work required by this project consists of four steps: sampling model selection, sampling model construction, label colour coding, and plan-perspective pair data collection, elaborated below.

3.1.1. Sampling Model Selection

The Barcelona Pavilion designed by Ludwig Mies van der Rohe in 1929 is an important building in the history of modern architecture, known for its continuous space and spectacular use of extravagant materials. It is referenced by Immanuel Koh to illustrate Kengo Kuma's concept about 'particlised architecture' (Kuma, 2012). And according to Koh, the particlised interpretation of Barcelona Pavilion is reinforced by its use of pilotis and plinth, as a formal means to isolate their architecture from the ground, especially when viewed as images (Koh, 2023). Based on this interpretation, in this research, we decide to select Barcelona Pavilion as the basis for the construction of sampling models.

3.1.2. Sampling Model Construction

In order to achieve a good neural network performance, training data with adequate visuospatial quality is the key, thus it is necessary to acquire different variations of Barcelona Pavilion to be sampled for training. The Barcelona Pavilion model is firstly broken down into 18 discrete elements with a 3 by 6 subdivision, which is an alternative reading resonate with the concept of 'particlised architecture' allude to a mechanic form of statistical seeing. The distinction of this manipulation is that figure-and-ground and part-to-whole are completely abandoned. Each decomposition unit is interchangeable without any predefined rules of assemble, where parts could be randomly recombined within the original grid to form a possible whole (Koh, 2023). In this sense, Barcelona Pavilion and its variations conceptually can be regarded as a 'training set' that embodies a hidden range of likely shapes within the same architectural vernacular. We ultimately end up with 6 more variation of the Barcelona Pavilion models. The 7 models, including the Barcelona Pavilion itself, are then used for the plan-perspective pair data collection in later stages (Figure 3).

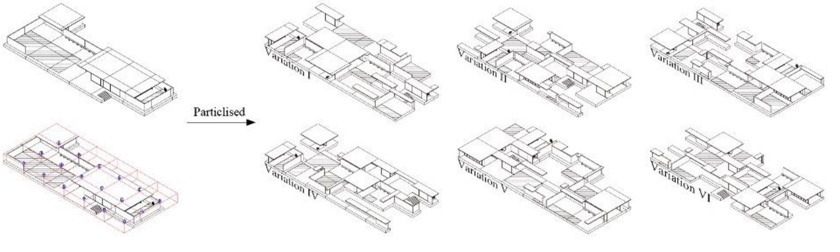


Figure 3. The diagram of sampling model construction

3.1.3. Label Colour Coding

The training of a pix2pix requires pairs of input and output images to learn the translation patterns. And from the perspective of a user of a trained pix2pix network, we offer an input image that conforms to some mapping convention and receive in return a synthetic image that results from the transformation of this input. In such a translation, the key is the correspondence between the input and output elements. Hence, the design of label colour coding to assign various unique RGB colours to each kind of architectural element on the plan becomes particularly important. Figure 4 shows ten colours assigned to represent green alpine marble wall, roman travertine marble wall and floor, transparent glass, grey frosted glass, area covered with carpet, pool with green alpine marble floor, pool covered with cobblestone, roof covered area and steel framework respectively leaving the background with pure white. The selection of colour was based on the principle to differentiate the vertical architectural elements with horizontal ground and canopy as far as possible. Each sampled plan was marked with an isosceles triangle representing the view cone of the camera in yellow of which the middle line represents the viewing direction, and the side lines represent the field of view of the camera in the corresponding perspective view. The view cone is positioned at three-quarters along the central axis of the sampled plan, enabling the assessment of environmental factors behind the camera, like reflections and shadows.

Figure 4. The diagram of label colour coding

3.1.4. Plan-Perspective Pair Data Collection

Since the objective is to translate architectural plans into their corresponding perspectives, a collection of plans-perspectives pairs is utilised to construct a dataset for training a pix2pix neural network. Our dataset is composed of 10,500 plan-perspective pairs. And the dataset pairs are sampled from Barcelona Pavilion and its particlised variation based on the labelling principle illustrated in previous section.

Our study utilises a Grasshopper program to automatically sample architectural plans. It captures plan elements in a specified location with appropriate range, rotation, and colour-coded labels for elements. Figure 5 illustrates the crucial steps of this process: rotation and cropping. This approach deconstructs the building architecturally, aligning it with the camera's viewpoint for uniform information capture. We ensure the sampled plan rotates to keep the camera facing upward during sampling. Cropping is used to enhance the neural network's efficiency in translation and its capacity for generalising and inferring missing elements, a concept we delve into in the evaluation section. We set the cropping size to 10x10 metres, balancing coverage with generalisation capabilities. For perspective renderings, they are generated in V-Ray, using the camera positioned and oriented as per the plan sampling.

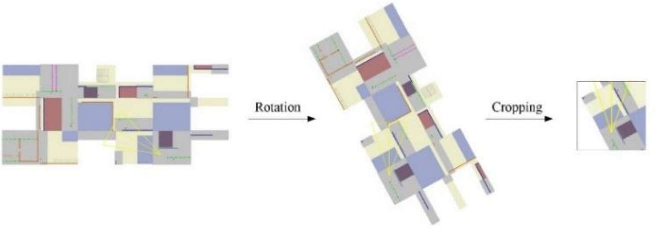


Figure 5. The diagram of rotation and cropping manipulations

3.2. NEURAL NETWORK TRAINING

This project employs the PyTorch version of pix2pix (256px x 256px) architecture. And to accommodate our purpose of experiment, some tunings are applied. We trained the network for 200 epochs with a start learning rate of 0.002 and a batch size of 1. The training, using a T4 GPU on Google Colab with Python 3, takes about 33 hours for a dataset of 10,500 images. Once trained, the neural network, following the pix2pix architecture, generates synthetic RGB images from input cropped-plans. The results and performance evaluation of this network are detailed in the subsequent section.

4. Evaluation of Trained Neural Networks

Three types of studies of the trained neural network are conducted to respectively evaluate its efficacy in translation, its ability to generalise, and its potential for creative outputs. Since the aim of this study is to explore the potential feasibility of leveraging neural networks to achieve an analogous mode of representation translation moving from a symbolic one, a perceptual survey rather than a quantitative method is adopted for evaluation.

A perceptual survey enables the comprehension of human opinions, facilitating the assessment of a crucial aspect of the proposed system: producing results that align closely with human perceptions of space. In this research, six volunteers are invited and divided into two groups - architectural professionals and novices. Each participant takes three questionnaires corresponding to three methods of the trained neural network exploration. Every questionnaire consists of twenty single-choice survey questions; therefore, sixty questions are taken in total. All participants will receive an overview of the research motivation, methodology, and the high-level objectives of each perceptual survey, along with an introduction to the labelling principles before participating in the survey. The following survey question is posed to the participants: “Which of the four perspectives in the right column filling into the blank area in the left best reflects its relationship with the plan above?” Only one of the four images in the right column is generated by the plan as input feeding to the trained neural network while the other three are generated by random input. The labelling principal is placed on the left side of each page for reference.

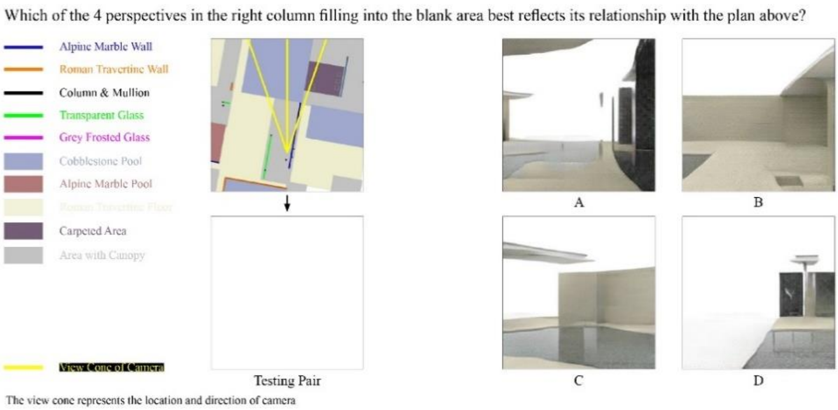


Figure 6. The format of perceptual survey questions

4.1. STUDY 1: TRANSLATION EFFICIENCY OF NETWORKS

The first study focuses on assessing the translation efficiency of a system that converts architectural plans into renderings without direct training on these inputs. Initial results, as shown in Table 1, indicate effective transformation, confirmed by a perceptual survey where both professional and novice groups performed similarly, with most participants only answering 2-3 questions incorrectly out of 20. An example survey (Figure 8) analyses the system's ability to differentiate between interior and exterior perspectives, identifying interior spaces through specific spatial elements like a roof or view cone placement. Despite some discrepancies, such as the absence of transparent glass or inaccurate carpet depiction, the system effectively translates plans into spatially coherent perspectives, highlighting its proficiency in interpreting architectural information without complete 3D designs.

Survey I	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	Correct	Wrong
Professional.1	√	×	√	√	√	×	√	√	√	√	√	√	√	√	×	√	√	√	√	√	**17**	3
Professional.2	√	×	√	√	√	×	√	√	√	√	√	√	√	√	√	√	√	√	√	√	**18**	2
Professional.3	×	×	×	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	**17**	3
Novice.1	√	√	×	√	×	√	√	×	√	√	√	√	√	√	√	√	×	√	√	√	**16**	4
Novice.2	√	√	√	×	√	×	√	√	√	×	√	√	√	√	√	√	√	√	√	√	**17**	3
Novice.3	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	×	√	√	√	**19**	1

Table 1. The results of the perceptual survey I

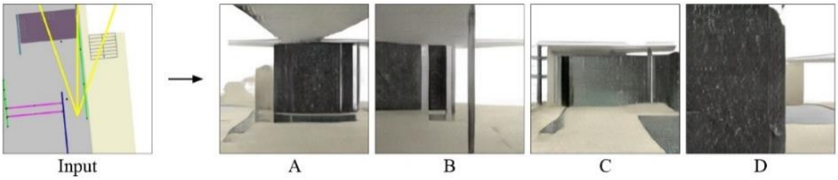


Figure 7. An example of the perceptual survey I

4.2. STUDY 2: GENERALISATION ADAPTABILITY OF NETWORKS

In the second study, plans based on the Barcelona Pavilion were used to assess the network's ability to adapt learned styles to new spatial configurations. The results, detailed in Table 2, revealed that professionals generally outperformed novices, with the latter group making more errors on average. An example survey (Figure 9) involved analysing a plan, where a view cone under a canopy pointed towards a small pool, helping to eliminate certain options. Despite some inconsistencies, such as the absence of grey frosted glass and the relative position of a green wall and pool, the chosen option best represented the spatial configuration, illustrating the neural network's effectiveness in generalising styles to new design contexts.

Survey II	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	Correct	Wrong
Professional.1	√	√	√	×	√	√	√	√	×	√	√	√	√	√	√	√	√	√	√	√	**18**	2
Professional.2	√	√	√	√	√	√	√	√	×	×	√	√	√	√	√	×	√	√	√	√	**17**	3
Professional.3	√	√	√	√	√	√	√	√	×	×	√	√	×	√	√	√	√	√	√	√	**17**	3
Novice.1	√	×	×	√	√	×	√	√	√	√	√	√	×	√	√	√	√	√	×	√	**15**	5
Novice.2	√	√	√	×	√	√	√	×	×	×	√	√	√	√	√	√	√	√	√	√	**16**	4
Novice.3	×	√	√	√	√	√	√	√	×	×	√	√	√	√	√	×	×	√	√	√	**15**	5

Table 2. The results of the perceptual survey II

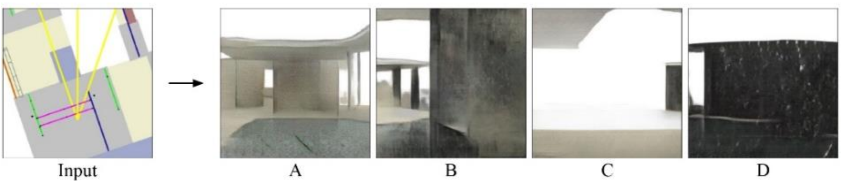


Figure 8. An example of the perceptual survey II

4.3. STUDY 3: CREATIVE FLEXIBILITY OF NETWORKS

The third study examines how trained neural networks handle creative tasks like cropping, colour alterations, and element removal, assessing their ability to infer missing details from the training dataset. Results in Table 3 show that professionals outperform novices in recognizing the network's inference capabilities. A detailed analysis of one survey (Figure 10) reveals swapped labelling of pool colours and marble floor, with a view of water and roof. Despite some confusion, option 'C' is identified as most accurate, showcasing the network's adeptness at managing missing information under complex input variations.

Survey III	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	Correct	Wrong
Professional.1	√	√	√	×	√	×	√	√	×	√	√	√	√	√	×	√	√	√	√	√	**16**	4
Professional.2	√	√	√	√	√	√	√	√	√	√	√	√	√	√	×	√	√	√	√	√	**19**	1
Professional.3	√	√	√	√	√	√	√	√	√	√	√	√	√	×	×	√	√	√	√	√	**18**	2
Novice.1	√	×	√	√	√	×	×	√	√	×	√	√	√	×	√	√	√	√	√	×	**14**	6
Novice.2	√	×	√	×	√	√	√	√	×	√	√	√	×	√	×	×	√	√	√	√	**16**	4
Novice.3	√	√	√	×	√	√	√	√	×	√	√	√	×	√	×	√	√	√	√	√	**16**	4

Table 3. The results of the perceptual survey III


Input A B C D

Figure 9. An example of the perceptual survey III

5. Conclusion and Reflection

This project explores the potential of pix2pix neural networks, a form of conditional Generative Adversarial Networks (cGANs), for transforming architectural plans into perspective renderings. The Pseudo-cross-modal Translation system developed in this study simplifies the design process by converting plans into visual perspectives without complex 3D modelling, offering a more accessible and creative approach to architectural design, particularly for novices. Our method stands out from conventional Computer-Aided Design techniques by its ease of use, adaptability to various design layouts, and its potential to enhance creativity, making it an invaluable brainstorming tool that democratises design capabilities and fosters innovative thinking in early-stage design.

The current state of research in generative design, particularly in translating various architectural representations, still faces limitations in terms of efficiency, accuracy, and diversity. The existing pix2pix model struggles with fine details in translation, such as columns and glass, suggesting a need for more robust or custom neural networks, possibly diffusion-based, for enhanced accuracy. Furthermore, the network's limited capability in generalising learned styles highlights the potential of multi-modal neural networks to improve efficiency and diversify outputs. Additionally, expanding the translation capabilities beyond just architectural plans to perspectives, including reverse translation and other spatial representations, would further streamline and inclusivize the design process.

The progress in AI and its understanding of spatial design significantly impacts not only architecture but also fields like virtual reality, game development, filmmaking, and robotics. This democratises design, allowing greater involvement from amateurs and reducing misunderstandings in design communication. Emerging technologies could shift architects' roles from routine tasks to developing design frameworks.

References

Coons, S. A. (1963, May). An outline of the requirements for a computer-aided design system. In *Proceedings of the May 21-23, 1963, spring joint computer conference* (pp. 299-304).

d'Espouy, H. (1905) *Greek and Roman Architecture in Classic Drawings*. Translated by Henry Hope Reed., 1999. New York: Dover Publications.

Goodfellow, I., J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. 2014. “Generative Adversarial Nets.” In *Advances in Neural Information Processing Systems*, 2672–80.

Huang, W., & Zheng, H. (2018). Architectural drawings recognition and generation through machine learning. In P. Anzalone, M. Del Signore, & A. J. Wit (Eds.), *Recalibration: on imprecision and infidelity: Proceedings of the 38th Annual Conference of the Association for Computer Aided Design in Architecture*, ACADIA 2018 (pp. 156-165).

Isola, P., Zhu, J. Y., Zhou, T., & Efros, A. A. (2017) Image-to-image Translation with Conditional Adversarial Networks, 217 *IEEE Conference on Computer Vision and Pattern Recognition* (pp.1125-1134). Available at: https://doi.org/1.119/CVPR.217.632

K. Kuma (2012), *Anti-object: the dissolution and disintegration of architectures*, Reprint., vol. 2. London: Architectural Association Publications.

Koh, I. (2023). Architectural sampling: three possible preconditions for machine learning architectural forms. *Architecture Intelligence*, 2(1), 7.

Nagakura, T. (1997) DIGITARAM. In: Ken Sakamura and Hiroyuki Suzuki, ed. 1997. *The Virtual Architecture*. Tokyo: Tokyo University Digital Museum.

Nagakura, T. (1998) Gushikawa Orchid Center. 17th Exhibition of Winning Architectural Models and Drawing. *SD Review*, December 1998, pp.36-38.

Nagakura, T. and Oishi, J. (2006) Deskrama. In: *Proceedings of ACM SIGGRAPH 1998, Emerging technologies*. Article No. 6. New York: ACM New York.

Nagakura, T., & Sung, W. (2014, November). Ramalytique: Augmented reality in architectural exhibitions. In *Conference on Cultural Heritage and New Technologies 19th Proceedings* (pp. 3-5).

Palladio, A. (1570) *The Four Books on Architecture*. Translated by Robert Tavernor and Richard Schofield., 1997. Cambridge, Massachusetts: MIT Press.

Rossi, G., & Nicholas, P. (2021). Encoded Images: Representational protocols for integrating cGANs in iterative computational design processes. In *Acadia 2020 Distributed Proximities: Proceedings of the 40th Annual Conference of the Association for Computer Aided Design in Architecture* (Vol. 1, pp. 218-227).

Steinfeld, K. (2019, October). Gan Loci. In *Proceedings of 39th Conference of the Association for Computer Aided Design in Architecture: Ubiquity and Autonomy* (pp. 392-403).

Vesely, D. (2004). *Architecture in the age of divided representation: the question of creativity in the shadow of production*. MIT press.

THE INTERSECTION OF TECHNOLOGY AND ARCHITECTURE: SMARTPHONE PHOTOGRAPHY IN CARBON ANALYSIS

IUAN KAI FANG[1] and SHEN GUAN-SHIH[2]
[1,2]*Department of Architecture, National Taiwan University of Science and Technology.*
[1]*iuankai.fang@gmail.com, 0009-0002-7693-2579*
[2]*sgshih@mail.ntust.edu.tw, 0000-0001-7108-9960*

Abstract. Our research introduces an innovative methodology that employs smartphone imaging for measuring dimensions and utilizes deep learning to estimate carbon emissions associated with facade materials. The dimensions of various components of building exteriors are obtained through smartphone imaging, and a network model on a cloud server automatically segments these components in the images, calculating their respective areas. By combining user-input material specifications such as thickness and density， with standard values of material carbon coefficients, estimations for each component's material carbon footprint are derived. This approach offers the advantage of individual estimations for diverse materials, aiding in the design of low-carbon facades. Additionally, it features a user-friendly interface enabling swift carbon estimation through portable devices. The method provides a convenient and efficient means for assessing carbon emissions in building facades, contributing to sustainable efforts and informed material selections for a greener future.

Keywords. Carbon Emission, Façade, Part Segmentation, Smartphone.

1. Introduction

Life Cycle Assessment (LCA) considers the carbon emissions associated with materials' production, transportation, installation, and usage phases. It calculates the overall carbon footprint of entire building materials, including those used for the exterior appearance of buildings. While the façade constitutes a part of the architectural framework, distinct from the main structure and mechanical systems, its design plays a pivotal role in enhancing environmental conservation, energy efficiency, and sustainable development. For instance, the Double-Skin Façades (DSF) design aims to enhance a building's energy efficiency, reduce environmental impact, and promote energy conservation and sustainability. (Andrea Zani et al., 2021). By isolating external climate influences, improving thermal performance, and minimizing energy wastage, it balances indoor comfort and environmental friendliness (Sabrina Barbosa & Kenneth Ip,2014) (Baldinelli, 2009).

However, improvement projects for façade in compliance with green building

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 189-198.

standards often achieve energy-saving objectives without accounting for the carbon footprint generated by components enveloping the façade. This oversight might inadvertently harm environmental sustainability despite achieving energy-saving effects(Zahra S. Zomorodian & Mohammad Tahsildoost,2018).

According to the Ministry of Environment in Taiwan's Voluntary Disclosure Review (VDR) report for sustainability from 2020 to 2022 released in 2023, there's a severe shortage of certified personnel in carbon footprint assessment, leading to numerous companies facing challenges in obtaining carbon footprint certifications. When considering ESG regulations, the carbon footprint of existing corporate office spaces, factories, and other buildings serves as a significant evaluation criterion for a company's compliance with ESG standards. There's an urgent need for a rapid and convenient assessment of building carbon emissions.

In light of those issues, we are endeavoring to explore a path from the domain of computer vision. We aim to research a method capable of using smartphone imaging to capture real-world architecture, subsequently enabling the identification and quantitative assessment of visible exterior materials of buildings with different textures. The carbon footprint of a product can be calculated using methods such as direct monitoring instruments, energy and mass balance, or emission factors. However, within the construction industry, the emission factor method is commonly adopted. This method employs the fundamental formula: the carbon footprint of activity equals the activity data (mass/volume/kilowatt-hours/kilometers) multiplied by the emission factor (per unit of carbon dioxide equivalent). Mass equals volume multiplied by density; hence, once the volume and density of a material are known, its carbon footprint can be calculated.

2. Methodology

2.1. DEVELOPMENT PROCEDURE

The development process of the entire system begins with smartphone imaging and measurement of building dimensions. The captured images and dimensional information are then transmitted to a cloud server. Within this cloud server, there are two integrated systems: one is a proprietary model that improves upon the Generative Adversarial Network (GAN)(Ian Goodfellow et al.,2014) using Pix2Pix(Phillip Isola et al.,2017) as a baseline. It incorporates self-attention and self-proliferation mechanisms to enhance feature weights, accomplishing the semantic segmentation task within the images, and enabling precise identification of building components in photos. The other system is an expert system composed of prior knowledge. It encodes different component material thicknesses, densities, and carbon emission coefficients into the program. Segmented image blocks are analyzed within this system to compute individual material areas, estimating their areas based on RGB pixel points.

The user interface on the smartphone will display recognized building components, presenting a menu for users to select potential materials and specifications based on their professional construction experience, as illustrated in (Figure 1.). For instance, if the component is a window, the user can choose the possible thickness values for the glass. Then, the total area of the segmented and identified glass windows multiplied by the selected thickness provides the total volume. The cloud server matches the density and carbon emission coefficients of that material from the expert system, multiplying them accordingly, and transmits the results to the smartphone's user interface. The interface lists the estimated carbon emissions of various materials captured in the smartphone's images of the facade, culminating in the final comprehensive data result.

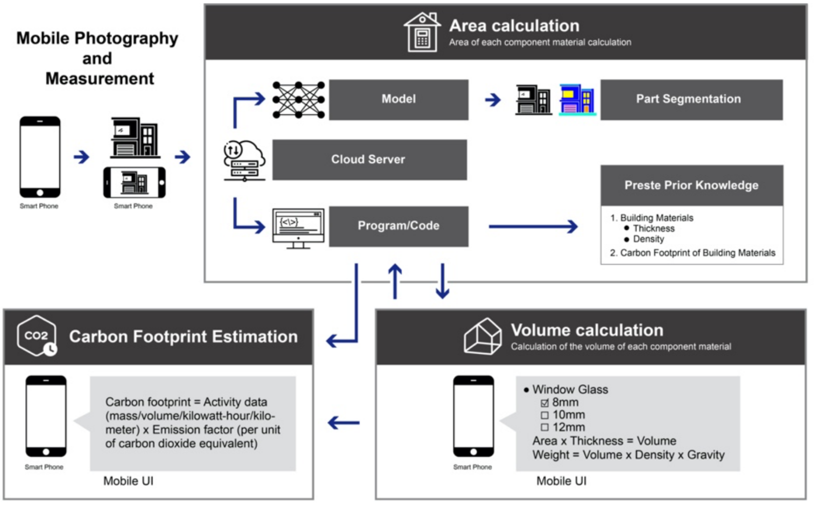


Figure 1. The technical development process diagram of our study.

Our research focuses primarily on the part segmentation of the facade throughout the development process. The functionalities of smartphone imaging and measurement are already well-established, and various countries have established standard reference carbon coefficients for different building materials. Hence, our emphasis lies in the automatic identification of building exterior components, followed by the computation of component quality. Assuming precise measurements of component quality, the accuracy of carbon emission estimation would also improve. Therefore, within the entire development process, the precision of part segmentation stands as our paramount objective in machine learning.

2.2. DATASET

The process of semantic segmentation involves assigning each pixel in an image to its corresponding class label. The value of semantic understanding lies in providing explanatory categorization for meaningful objects in the real world. In contrast to object detection and recognition, semantic segmentation achieves pixel-level classification of images. However, current semantic understanding has not shown substantial development and progress in the tasks of component identification and segmentation in real-world object images. This implies that apart from leveraging large-scale image datasets with diverse labels such as ImageNet (Deng, Jia, et al.,2009), ShapeNet(Angel X.Chang et al.,2015), COCO (Tsung-Yi Lin et al.,2014), ADE20K(Bolei Zhou et al.,2017), etc., for the recognition and understanding of real-world objects, there's still limited inference regarding the finer constituents and functionalities of these objects.

ImageNet, ShapeNet, COCO, ADE20K, and others are renowned large-scale datasets comprising over 200,000 images. However, in comparison to all the objects in the world, their image data for building exteriors is relatively limited. Moreover, part annotations for objects are absent in ImageNet, ShapeNet, COCO, and ADE20K. Our self-built dataset enhances the recognition of information regarding the exterior components of buildings.

Firstly, we collected 500 photographs of building exteriors, forming what we call Raw Data. Next, we removed backgrounds and unnecessary obstructions from the building images, creating what we term Object Data(Figure 2). Subsequently, the components within the building exterior were categorized into ten main types: doorway, window, exposed beam, exposed column, exposed slab, projecting balcony slab, railing, eaves, decorative wall panel, and curtain wall. These ten types were manually annotated based on functional attributes, establishing what we refer to as Notation Data (Figure 2).

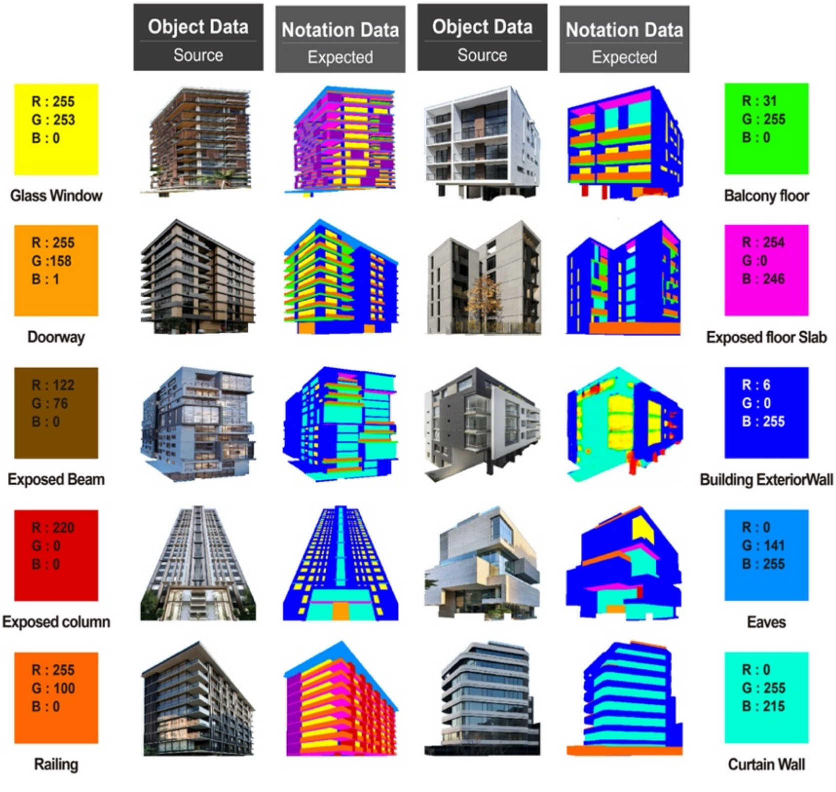


Figure 2. RGB Setting, Object data samples and Notation data samples.

The Façade dataset from UC Berkeley's official directory of Pix2Pix Datasets contains 506 photos of building façades alongside corresponding annotated images. However, aligning with our research objectives, capturing entire building façades conveniently with a mobile phone camera might not always be feasible due to perspective angles. Hence, the Façade dataset might not be entirely suitable for our machine-learning purposes. Moreover, the dataset lacks finer categorization for external building components. Additionally, the BuildingNet (Pratheba Selvaraju et al.,2021) dataset comprises 2,000 annotated architectural models, but it's a 3D model dataset, differing from our 2D real image data. The annotation methods also differ logically. Nonetheless, in the future, we could render BuildingNet's 3D architectural models into 2D images to expand our dataset. However, currently, we prefer using actual images as our experimental foundation to approach real-world scenarios. Annotating real-world photos for component classification aims to enable our trained network models to be practically applied and achieve high precision in real-world scenarios. The genuine composition of real objects offers reliable volumetric forms for carbon footprint calculations.

2.3. MODEL

Once the components of a facade are accurately identified, estimating carbon emissions through the calculation of material coverage areas becomes feasible. This approach offers a rapid initial estimation for broad-scale carbon calculations. To achieve this goal, we propose a novel part segmentation model. Given Pix2Pix's room for improvement in accuracy regarding target details and boundaries, we introduced two techniques(Figure 3): self-proliferation(Yuan-Fu Yang & Min Sun,2021) and self-attention(Ashish Vaswani et al.,2017). The self-proliferation aids in generating meaningful feature maps, while self-attention provides a more refined way to enhance features for improved precision in semantic understanding and part segmentation.

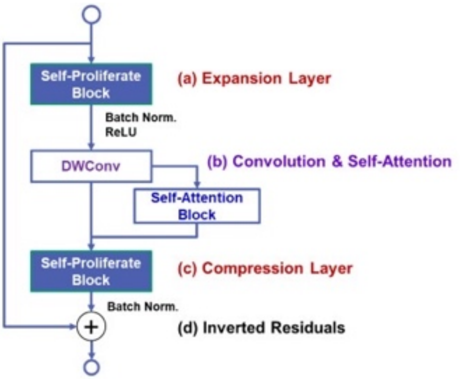


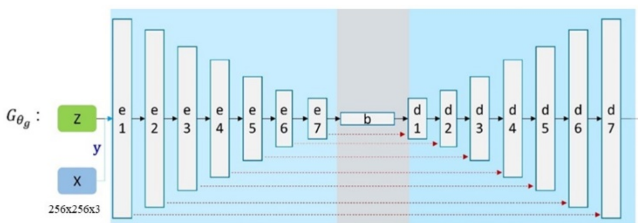
Figure3. incorporated self-attention and self-augmentation mechanisms into our baseline model.

Image segmentation requires positional information for each pixel. It demands full-resolution semantic estimation, making it impossible to reduce computational complexity using pooling or dilated convolution networks as in classification tasks. Therefore, we adopted an encoder-decoder model structure. In the Pix2Pix framework, the encoder part utilizes downsampling to reduce spatial resolution, generating low-resolution feature maps. We employed an embedded attention mechanism to alleviate the limitations of pooling. Furthermore, the self-expansion mechanism is an extension derived from MobileNet(Andrew G. Howard et al.,2017), a lightweight network structure based on depth-wise separable convolutions. We embedded it into the encoder to achieve low power consumption and enhanced speed.

Pix2Pix, developed in 2014, emerged from the Conditional Generative Adversarial Network (CGAN)(Mehdi Mirza & Simon Osindero. 2014.), specifically designed for image translation tasks. CGAN extends the basic GAN, enabling the generation of images that satisfy specific conditions or features, facilitating the transformation from patterns to images. Despite being an older development, Pix2Pix remains powerful in its image-to-image tasks and even exhibits faster output generation compared to the current prevalent Diffusion models. For large-scale estimations of carbon emissions based on building exteriors, the focus is on approximate calculations, eliminating the need for excessive granularity in image quality. The emphasis lies in the immediate availability of information through mobile devices for real-time estimation of building carbon footprints, prioritizing rapid results. Hence, we chose Pix2Pix as the baseline model for development.

In addition to integrating the generator and decoder based on these foundational principles, this model architecture also incorporates the concept of U-Net (Ronneberger, O. et al.,2015).This involves establishing skip connections between multiple levels in the encoder and decoder to address information loss and resolution reduction issues in semantic segmentation tasks. Conventional encoder-decoder structures, due to repetitive downsampling and upsampling operations, tend to reduce resolution gradually, potentially leading to the loss of fine features. Skip connections allow rich semantic information transfer from low-level to high-level feature maps into the decoder, preserving more details while maintaining high resolution. By integrating CGAN's ability to conditionally generate meaningful feature maps, combining PatchGAN's (Phillip Isola et al.,2017) detailed discriminator, and incorporating U-Net's method for retaining fine features at high resolutions, Pix2Pix demonstrates flexibility and robust performance in various image-to-image translation tasks. These include both image-to-image and image-to-label transformation tasks.

Our network model is based on the Pix2Pix generative adversarial architecture, enhancing feature quality through improvements in the generator. Firstly, the model utilizes a series of linear transformations to generate additional feature maps with lower computational costs. This segment incorporates a self-augmentation mechanism. Subsequently, it captures the long-term dependencies of feature maps through self-attention mechanisms in both channel and spatial domains. Our model consists of a seven-layer Convolutional Neural Network (CNN)(Figure 4). As data passes through each layer of the CNN, the model duplicates the same number of feature maps and incorporates them into the self-attention mechanism, where the generated feature maps, processed through softmax, are added back to the original data.



Layer	e1	e2	e3	e4	e5	e6	e7	d1	d2	d3	d4	d5	d6	d7
Feature Maps	64x64	32x32	16x16	8x8	4x4	2x2	2x2	2x2	2x2	4x4	8x8	16x16	32x32	64x64
Filters	64	128	256	512	512	512	512	512	512	512	512	256	126	64

Figure 4. This diagram shows the size of feature maps and the number of filters in each layer of encoder and decoder.

3. Experiment

In the training process of machine learning, we continually adjusted several hyperparameters, including epoch, batch size, and learning rate. To achieve better computational efficiency, convergence, and prevent overfitting, we settled on setting 50 epochs for the model to train on different parts of the dataset each time. Additionally, we fixed the Batch Size at 75, meaning that during each training step, the model randomly selects 75 images from the training data as a batch for gradient computation and weight updates. The gradients of the model's weights are calculated based on the loss function of these 75 images, and the model's weights are updated accordingly. Throughout our training process, both the generator and discriminator had a learning rate of 0.0002. We utilized the Adam optimizer and set the loss weight for the generator and discriminator to 0.5. For hardware specifications, we employed a TESLA T4 16G GPU and conducted the training using Keras.

About the evaluation metric, we utilized PSNR (Peak Signal-to-Noise Ratio) in our part segmentation experiments. A higher PSNR value indicates greater similarity between images, suggesting lower distortion. PSNR measures the ratio between the maximum possible power of a signal and the power of noise that affects its accuracy. In the context of images, PSNR offers a relatively objective and quantifiable means of assessing image distortion. This allows for a more comprehensive evaluation of the quality of our segmentation results.

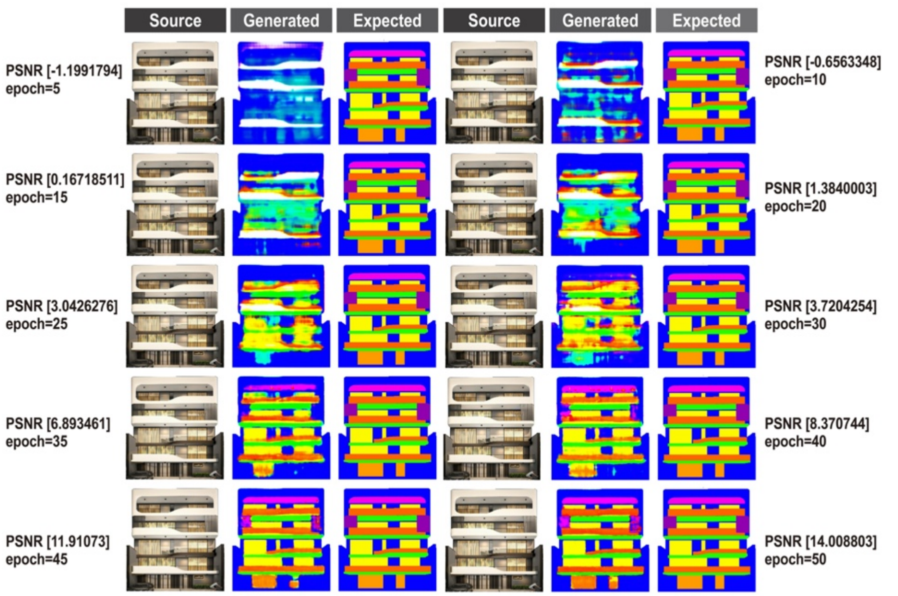


Figure 5. Segmented images selected at different training epochs in the recognition network.

We partitioned the data into training and testing sets with a ratio of 9:1. We evaluated the impact of different epoch counts on model improvement. As shown in (Figure 5), the results of part segmentation become increasingly similar to the expected values with more training epochs. Our model exhibits better PSNR values at 50 epochs, indicating its significant capability in the task of part segmentation.

In addition to training the machine using datasets, we conducted real-time tests on ongoing construction projects. After capturing the exterior of the buildings, we input the images into our self-developed convolutional neural network model. The automatically segmented results are illustrated in (Figure 6). We verified the results using a pixel-based comparison method, where the total sum of similar-colored pixels represents the area of the component. In this case, the pixel comparison error rates for the main building's exterior walls and metal curtain walls were below 10% (Table 1). For windows and railings, the pixel comparison error rates were below 30% (Table 1). This suggests that when calculating areas based on pixel counts, the segmentation results closely approximate the actual component areas.

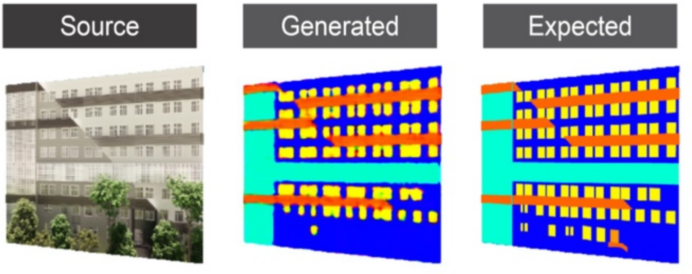


Figure 6. Mobile photography of an ongoing construction site and its generation result of part segmentation.

Category / Component	RGB	**Generated** / unit : Pixels	**Expected** / unit : Pixels	**Error Rate**
Glass Window	(255,253,0)	7946	11414	**32.9%**
Railing	(255,100,0)	6164	8499	**27.47%**
Curtain Wall	(0,255,215)	11194	12284	**8.87%**
Building ExteriorWall	(6,0,256)	22688	23780	**4.59%**

Table 1. We conducted a pixel-based analysis by comparing the segmented results from *Figure 6.* with the expected values to assess the error rates in pixel estimation. Pixel-based calculations provide estimations of areas, allowing us to compare them with actual area measurements.

4. Result & Discussion

4.1. PIXEL COUNTING VS. PSNR.

The task of our segmentation model is to accurately identify and delineate architectural components. The best performance of our component segmentation model in terms of PSNR values ranges between an average of 18-20. While not exceptionally high, the precision in comparison between detected RGB pixel values and actual color block areas is high. Consequently, from such experimental results, we observe that the precise shape of the color block segmentation isn't as crucial. What matters more to us is the area covered by these color blocks. If the overall area approximation tends towards reality, the area calculation becomes more realistic. Hence, in pursuit of automatically deriving areas through segmentation, the error rate in RGB pixel estimation holds higher importance and reference value compared to the distortion rate evaluated through PSNR between images.

4.2. DATASET REVIEW

During our research, we identified lower recognition rates for certain architectural components, such as the eaves. This is primarily due to the limited representation of eaves data in our training dataset, resulting in what is known as an 'Imbalanced Dataset.' This refers to a significant disparity in the number of samples across different categories in the dataset, impacting the model's performance in predicting and classifying minority categories. Additionally, in semantic understanding, the definition of eaves might be multifaceted. Therefore, in labeling, we might need further delineation regarding forms and functions, such as styled eaves, shading eaves, and traditional sloping roof eaves, among others.

4.3. COMPARISON OF DATASETS

The Façade dataset from UC Berkeley consists entirely of Western traditional architectural facade photos, all limited to buildings with seven floors or less. These images exhibit a higher degree of order in façade design, making it relatively easier for the model to discern various components. In contrast, our dataset emphasizes the collection of photographs from modern architectural settings, where there's more design variation and a combination of diverse materials in components. This diversity presents considerable ambiguity in defining components, posing a significant challenge for model training. However, it is precisely for this reason that we need to gather a vast amount of data from modern architectural contexts to meet the current demand.

5. Conclusion & Future Work

Overall, our research provides a rapid and convenient way to perform preliminary assessments of carbon emissions in physical constructions: (1). It offers an initial estimation of the total carbon emissions from façade components made of composite materials, allowing a deeper understanding of an object's environmental impact. (2). It enables architects designing architectural exteriors to improve and optimize designs with a lower carbon footprint, encouraging sustainable practices. (3). It provides a highly convenient method where, through the camera and measurement functions on a mobile phone, one can upload data to a cloud server for computation and then transfer results back to the phone, enabling assessments of component material carbon emissions, addressing the shortage of experts in carbon footprint assessment.

The focus of our paper lies in our precise calculation of the area of architectural exterior components to obtain the material quantity necessary for carbon footprint calculations. Future research aims not only to enhance the segmentation model's accuracy but also to develop a complete system focusing on expert systems and user interfaces, facilitating practical industrial applications.

REFERENCES

Andrea Zani1, Oluwateniola Ladipo, Antonio D'Aquilio, Carmelo Guido Galante, Matthew Tee and Tania Cortés Vargas.(2021). Facade Design Process to Establish and Achieve Net Zero Carbon Building Targets. *8th International Building Physics Conference (IBPC 2021) 25-27 August 2021, Copenhagen, Denmark.*

Zahra S. Zomorodian, Mohammad Tahsildoost. (2018). Energy and carbon analysis of double skin façades in the hot and dry climate. *Journal of Cleaner Production. Volume 197, Pages 85-96.*

Sabrina Barbosa, Kenneth Ip (2014). Perspectives of double skin façades for naturally ventilated buildings:A review. *Renewable and Sustainable Energy ReviewsVolume 40, December 2014, Pages 1019-1029.*

G. Baldinelli,(2009). Double skin façades for warm climate regions: Analysis of a solution with an integrated movable shading system. *Building and Environment. Volume 44, Issue 6, Pages 1107-1118*

Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio. (2014). Generative Adversarial Nets.*Neural Information Processing Systems (NeurIPS).*

Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, Alexei A. Efros. (2017).Image-To-Image Translation With Conditional Adversarial Networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Deng, Jia, et al. (2009). Imagenet: A large-scale hierarchical image database. *IEEE conference on computer vision and pattern recognition*.

Angel X.Chang, Thomas Funkhouser, Leonidas Guibas, Pat Hanrahan, Qixing Huang, Zimo Li, Silvio Savarese, Manolis Savva, Shuran Song, Hao Su, Jianxiong Xiao, Li Yi, and Fisher Yu.(2015). ShapeNet: AnInformation-Rich3DModelRepository.

Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár & C. Lawrence Zitnick(2014). Microsoft COCO: Common Objects in Context. *European Conference on Computer Vision(ECCV)*, *Pages740-755.*

Bolei Zhou, Hang Zhao, Xavier Puig, Sanja Fidler, Adela Barriuso, Antonio Torralba. (2017).Scene Parsing Through ADE20K Dataset. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).*

Pratheba Selvaraju, Mohamed Nabail, Marios Loizou, Maria Maslioukova, Melinos Averkiou, Andreas Andreou, Siddhartha Chaudhuri, Evangelos Kalogerakis.(2021) BuildingNet: Learning to Label 3D Buildings. P*roceedings of the IEEE/CVF International Conference on Computer Vision (ICCV).*

Yuan-Fu Yang, Min Sun. (2021). A Novel Deep Learning Architecture for Global Defect Classification: Self-Proliferating Neural Network (SPNet).*32nd Annual SEMI Advanced Semiconductor Manufacturing Conference (ASMC).*

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, Illia Polosukhin. (2017). Attention Is All You Need. *arXiv:1706.03762v7*

Andrew G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, Hartwig Adam. (2017).MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. *Conference on Computer Vision and Pattern Recognition (CVPR). arXiv.1704.04861.*

Mehdi Mirza, Simon Osindero. (2014).Conditional Generative Adversarial Nets. *Conference on Neural Information Processing Systems(NIPS).arXiv:1411.1784.*

Ronneberger, O., Fischer, P., & Brox, T. "U-net. (2015).Convolutional networks for biomedical image segmentation. *International Conference on Medical Image Computing and Computer-Assisted Intervention(MICCAI)pp. 234-241, Springer.*

Building Information Modelling

BATCHPLAN: A LARGE SCALE SOLUTION FOR FLOOR PLAN EXTRACTION

BURAK YILDIZ[1], JAVIER CUARTERO[2], FATEMAH MOSTAFAVI[3] and SEYRAN KHADEMI[4]
[1,3,4]*Technical University of Delft.*
[2]*Kaan Architecten.*
[1]*b.yildiz@tudelft.nl , 0000-0001-9932-4221*
[2]*j.Cuartero@kaanarchitecten.com, 0009-0007-5981-4112*
[3]*f.mostafavi@tudelft.nl, 0000-0002-8047-2168*
[4]*s.khademi@tudelft.nl, 0000-0003-4623-3689*

Abstract. The development of Building Information Modelling (BIM) has enabled new opportunities, such as standard data storage and collaborative building design. Moreover, there exist many Life Cycle Assessment (LCA) tools and Building Energy Performance (BEP) simulators that use the Industry Foundation Classes (IFC) exports of BIM platforms as input for further operational analysis. While the extracted IFC files contain numerical and tabular data from the BIM model, the visual data including floor plans and section drawings is often obtained directly from the original 3D software such as REVIT. In this study, we introduce an open-source solution, BatchPlan, for batch processing IFC files of medium- and high-rise building projects, leading to floor plan extraction on a large scale. Furthermore, we have designed a user-friendly graphical interface that allows users to select floors manually. BatchPlan is based on open-source Python packages; thus users can easily edit and adapt it to their specific requirements. The presented solution enables a scalable data generation pipeline for downstream tasks that require extensive quantitative analysis, such as machine learning models to perform material detection, volume estimation, and environmental impact prediction.

Keywords. Floor plan extraction, Industry Foundation Classes (IFC), Building Information Modelling (BIM), Architectural Technical Drawings, Big Data

1. Introduction

The emergence of Building Information Modelling (BIM) has improved building design efficiency by managing all the project data. To improve sustainable assessments of buildings in the design and evaluation phases, BIM tools are becoming more popular in the construction industry (Santos et al., 2019) and (Yeung et al., 2023). The Industry Foundation Classes (IFC) format file is an open file format used by BIM programs, including spatial elements, materials, and shapes of a building or facility (ISO 16739-1, 2018). The majority of BIM modelling software, including Revit, ArchiCAD, and Rebro, currently supports the import and export of IFC files (Jiang et al., 2019). IFC files are intended to be platform-independent, and therefore play an information exchanger role. This would allow interoperability between various BIM programs (Liu et al., 2021), as well as between BIM and other tools such as LCA (Xu et al., 2022).

To adapt to more diverse applications, IFC files can be converted to various file types such as Green Building XML (gbXML) (Elagiry et al., 2020). Moreover, since IFC files contain machine-readable unique identifiers alongside semantics (e.g., object type or function), characteristics or attributes (e.g., materials, color, and thermal properties), relationships (e.g., locations, connections, and ownership), and objects (e.g., columns and slabs), there is an opportunity to make use of them in automated procedures (Lai et al., 2019). In addition, given the difficulties stemming from multi-resource heterogeneous BIM software, IFC files have been used in automated building performance calculations (Deng & Lu, 2023).

Despite all the possibilities, not all BIM software is free to use or open source. In an open-source software, the source code is accessible, and licensing allows to modify and extend the software's functionalities. As a result of such transparency, users can improve and customize the software according to their own needs. Accordingly, this study introduces BatchPlan, a large-scale solution for floor plan extraction of buildings' IFC files. The proposed solution is developed to address the challenge of analysing big data during a data-driven design process, utilizing the conventional method of data storage in the building industry in an accelerated manner. The main contributions of this study are therefore as follows:

- Introducing an open-source, customizable, and flexible solution for floor plan extraction in large scale
- Enabling visual (floor plan maps), geometrical (room- and unit-wise coordinates), and numerical (area, volume, etc.) outputs data format
- Batch-processing large IFC files
- Facilitating large-scale quantitative analysis and ML-based workflows

The remainder of the papers is divided into three main subjects: 1) discussing related work regarding the floor plan extraction and the available open-source solutions using IFC files, 2) investigating the development of BatchPlan and the underlying dataset, 3) presenting the features of BatchPlan in comparison to other solutions for the task of floor plan extraction.

2. Related Work

2.1. FLOOR PLAN EXTRACTION

The plan, elevation, and section views of a construction structural drawing correspond to the different engineering data including dimensions, construction entities, and annotations, respectively (Björk & Laakso, 2010). Building plans represent a building on a horizontal plane at a particular floor level and convey information about architectural objects' distribution, extension, and outline (Yin et al., 2020).

The idea of extracting indoor spatial information from floor plans has gained a lot of attention in previous studies (Kim et al., 2021). The extraction of technical architectural drawings on the other hand is still often done manually. BIM software offers a user-friendly graphical interface for the manual extraction of floor plans at the scale of a single building design. However, dealing with a substantial volume of files simultaneously is not practical using this software. Also, manually extracting 2D architectural plans for all stories in large-scale projects can be a cumbersome process and is prone to errors. The mentioned limitations shed light on the advantages of automated extraction of floor plans, more preferably on a large scale.

2.2. IFC-BASED OPEN-SOURCE SOLUTIONS

There exist several solutions that take the IFC files as input and give the user the possibility to view and edit BIM files. These solutions range from toolkits to programming packages and libraries. Several available solutions are explained in this section.

IfcOpenShell is an open-source toolkit that helps in dealing with IFC files (IfcOpenShell, 2023). Viewing models, extracting attributes, moving objects, and generating 2D drawings are among the toolkit's possible functionalities. More specifically, IfcConvert is a command-line application of IfcOpenShell for converting IFC geometry into file formats such as OBJ, DAE, GLB, STP, IGS, XML, SVG, H5, and IFC itself. IfcOpenShell has been implemented for IFC-based BIM reconstruction in previous studies (Pan et al., 2023).

Xbim is an open-source BIM toolkit that supports the IFC file format (xbim Toolkit, 2023). Building Information Models (BIMs) can be read, created and viewed using Xbim. In addition to geometric and topological operations, there is also support for visualization of IFC files. Xbim Utilities, an example application of the toolkit allows performing bulk functions on IFC files. Another application, Xbim Exchange, makes the conversion of IFC files to COBie format possible.

Other than toolkits with multiple applications, there are few related solutions offering more limited functionalities. IFC.js is an open-source library that allows for viewing IFC models in web browsers (Viegas, 2023). Through its role as a bridge between different software, including commercial options, it demonstrates how open-source solutions can improve interoperability. In addition, BIMvision is a free IFC model viewer (Datacomp IT Sp. z o.o., 2023), allowing for viewing a virtual model from Revit, Archicad, or other CAD systems. Cutting the plane of the selected model face, edge, or the global axis is among the possibilities that BIMvision offers.

Most of the tools and solutions have been developed mainly with the purpose of either viewing or converting IFC files, whereas BatchPlan is developed to not only process big IFC files, but also provide other visual, geometrical, and numerical outputs which will be used in many design-related downstream tasks. This is an urgency to be addressed given the fast-growing development of data-driven methods in architectural design discipline and industry. Hence, the creation of workflows and pipelines for generating extensive and machine-readable datasets is crucial. Example measures are taken in the field for the shift from conventional databases (Pizarro et al., 2022) to ML-ready datasets (Van Engelenburg et al., 2023). Implementing BatchPlan, a large-scale solution for floor plan extraction of buildings, not only helps accelerate the process of outputting visual and numerical data from IFC files but also contributes to developing proper datasets for large quantitative analysis and ML-based tasks.

3. BatchPlan

We introduce BatchPlan as a robust large-scale floor plan extraction tool designed to be highly customizable, extensible, and pluggable in various capacities. The design decisions are meticulously crafted, particularly for the processing of extensive BIM data stored in IFC files. Customization is facilitated through the provision of user-defined element filters and styling. Filter functions serve the purpose of selectively excluding elements from the final outputs, allowing users to specify the elements they wish to exclude. Styling functions, on the other hand, are employed to define the appearance of elements, encompassing attributes such as colour coding in the generated outputs. Both filter and styling functions operate on an IFC element and its corresponding OpenCASCADE Boundary Representation (BRep) geometry. The filter functions return a Boolean value, while the styling functions provide the desired element style. Users have the flexibility to enhance BatchPlan by supplying a user-defined output formatter to tailor the output to their preferred format. An example output of this formatter can be a list of geometrical details of each room in WKT (well-known text) format. Furthermore, BatchPlan seamlessly integrates into other programs or data pipelines through its Application Programming Interface (API). The architectural framework of BatchPlan along with the data flow within it are shown in Figure 1.

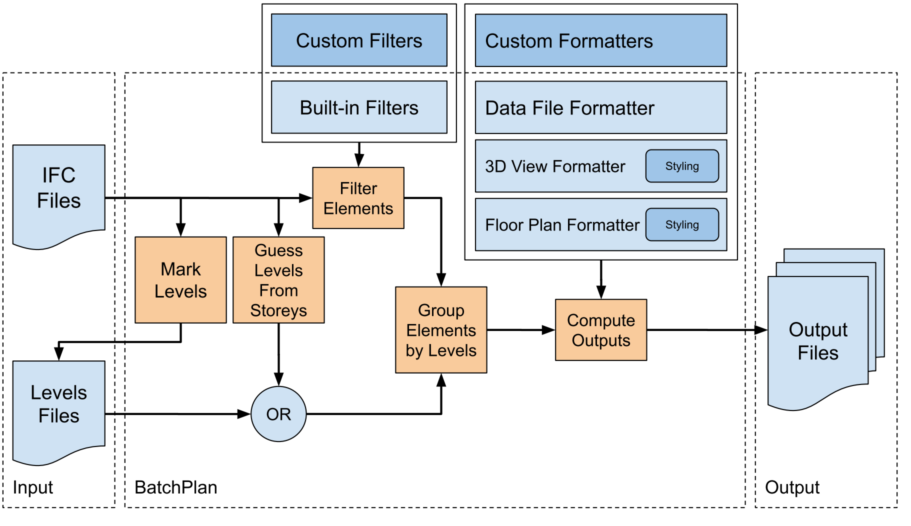


Figure 1. BatchPlan architecture and data flow of one single project from input to output. All orange boxes indicate main computing points where only level marking can be done using BatchPlan's Graphical User Interface (GUI) and others using Command Line Interface (CLI). Dark blue areas are

where user-defined codes can be injected.

BatchPlan accepts one mandatory and one optional input. The mandatory input comprises IFC files, while the optional input involves level files. These level files are formatted as comma-separated vector (CSV) files and encompass heights at which BatchPlan computes floor plans. Users can generate these files conveniently through BatchPlan's Graphical User Interface (GUI). Alternatively, one may manually create a CSV file or employ a script to populate it with the desired heights. This deliberate design choice offers users the flexibility to extract floor plans at various heights on a scalable basis. The level files are optional since BatchPlan can use storey information to automatically calculate the heights. The data flow within BatchPlan, illustrated in Figure 1, unfolds as follows:

- IFC files given in command line arguments are read.
- IFC "Cutting" planes on which floor plans are generated are determined by employing heights from either the corresponding level files or storey information.
- The filters, either provided or selected by user, are executed on the elements.
- The remaining elements, post-filtering, are categorized by levels. This categorization involves grouping elements with a non-empty intersection with the same plane into a common bucket.
- The intended outputs are computed for each of the buckets, created at the previous step, using the selected and/or provided output formatters, subsequently saved to output files.

BatchPlan is implemented in the Python programming language, and its source code (Yildiz et al., 2023) is made available under the MIT License, known for its permissive nature as an open-source license. The selection of Python as the programming language aimed at facilitating ease of modification and contributions for future users. The code leverages two primary open-source packages, namely IfcOpenShell (IfcOpenShell, 2023) and pythonocc (pythonocc-core, 2023). IfcOpenShell serves the purpose of reading and processing IFC files, while pythonocc is utilized for identifying intersections between planes and elements. Additionally, built-in formatters are implemented using pythonocc.

3.1. DATASET

For the development and the test phase of the BatchPlan, we curated a dataset from KAAN Architecten (KAAN Architecten, 2023) projects. Six distinct residential projects were selected to encompass various design stages, ranging from early design to project construction, including building permit documentation and design development stages.

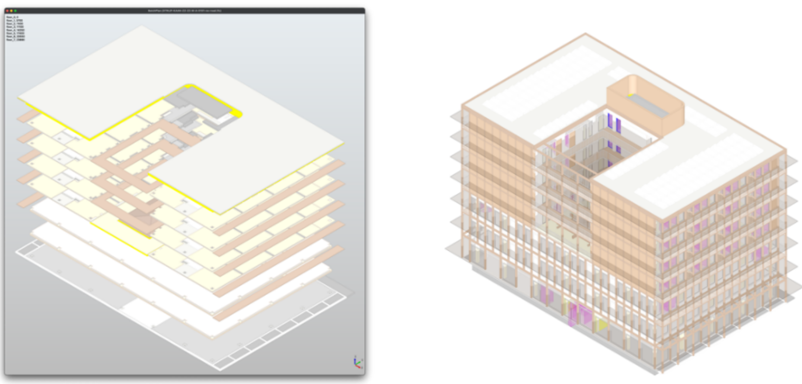
To ensure consistency and facilitate meaningful comparisons among dataset assets, a decision has been made to focus on both medium- and high-rise housing architectural examples. These projects have been chosen to address different scenarios in which the evolution of the geometric (detailed) aspects of architecture could potentially impact the results of BatchPlan or its processes. Details of the selected projects are provided in Table 1.

Table 1. Details of the selected projects from the KAAN Architecten dataset.

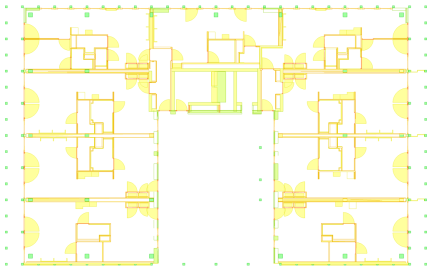
Project Name	Category	Design Start	Design End	Gross Floor Area (m2)	Number of Storeys	Number of Elements
SPOT	Construction	2017	2021	40,000	35	47,560
Lumiere	Design	2018	-	50,000	46	33,865
Zalmhaven	Built	2015	2018	49,000	24	27,900
Overhoeks	Construction	2017	2019	11,400	9	11,720
Blok O	Design	2022	-	150,000	9	10,270
Strijp S	Design	2018	2023	4,800	7	5,340

4. Results

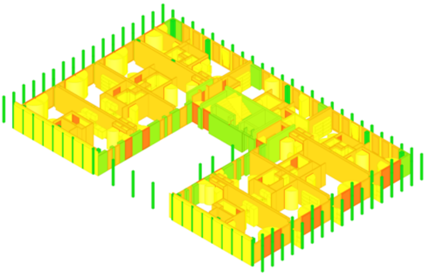
BatchPlan is designed to efficiently process large-scale data while maintaining ease of customization. To assess its capabilities, the KAAN Architecten buildings dataset has been employed. Sample outputs of the BatchPlan's GUI are presented in Figure 2, highlighting the features of the proposed solution, from marking the levels at which the floor plans are desired to be extracted, to customized 2D or 3D color-coding of the outputs depending on a certain criterion.



(a) (b)



(c)



(d)

Figure 2. A sample GUI view and sample outputs of BatchPlan. (a) BatchPlan GUI for marking the levels at which the floor plans are extracted, (b) 3D view of the whole building with planes indicating the levels where the floor plans are extracted, (c) Customized color-coding of a floor plan using an external data source (i.e., carbon footprint values). (d) Customized color-coding of the 3D view of the single floor plan in (c).

A comparative analysis of BatchPlan's features with those of selected software and tools capable of floor plan extraction was also conducted. The results of these comparisons are presented in Table 2. The MIT license has been chosen for BatchPlan, affording users the right to modify, customize, and redistribute it under both open source and proprietary licenses without hesitation. This approach is advantageous for both the open-source community and the industry, serving non-profit as well as for-profit purposes. One distinctive feature of BatchPlan is its customization capability, allowing users to provide their own styling functions to alter the styling of floor plans or 3D output views. These styling functions can also incorporate external data sources to manipulate the appearances of the elements. For example, users can create a heatmap by coloring the elements according to their respective environmental footprint.

Table 2. Comparison of the software and tools related to floor plan extraction

Program	IfcConvert	BlenderBIM	Revit	BatchPlan
License	LGPL-3.0-or-later	GPL-3.0-or-later	Proprietary	MIT
Source Code	Open source	Open source	Closed source	Open source
User Interface	CLI	GUI	GUI	CLI
Customizable	Not easy	Not easy	Limited	Easy
Extensible	No	Not easy	Not easy	Easy
Scalable	Yes	No	No	Yes
Platform-Independent	Yes	Yes	No	Yes

Another feature of BatchPlan is its extensibility, signifying that in addition to built-in output formatters such as rendered images and 3D views, users can introduce customized output formatters. These output formatters are functions that accept a set of IFC elements and their corresponding OpenCASCADE BRep representations, generating customized outputs such as data files and rendered images. The data files may encompass associated properties of the elements, such as their volume and locations. Another example of a customized formatter could be one that produces rendered images with bounding boxes around elements.

Sample floor plans have been extracted using the software and tools listed in Table 2, as illustrated in Figure 3. IfcConvert produces floor plans in the form of vectorized images, particularly in the Scalable Vector Graphics (SVG) format. Achieving a satisfactory floor plan in IfcConvert necessitates setting specific parameters and modifying the output SVG file. BlenderBIM, equipped with a user-friendly GUI, requires manual effort for extracting floor plans for each building and its individual levels. Revit, being a proprietary software, also entails manual intervention for floor plan extraction. In contrast, BatchPlan offers the advantage of extracting floor plans for an entire set of projects with a single command, streamlining the process significantly.

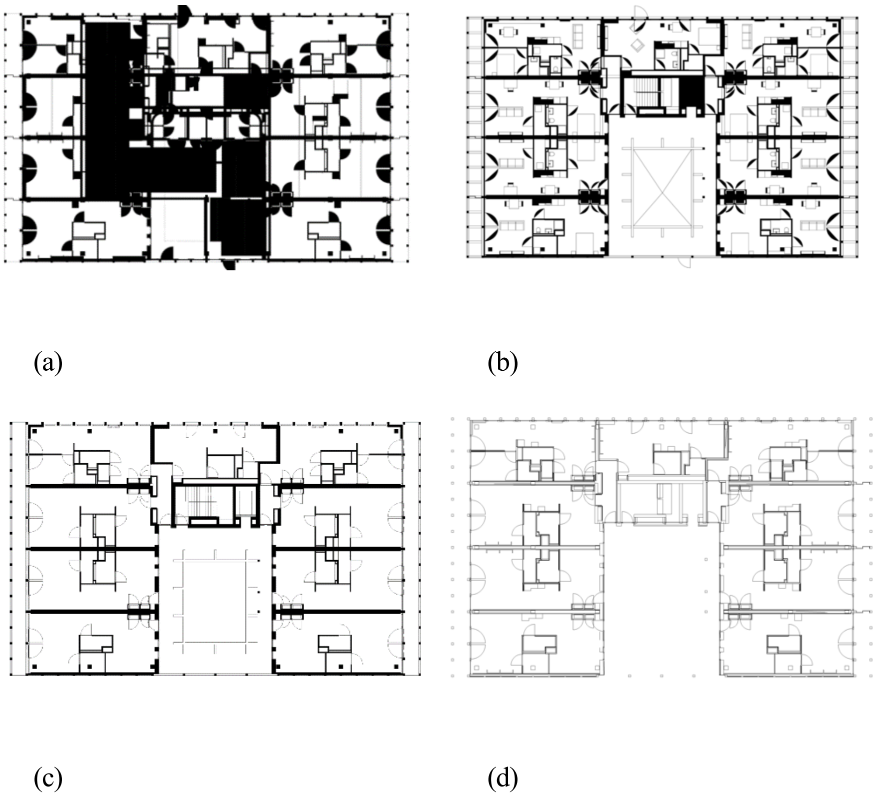


Figure 3. Comparison of the default floor plans of a sample building extracted using IfcConvert (a), BlenderBIM (b), Revit (c), and BatchPlan (d).

5. Findings and Discussion

Floor plans serve as the orthogonal projection of a building level, depicting building elements such as walls, doors, and windows to convey the architect's design. With the development of data-driven design methods, the need for processing big architectural data arises. In this study, an open-source, customizable, and extensible solution for floor plan extraction, called BatchPlan, is introduced. The development of the proposed solution was presented, and the results were compared to the similar available solutions. The presented solution can serve the research on big data resulting from large amount of existing IFC files. Moreover, to test the features of the BatchPlan, the solution was applied to six medium- and high-rise housing projects.

The development of BatchPlan facilitates flexibility, while concurrently prioritizing the efficient processing of large-scale data. Since BatchPlan offers processing and outputting large amounts of data, it serves further downstream tasks in which big data is needed. For instance, a possible application is developing a diverse collection of floor plans to create a dataset, enabling the training of an AI-driven model. The outputs of this study will not only assist designers in generating floor plans at the desired building levels but also enable researchers to perform quantitative analyses on district or urban scales.

Besides the mentioned features, BatchPlan has some limitations, mostly depending on the familiarity of the users with various interfaces. Although BatchPlan has a GUI for level marking, it is executed through its Command Line Interface (CLI). While the CLI offers advantages, particularly when running the program for batch processing on High-Performance Computing (HPC) clusters, some users may find it less convenient. Further developments of the proposed solution include enhancing the GUI tailored towards the architectural design process and allowing other types of output formats.

Acknowledgements

We thank our collaborators from the KAAN Architecten firm who provided the required data of building projects that assisted the research. This study was funded by NWO (Dutch Research Council) under the file number of GOCI.KIEM.02.017.

References

Björk, B. C., & Laakso, M. (2010). CAD standardisation in the construction industry — A process view. Automation in Construction, 19(4), 398–406. https://doi.org/10.1016/J.AUTCON.2009.11.010

Datacomp IT Sp. z o.o. (2023). BIMvision. https://bimvision.eu/

Deng, X., & Lu, K. (2023). Multi-level assessment for embodied carbon of buildings using multi-source industry foundation classes. Journal of Building Engineering, 72. https://doi.org/10.1016/j.jobe.2023.106705

Elagiry, M., Bourreau, P., Charbel, N., & De Angelis, E. (2020). IFC to Building Energy Performance Simulation: A systematic review of the main adopted tools and approaches. https://www.researchgate.net/publication/345983333

IfcOpenShell. (2023). The open source IFC toolkit and geometry engine. https://ifcopenshell.org/

ISO 16739-1. (2018). Industry Foundation Classes (IFC) for data sharing in the construction and facility management industries.

Jiang, S., Jiang, L., Han, Y., Wu, Z., & Wang, N. (2019). OpenBIM: An Enabling Solution for Information Interoperability. Applied Sciences 2019, Vol. 9, Page 5358, 9(24), 5358. https://doi.org/10.3390/APP9245358

KAAN Architecten. (2023). KAAN Architecten - Netherlands based architectural firm. https://kaanarchitecten.com/

Kim, H., Kim, S., & Yu, K. (2021). Automatic Extraction of Indoor Spatial Information from Floor Plan Image: A Patch-Based Deep Learning Methodology Application on Large-Scale Complex Buildings. ISPRS International Journal of Geo-Information, 10(12). https://doi.org/10.3390/ijgi10120828

Lai, H., Deng, X., & Chang, T.-Y. P. (2019). BIM-Based Platform for Collaborative Building Design and Project Management. Journal of Computing in Civil Engineering, 33(3), 05019001. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000830/ASSET/A68C72D6-F281-4755-8C15-EA811AA6B442/ASSETS/IMAGES/LARGE/FIGURE11.JPG

Liu, L., Li, B., Zlatanova, S., & van Oosterom, P. (2021). Indoor navigation supported by the Industry Foundation Classes (IFC): A survey. Automation in Construction, 121, 103436. https://doi.org/10.1016/J.AUTCON.2020.103436

Pan, Z., Yu, Y., Xiao, F., & Zhang, J. (2023). Recovering building information model from 2D drawings for mechanical, electrical and plumbing systems of ageing buildings. Automation in Construction, 152, 104914. https://doi.org/10.1016/J.AUTCON.2023.104914

Pizarro, P. N., Hitschfeld, N., Sipiran, I., & Saavedra, J. M. (2022). Automatic floor plan analysis and recognition. Automation in Construction, 140(April), 104348. https://doi.org/10.1016/j.autcon.2022.104348

pythonocc-core. (2023). GitHub - tpaviot/pythonocc-core: Python package for 3D CAD/BIM/PLM/CAM. https://github.com/tpaviot/pythonocc-core

Santos, R., Costa, A. A., Silvestre, J. D., & Pyl, L. (2019). Integration of LCA and LCC analysis within a BIM-based environment. Automation in Construction, 103, 127–149. https://doi.org/10.1016/J.AUTCON.2019.02.011

Van Engelenburg, C., Khademi, S., Mostafavi, F., Standfest, M., & Franzen, M. (2023). Modified Swiss Dwellings: A ML-ready Floor Plan Dataset of Residential Building Complexes. https://doi.org/10.4121/e1d89cb5-6872-48fc-be63-aadd687ee6f9.v1

Viegas, A. G. (2023). IFC.js Open source IFC library. https://github.com/IFCjs

xbim Toolkit. (2023). making building information flow. https://docs.xbim.net/

Xu, J., Teng, Y., Pan, W., & Zhang, Y. (2022). BIM-integrated LCA to automate embodied carbon assessment of prefabricated buildings. Journal of Cleaner Production, 374, 133894. https://doi.org/10.1016/J.JCLEPRO.2022.133894

Yeung, J., J Hahn Menacho, A., Marvuglia, A., Navarrete Gutiérrez, T., Beach, T., & Rezgui, Y. (2023). An open building information modelling based co-simulation architecture to model building energy and environmental life cycle assessment: A case study on two buildings in the United Kingdom and Luxembourg. Renewable and Sustainable Energy Reviews, 183, 113419. https://doi.org/10.1016/J.RSER.2023.113419

Yildiz, B., Mostafavi, F., Cuarteo, J., & Khademi, S. (2023). Carbon Image Project. https://github.com/byildiz/BatchPlan

Yin, M., Tang, L., Zhou, T., Wen, Y., Xu, R., & Deng, W. (2020). Automatic layer classification method-based elevation recognition in architectural drawings for reconstruction of 3D BIM models. Automation in Construction, 113. https://doi.org/10.1016/j.autcon.2020.103082

BIM-ENABLED REGULATORY DESIGN RULE CHECKING FOR BUILDING CIRCULATION

CHERYL LEE[1], KHYSTELLE YEO[2] and STYLIANOS DRITSAS[3]
[1,2,3]*Singapore University of Technology and Design.*
[1]*cheryl_lee@sutd.edu.sg, 0009-0001-0965-9368*
[2]*khystelle_yeo@sutd.edu.sg, 0009-0007-3478-0876*
[3]*dritsas@sutd.edu.sg, 0000-0002-9609-2784*

Abstract. Automated design rule checking (RDC) in Building Information Modelling (BIM) can be challenging especially when dealing with qualitative aspects and intricate regulations like building circulation. The study proposes a novel method for Regulatory Design Rule Checking (RDRC) for building circulation, addressing challenges in translating regulations to computational constructs and extracting relevant information from complex BIM models. Through a tiered compliance assessment, the investigation considers preventive rule-based checks for doors and corridors and explores constraint-based regulatory incentive schemes such as through-block pedestrian links (TBPL). The RDRC analysis identifies non-compliance and concludes with a recommendation for potential adjustments. This work aims to benefit designers and regulators, providing productivity enhancements and a deeper understanding of regulatory intricacies in the context of building circulation.

Keywords. Design Rule Checking, Building Information Modelling, Building circulation, Network analysis

1. Introduction

One of the statutory roles of architects, engineers and contractors is to ensure buildings are designed following the local regulations so that buildings are safe and serviceable for their occupants and environmentally sustainable (Dimyadi & Amor, 2013). However, the numerous criteria mandated by various government departments render regulatory compliance checking a time-intensive process traditionally manually carried out by professionals in the Architecture, Engineering and Construction (AEC) industry (Lee et al., 2014; Peng & Liu, 2023). Nevertheless, the use of building information modelling (BIM) offers a significant opportunity for automating design rule checking (DRC) (Eastman et al., 2018) so that designers and regulators can quickly identify non-compliances and make necessary adjustments, increasing the efficiency of compliance checking saving time and resources (Eastman et al., 2009; Choi et al., 2013).

While DRC is not a new concept, its adoption faces two primary obstacles: (a) converting regulations from natural language to computational constructs, and (b)

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 211-220.

correctly identifying and extracting relevant information from the BIM model. Quantitative rules assessing explicit BIM entity properties are straightforward but compliance analyses using implicit information, like circulation paths, are challenging. Circulatory networks, reflecting the routes traversed by people within buildings and urban spaces, is an implicit information derived from the remainder of the building components within a model. In addition, while quantitative aspects of circulation can be defined and measured, evaluating the qualitative aspects, such as ease of access, safety and sufficiency, can be subjective, multi-dimensional and lacking objective metrics (Lee et al., 2009). Hence, determining whether the circulation configuration for a building and its external spaces meets the authorities' requirements requires critical judgment and interpretation, making it a complex task to automate.

This work investigates regulatory design rule checking (RDRC) for building circulation. We address the challenge of identifying circulatory networks by analysing complex BIM databases and translating three-dimensional entities into spatial graphs. RDRC assessment is done using a tiered approach starting from (a) basic compliance testing, (b) deep-insight analysis to identify reasons for non-compliance, and (c) concluding with design recommendations. Our work extends beyond preventive rule-based compliance to explore regulatory incentive schemes which aim to enhance the built environment. We hope our work is useful for designers and regulators by improving productivity and deepening understanding of regulatory frameworks.

2. Relevant Work

The development of rule-based systems in models began in the 1980s (Garrett & Fenves, 1987), and with technological advancements, automated rule checking has progressed but primarily within the research domain (Dimyadi & Amor, 2013). We summarise the various approaches implemented to overcome key challenges in circulatory DRC: (a) rule language, (b) model checking and (c) circulation analysis.

DRC involves translating natural language rules to a computer-readable format. Domain-specific languages are used to extract and parse information from regulations into computer executable code, such as KBim. Another approach is natural language processing which uses artificial intelligence for automatic code generation. Despite these advancements, the ambiguity and complexity of regulations pose challenges for complete automatic conversion into computer code, necessitating human intervention for accurate rule interpretation and translation (Sydora & Stroulia, 2020).

Solibri Model Checker (SMC) is a commonly used DRC software that analyses Industry Foundation Class (IFC) models based on international standards with the flexibility to adjust values to suit local codes (Solibri Inc., 2023). Another example is the e-PlanCheck component in CORENET (Construction and Real Estate Network), a web-based integrated hub for the Singapore construction industry (CORENET, 2016). Both examples use object libraries with coded regulations to check compliance of the IFC models and generate reports, simplifying DRC (Eastman et al., 2009). However, e-PlanCheck has predefined rules that limit users and the report in SMC does not track compliant instances, making it difficult to review which instances have been processed.

Determining a consistent method to represent the paths around the building for circulation analysis is challenging due to varying human movement patterns. A graph-based approach is commonly used due to its simplicity and efficiency (Werner et al, 2000). Various methods have been introduced and the two main types of circulation graphs are (a) topological graphs which show simple room connections without indicating the actual path taken, and (b) distance-measured graphs, which define the circulation path through each space. Traditional methods measure the distance through the center of spaces, while others, like Kannala's evacuation-graph model (2005) and Lee's Universal Circulation Network (UCN) method (2009), propose shortest distance measurements and diagonals to emulate human circulation patterns. We integrate these concepts in our work to create a metric graph structure that provides a unified circulation path through all spaces, following a more natural human circulation pattern.

Various studies have demonstrated the use of circulatory graphs with embedded spatial and physical information for circulatory DRC. The Georgia Institute of Technology developed an SMC plug-in to assess circulation and security compliance for courthouse designs (Eastman et al., 2009; Lee, 2010). Lee et al. introduced Numeric Data of Building Circulation (NDBC) to quantitatively evaluate building circulation by comparing NDBC values of different design options (Lee et al., 2014).

3. Methodology

The process for building circulation RDRC follows a similar four-stage structure that C. Eastman et al. suggest for DRC: (1) rule interpretation; (2) building model preparation; (3) rule application and (4) reporting of results. Based on the four stages, we delve deeper into the specific challenges and considerations associated with each RDRC stage for circulation and mobility within and around buildings.

3.1. RULE INTERPRETATION

Interpreting and translating regulations on circulation and mobility into a computer-readable format poses challenges due to its multiple layers of conditions. To tackle this, a tiered approach is employed sorting the regulations into two groups: simple rules and complex rules. The simpler rules are mostly prescribed preventive quantitative rules and are relatively straightforward to interpret. On the other hand, complex rules are commonly qualitative constraint-based regulations and may require the consideration of multiple conditions before a check can be done. For instance, the minimum covered walkway width varies according to its location. Qualitative regulations are usually more complicated as they tend to lack clarity. For example, the regulation may state that a public space should have frontage onto a major street or pedestrian thoroughfare which is difficult to define. To deal with the complexity, the rules are broken down into parts to ensure the nuances and intricacies of the regulations are captured accurately and that the computerized version reflects the original intent and meaning.

3.2. BUILDING MODEL PREPARATION

Standardised BIM models are crucial to facilitate the extraction and utilisation of the relevant information for building circulation RDRC. Unfortunately, there is a lack of quality control as many do not adhere to the modelling conventions or best practices, (Migilinskasa et al. 2013; Wang et al. 2014; Chen and Luo, 2014). Some examples include (a) misuse of modelling object-model such as using a floor-type object to model the roof, (b) limitations of BIM object-model such as roads and urban features, (c) partial or incorrect labelling entities with non-standard categorical data such as door width can be stored as "width" or "rough width", (d) inconsistent use of geometric and semantic information such as elevators or parking lots can be modelled as floors, lines or hatches, and (e) modelling errors such as unenclosed spaces with gaps between walls or walls and floors. Consequently, identifying the appropriate elements for checking and preparing the model for data extraction becomes more problematic.

For this purpose, we developed an application to automatically extract the essential information from the model to generate the implicit circulation path for circulatory RDRC. The application reconstructs the 3D model into 2D polygons with assigned attributes to identify the elements (e.g., walls, doors or windows), their properties (e.g., opaque or porous, indoors or outdoors) and the type of space (e.g., shop, restaurant, toilet or stairs), see Figure 1. However, due to modelling inconsistencies, much manual polygon reconstruction is required. To compute the circulatory graph for analysis, the polygons are skeletonized using the Medial Axis Transform. This graph transverses through the middle of all walkable spaces and doors, see Figure 1, and is composed of nodes and edges with embedded physical data organised in Python dictionaries.

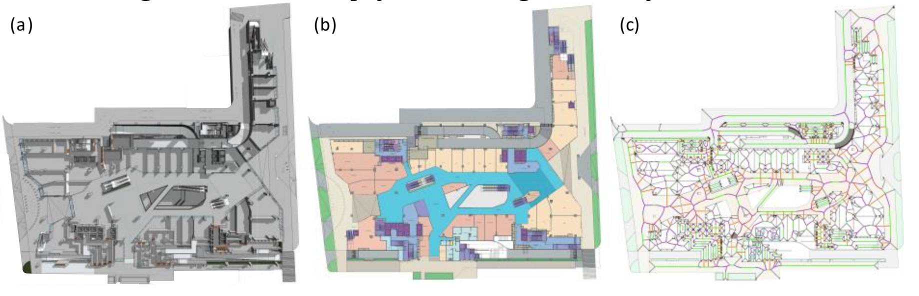


Figure 1. Extraction of circulation paths from BIM model: (a) Perspective plan view of 3D model. (b) Reconstructed 2D polygons. (c) Extracted circulatory graph through walkable spaces

3.3. RULE APPLICATION

Using the extracted spatial properties and attributes of the model's circulatory graph, the design is checked for compliance. The circulatory RDRC is formulated and implemented in a Python notebook using NetworkX functions to analyse the extracted circulation paths. Depending on the complexity of the rules, the execution process and required data for rule checking may vary. Typically, RDRC for simple quantitative rules, such as minimum or maximum dimensions for a particular element, can directly utilise the data in the dictionaries of the relevant nodes and edges. Advanced rules with multiple requirements will need to be broken down into parts for analysis and additional analytical steps to derive implicit properties from the model. For example, determining the shortest distance to an exit door would require identifying the nearest exit door and measuring the distance from that room to the exit door.

3.4. REPORTING OF RESULTS

This final step reports the results of the rule checking, identifying which elements have passed or failed the requirements of the rules. We use Matplotlib in Python to generate graphical reports which visualise the items that passed or failed the regulations at their location. Additionally, further classification to indicate the severity of the violations can be implemented, such as colour coding, numbering or other means. This allows designers to promptly identify the areas that require rectification to ensure compliance.

4. Results

Based on the proposed methodology, the application of RDRC for circulatory-related rules in Singapore has been tested on various typologies. The BIM model of a shopping mall in Singapore is used as a case study to exemplify the complexity, approach and any assumptions made.

4.1. BASIC COMPLIANCE TEST

The circulatory RDRC assessment starts with simple prescriptive quantitative regulations. An example applicable across most countries would be a minimum door or corridor width requirement. Under Singapore's regulations, the minimum clear opening of doors is 850mm and the minimum width of corridors is 1200mm. These design parameters are set to ensure that the functionality of the doors and corridors is met and adequately sized for the occupants to pass through.

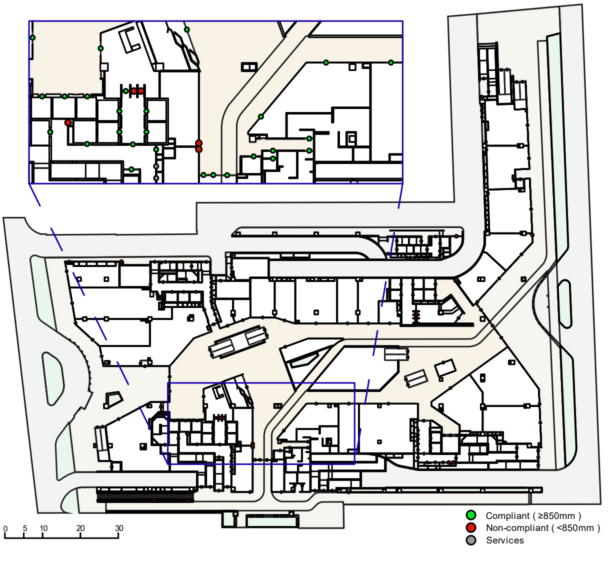


Figure 2. Checking for door compliance

Door width RDRC is a relatively straightforward task since the doors are specific object types with fixed dimensions. Using the extracted circulatory graph, all door nodes are filtered out and their width is checked if they meet the minimum width requirement. The results are visually reported in a plan to show which falls below or meets the minimum, see Figure 2, enabling designers to quickly identify areas that need rectification. Moreover, there may be additional conditions that affect the minimum door width requirement, such as the type and capacity of the room which the door leads to. For instance, cabinet or riser doors, although considered doors within the model, are exempt from following the minimum width requirements. Consequently, each door node has attributes which accurately identify the destination of the door to allow a more precise investigation, corresponding to its minimum door width requirements.

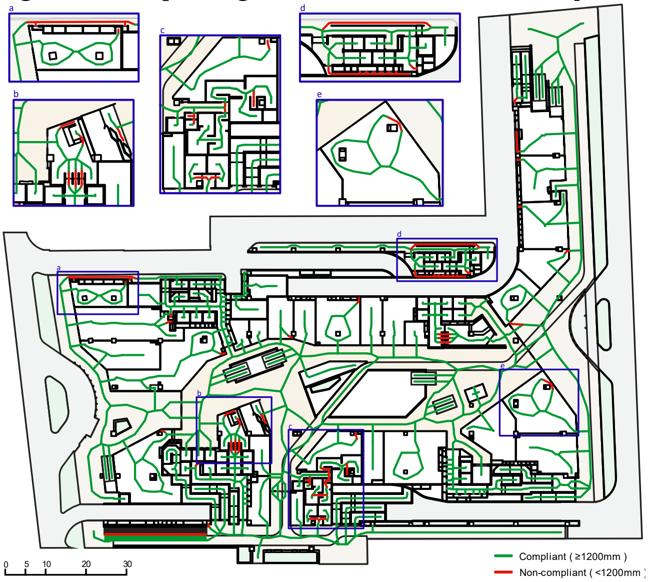


Figure 3. Checking for corridor compliance

On the other hand, corridor widths are not readily available and checking for compliance requires additional effort. Corridors are typically a straight passage formed between two parallel walls that provide access to various rooms. However, there may be instances where the space is irregularly shaped and the widths vary. Therefore, it is crucial to measure the widths of all circulation spaces to pick up potential areas that do not meet the regulatory requirements. Our circulatory graph representation derives the central line within all spaces and the measured dimensional information at the ends of each edge is stored in the node attributes, allowing us to evaluate the width of all the network paths. Similar to the assessment and reporting process for the door widths, corridors which violate the rules are filtered and highlighted, see Figure 3.

4.2. THROUGH-BLOCK PEDESTRIAN LINKS

Through-block pedestrian link (TBPL) is a regulatory incentive scheme by URA that encourages better urban circulatory design. TBPL is a pathway that cuts across a private development increasing ground-level permeability, creating an extension of the public space, increasing connectivity between public spaces and enhancing pedestrian convenience by shortening distances travelled.

Technically a TBPL is defined as “an internal covered walkway space between 4m to 7m wide that runs through a building, connects two parcels of public areas and is always kept open for public use”, (URA, 2020) as depicted in Figure 4. According to the code, all covered floor areas within a development are counted as gross floor area

(GFA) and are taxable unless exempted, meaning TBPL areas are also taxable. Hence, developers tend to avoid wide corridors or walkways in their projects since these circulation spaces are not commercially lucrative as they are non-rentable and non-sellable spaces. Therefore, to incentivise designers and developers to incorporate TBPL in their developments, the calculation of taxable GFA for compliant TBPL is excluded. The specified minimum width ensures sufficient space for pedestrians' comfort, preventing narrow corridors. Simultaneously, the maximum width ensures excessive corridor size and prevents exploitation of the incentive.

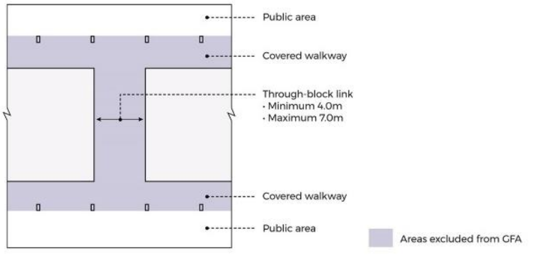


Figure 4. Explanatory concept diagram showing a plan view of Through-Block Pedestrian Link (TBPL) from Covered Walkway and Linkages, by Urban Redevelopment Authority, 2023.

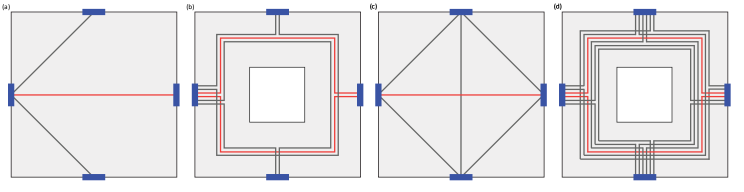


Figure 5. All routes from one door to all doors: (a) in a space without cycles vs (b) in a space with cycles. All routes between all doors: (c) in a space without cycles vs (d) in a space with cycles

Since the TBPL is a constraint-based regulation and is generally annotated in the submission plans by the architect, there is no certainty whether it is correctly done. For both architects and regulators, the verification of the validity of TBPL is non-trivial and should not be overlooked. Hence, the challenge from the perspective of DRC is (a) Determining what spaces may be a candidate for TBPL; this implies filtering out private from quasi-private spaces based on room usage. (b) Determining what source-destination pairs for the doors; differentiating the main and service doors. (c) Computing all possible routes between them; for which spaces without internal circulatory cycles, it is a quadratic complexity problem, O(n2), where 'n' represents the number of doors. However, when dealing with spaces containing cycles, the complexity quickly escalates to factorial, O(n!), refer to Figure 5.

Determining the potential public spaces for the TBPL can be inferred by either (a) using the architect's annotation or (b) selecting the spaces with the highest concentration of doorways. The latter heuristic is highly effective as lobbies and common building areas tend to have at least one order of magnitude higher number of doors compared to regular rooms, see Figure 6. To determine which doors are suitable to perform TBPL analysis we may either (a) use the architect's annotation as to which constitute main entrances or (b) infer those from space usage characteristics. To efficiently compute the TBPL we perform a route filtering by pruning the interior graph of the space to exclude all ineligible edges, as shown in Figure 7. This simplification reduces the computational complexity of finding all simple paths between source-destination pairs. Simple paths are paths in a graph that does not have repeated nodes and is suitable since the TBPL should not go in a loop. For instance, in the case study, the number of all simple paths without filtering is 2016 while after the filter is 20, computed in 6346 milliseconds and 27 milliseconds, respectively.

Figure 6. Comparison of number of doors per space

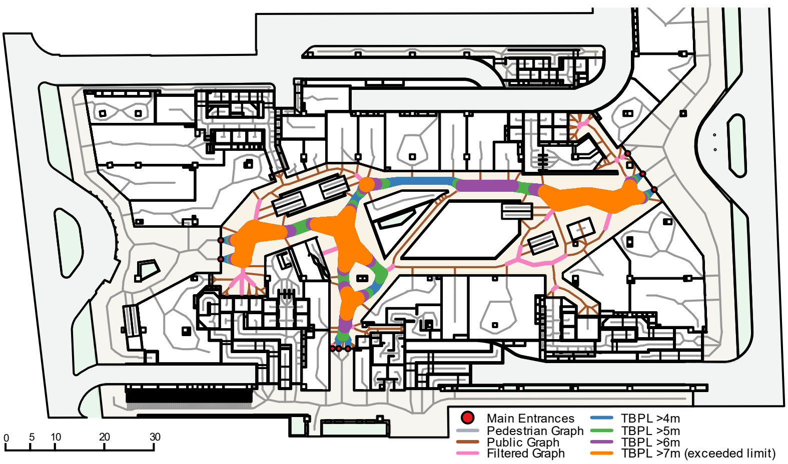


Figure 7. Filtered Graph that identifies all unique and valid TBPLs

Besides successfully accomplishing the primary goal of identifying all unique and valid TBPLs, as shown in Figure 7, our investigation uncovers additional observations related to code interpretation and design feedback. The following are our key findings from the analyses and insights of our research: (a) There may be more than one TBPL in a building which is interesting since it is not assumed in the code. (b) We can determine the straightness of a TBPL which is implied by the URA diagram, Figure 4, but not explicitly mandated. As seen in Figure 7, the identified TBPL is not straight but goes through a few bends and curves. (c) Our research provides valuable design feedback to architects, offering insights into paths that are potential TBPLs but are interrupted by local bottlenecks. (d) Finally, we can evaluate and compare the efficiency of the TBPL by computing the GFA gained or lost in the process. Through our research, we have provided a more measured approach to assess and evaluate the TBPL which offers insights to both regulators and designers.

5. Conclusion

In conclusion, we have presented a systematic approach to building circulation RDRC and demonstrated a plausible solution to resolve the issues faced. The study introduced a method for extracting circulation paths from BIM models and used the spatial structure of the graph to perform circulatory RDRC assessment for both conventional preventive rule-based compliance and regulatory incentive schemes. Achieving better BIM standardisation and quality control is crucial to fully automate the generation of the circulation network from BIM and reduce manual intervention. The extracted circulatory graph has the potential to be employed in various network analyses such as identifying the heavily traversed paths and assessing corridor capacities to prevent crowding.

Acknowledgements

This research is supported by the National Research Foundation, Prime Minister's Office, Singapore under its Cities of Tomorrow R&D Programme (COT-H1-2020-2). Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not reflect the views of National Research Foundation, Singapore and Ministry of Natural Development, Singapore.

References

BCA: All set for 2015: The BIM Roadmap. Build Smart Magazine, no. 9, p. 4 (2011)

Chen, L., & Luo, H. (2014). A BIM-based construction quality management model and its applications. *Automation in Construction*, 46, 64-73. https://doi.org/10.1016/j.autcon.2014.05.009

Choi, J., Choi, J., Cho, G., Kim, I. (2013). Development of Open BIM-based Code Checking Modules for the Regulations of the Fire and Evacuation. In *buildingSMART Korea 2013*. buildingSMART Korea.

Dimyadi, J., Amor, R. (2013). Automatic building code compliance checking–where is it at? *Proceedings of the 19th CIB World Building Congress*, pp. 172–1853. https://doi.org/10.13140/2.1.4920.4161 URL http://hdl.handle.net/2292/23574.

Eastman, C. (2009), Automated Assessment of Early Concept Designs. *Architectural Design*, 79: 52-57. https://doi.org/10.1002/ad.851

Eastman, C., Lee, J.-m., Jeong, Y.-s., Lee, J.-k. (2009). Automatic rule-based checking of building designs. *Automation in Construction*, 18 (8), 1011–1033. https://doi.org/10.1016/j.autcon.2009.07.002.

Eastman, C. M., Teicholz, P., Sacks, R., Liston, K. (2018). *BIM Handbook: A Guide to Building Information Modeling for Owners, Managers, Architects, Engineers, Contractors, and Fabricators.* John Wiley & Sons, Inc.

Garrett, J.H., Fenves, S.J. A knowledge-based standards processor for structural component design. *Engineering with Computers 2*, 219–238 (1987). https://doi.org/10.1007/BF01276414

Government of Singapore. (2016). *CORENET e-Information System.* Retrieved September 4, 2023, from https://www.corenet.gov.sg/general/e-info.aspx

Kannala, M. (2005)). Escape route analysis based on building information models: design and implementation. MSc thesis, *Department of Computer Science and Engineering, Helsinki University of Technology*, Helsinki.

Lee, H., Kim, J., Shin, M., Kim, I., & Lee, J.-k. (2014). A Demonstration of BIM-enabled Quantitative Circulation Analysis using BERA Language. In *The 31st International Symposium on Automation and Robotics in Construction and Mining, ISARC 2014* (pp. 202-209). The International Association for Automation and Robotics in Construction (IAARC).

Lee, H., Shin, J., Lee, J. K. (2016). BIM-enabled definition of a path object and its properties to evaluate building circulation using numerical data. *Journal of Asian Architecture and Building Engineering*, 15:3, 425-432. https://doi.org/10.3130/jaabe.15.425.

Lee, J. (2010). Automated checking of building requirements on circulation over a range of design phases (dissertation). *Georgia Institute of Technology*, Atlanta, GA.

Lee, J.-K., Eastman, C. M., Lee, J., Kannala, M., & Jeong, Y.-S. (2010). Computing Walking Distances within Buildings Using the Universal Circulation Network. *Environment and Planning B: Planning and Design*, 37(4), 628-645. https://doi.org/10.1068/b35124

Lee, J., Eastman, C.M., & Lee, Y. (2015). Implementation of a BIM Domain-Specific Language for the Building Environment Rule and Analysis. *Journal of Intelligent & Robotic Systems*, 79, 507-522. https://doi.org/10.1007/s10846-014-0117-7

Lee, J.K., Kim, M.J. (2014) BIM-Enabled Conceptual Modelling and Representation of Building Circulation. *International Journal of Advanced Robotic Systems,* 2014;11(8). https://doi.org/10.5772/58440

Migilinskas, D., Popov, V.A., Juocevičius, V., & Ustinovichius, L. (2013). The Benefits, Obstacles and Problems of Practical Bim Implementation. *Procedia Engineering*, 57, 767-774. https://doi.org/10.1016/j.proeng.2013.04.097

Peng, J., Liu, X. (2023). Automated code compliance checking research based on BIM and knowledge graph. *Scientific Reports,* 13:7065. https://doi.org/10.1038/s41598-023-34342-1.

Solibri Inc. (2023). *Solibri BIM software for architects, engineers and construction.* Retrieved September 6, 2023, from https://www.solibri.com/

Sydora, C., Stroulia, E. (2020). Rule-based compliance checking and generative design for building interiors using BIM. *Automation in Construction,* 120, 103368. https://doi.org/10.1016/j.autcon.2020.103368.

Urban Redevelopment Authority. (2023, July 26). *Gross Floor Area Covered Walkway and Linkages.* Retrieved August 14, 2023, from https://www.ura.gov.sg/Corporate/Guidelines/Development-Control/gross-floor-area/GFA/CoveredWalkwayandLinkages

Wang, J., Wang, X., Shou, W., Guo, J., & Hou, L. (2014). Development of BIM model fitness review system for modelling quality control. In *2014 International Conference on Computing in Civil and Building Engineering, 2014* (pp. 577-584). Orlando, Florida, USA: ASCE.

Werner, S., Krieg-Bruckner, B. & Herrmann, T. R. (2000). Modelling navigational knowledge by route graphs. *Lecture notes in computer science: Vol. 1849*. Spatial cognition 2 (pp. 295-316). Berlin: Springer.

BRIDGING BIM AND AI:

A Graph-BIM Encoding Approach For Detailed 3D Layout Generation Using Variational Graph Autoencoder

JIADONG LIANG[1], XIMING ZHONG[2] and IMMANUEL KOH[3]
[1]*Architectural Association School of Architecture.*
[2]*Aalto university school of art design and architecture.*
[3]*Singapore University of Technology and Design.*
[1]*liangjiadongarch@gmail.com, 0009-0005-9244-8674*
[2]*ximing.zhong@aalto.fi, 0000-0002-3060-6981*
[3]*immanuel_koh@sutd.edu.sg, 0000-0002-1181-1082*

Abstract. Building Information Modelling (BIM) data provides an abundant source with hierarchical and detailed information on architectural elements. Nevertheless, transforming BIM data into an understandable format for AI to learn and generate controllable and detailed three-dimensional (3D) models remains a significant research challenge. This paper explores an encoding approach for converting BIM data into graph-structured data for AI to learn 3D models, which we define as Graph-BIM encoding. We employ the graph reconstruction capabilities of a Variational Graph Autoencoder (VGAE) for the unsupervised learning of BIM data to identify a suitable encoding method. VGAE's graph generation capabilities also reason for spatial layouts. Results demonstrate that VGAE can reconstruct BIM 3D models with high accuracy, and can reason the entire spatial layout from partial layout information detailed with architectural components. The primary contribution of this research is to provide a novel encoding approach for bridging AI and BIM encoding. The Graph-BIM encoding method enables low-cost, self-supervised learning of diverse BIM data, capable of learning and understanding the complex relationships between architectural elements. Graph-BIM provides foundational encoding for training general-purpose AI models for 3D generation.

Keywords. BIM, Graph-Structured, Encoding Method, VGAE, Graph Reconstruction and Generation

1. Introduction

The current challenges in AI learning of 3D models stem from a lack of high-quality 3D datasets and the complexities of encoding 3D models to capture spatial layout features. Building Information Modelling (BIM) provides an abundant source, including 3D scanning and design and construction data. BIM encompasses detailed

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 221-230.

and hierarchical information on architectural elements, such as geometric data, spatial relationships, and material properties. Nevertheless, transforming BIM data into an understandable format for AI to generate controllable and detailed 3D models remains a significant research challenge(Bassir et al., 2023).

Encoding refers to transforming raw data into a format suitable for processing and learning by deep learning models. In this paper, the encoding method we discuss involves converting three-dimensional architectural models into data for AI learning. Typical methods of 3D model representation such as point clouds, voxels, and neural fields, can be encoded to facilitate AI learning(Tang et al., 2022; Zhong et al., 2023; Mildenhall et al., 2022). These methods primarily focus on the spatial attributes of individual elements but struggle to learn the interrelationships between elements within spatial layouts and face challenges in complex architectural spatial arrangements (Zhong et al., 2023).

We propose an encoding approach that transforms BIM data into graph-structured data for AI to learn 3D models, referencing encoding based on graph representation to align with the data structure of the BIM model, which we define as Graph-BIM. Graph-structured data, comprised of nodes and edges, serve to represent relationships between entities (Xu et al., 2018). Concurrently, BIM data includes detailed information about each architectural element, such as walls, doors, and windows, and spatial relationships between these elements. These two data structures exhibit inherent compatibility. In applying BIM data to graph-structured data, architectural elements are encoded as nodes. The physical connections, spatial relationships, or functional dependencies among these components are conceptualized as edge. This graph-structured data enables AI to capture the details of each architectural element and to articulate the interrelationships between these elements and the spatial layout. Recent studies have demonstrated the potential of graph-structured encoding methods in learning architectural 3D spatial layouts (Nauata et al., 2020; Zhong et al., 2023). However, they are beset with two principal issues. Firstly, encoding approaches from the aforementioned studies are constrained in the diversity of learnable data types, with irregular architectural spatial layouts posing significant challenges to learning. Secondly, these encoding methods predominantly rely on room function bubble diagrams or space voxels, limiting their direct control and generation capabilities for architectural components such as columns and walls. In contrast, Graph-BIM facilitates AI's comprehension and learning of multiple types of BIM data and reconstructs and generates precise, controllable detailed BIM models. To validate the feasibility of the Graph-BIM, we select the Variational Graph Auto-Encoders (VGAE) within the Graph Neural Networks (GNNs) framework for the experiment.

GNNs lie in their capability to process graph-structured data, effectively learning the intricate relationships among nodes(Xu et al., 2018). For instance, tools such as shear wall analysis tools have demonstrated the capability of GNNs to learn graph-structured data and reason spatial layouts (Zhao et al., 2023). However, GNNs in the aforementioned studies require extensive labelling of data during training. To facilitate more efficient learning of BIM data, we select the VGAE model. Comprising an encoder and decoder, VGAE facilitates self-supervised learning of BIM data, capturing the nodes' complex features and their topological interrelations within the graph, thereby generating a new graph (Kipf & Welling, 2016). Recent studies have demonstrated the successful application of VGAE in graphically representing molecular structures, enabling prediction edge generation and edge attributes and generating novel molecular structures (Bresson & Laurent, 2019). This suggests that VGAE also has the potential to learn BIM models represented as graph-structured data, understanding the relative spatial relationships between different architectural components. Therefore, we aim to explore suitable Graph-BIM encoding methods and predict building layouts by assessing the accuracy of spatial reconstructions by VGAE. Figure 1 illustrates the overall application process of Graph-BIM and the employed VGAE model, where users obtain detailed Revit models by inputting building outlines and inner control points.

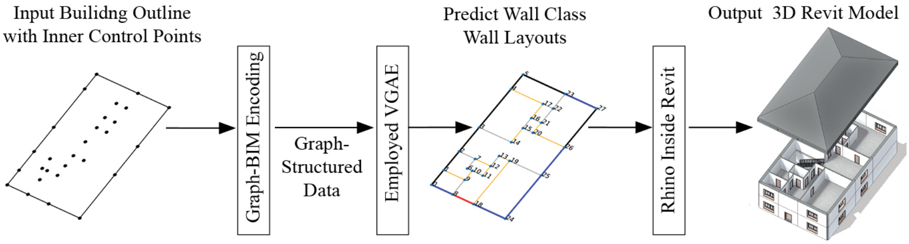


Figure 1. User Interface for 3D Layout Generation Using Graph-BIM and Employed VGAE

The principal contribution of the Graph-BIM encoding approach is to bridge AI and BIM. It empowers AI with the ability to self-supervise efficiently in the learning of BIM data features, facilitating cost-effective learning across multiple types of BIM-3D datasets. Graph-BIM encoding enables AI to grasp the relationship between architectural elements and to generate detailed 3D models.

2. Related Work

Nauata et al. (2020) develop House-gan, a system for generating housing layouts that comply with graph constraints. In House-gan, rooms are represented as nodes and their adjacencies are depicted as edges in the graph structure. Its abstraction oversimplifies the representation of detailed architectural elements such as columns, walls, and windows. Building-GNN encodes model voxels as nodes and their spatial interrelations as edges, allowing for precise spatial semantics control at the voxel level(Zhong et al., 2023). However, it faces challenges in learning irregular architectural layouts and generating architectural components within the voxel space. A GNNs-based method for shear wall layout prediction encodes structural elements like walls and windows as edges, with nodes representing component intersections(Zhao et al., 2023). This approach accurately predicts shear wall layouts, demonstrating the potential for architectural element control and prediction. Nevertheless, it primarily focuses on shear wall attributes, not encompassing the prediction of the overall spatial layout. The limitations in these studies in detailing architectural features, learning data types, and reason overall spatial layouts led to the development of the Graph-BIM encoding approach. Our method encodes architectural components such as columns and walls as graph nodes and edges, assigning them coordinates and specific positional attributes relative to the entire layout. Furthermore, to understand relationships among architectural elements, we propose incorporating room function connections as special attributes within our encoding framework.

Simonovsky and Komodakis (2018) demonstrate the use of GraphVae to output a predefined probabilistic fully connected graph in a single step using the decoder. By setting probability thresholds, it controls the likelihood of the existence of nodes and edges, thereby generating molecular graphs. This inspires us to test multiple threshold values to retain more complete reconstructed results of spatial layouts. The research of Permutation-Invariant Variational Autoencoder focuses on graph-level representation learning, introducing a novel model that addresses the graph reordering conundrum, and illustrating the management of node sequence uncertainty within graph structures(Winter et al., 2021). This provides specific technical guidance for handling the node sequence in various architectural layouts using VGAE. Shi et al. (2020) introduce a masking label prediction strategy in training the Message Passing Model, which involves randomly masking a certain proportion of input label information before making predictions. This inspires us to hide parts of the spatial layout information to reason the overall space, testing whether the Graph-BIM encoding allows the VGAE model to accurately learn the spatial features. Given these theoretical explorations and practical applications, we resolve to utilize the VGAE model's graph reconstruction capabilities to test suitable Graph-BIM encoding methods, while also exploring the potential of the VGAE model for reasoning architectural spatial layouts using Graph-BIM encoding BIM data.

Our research aims are to employ the VGAE model for reconstructing architectural spatial layouts and test appropriate Graph-BIM encoding methods, and to utilise VGAE 's graph generation capabilities to infer comprehensive architectural spatial layouts from localized information and generate new architectural spatial layouts.

3. Methodology

The Graph-BIM encoding experiment workflow is shown in Figure 2.

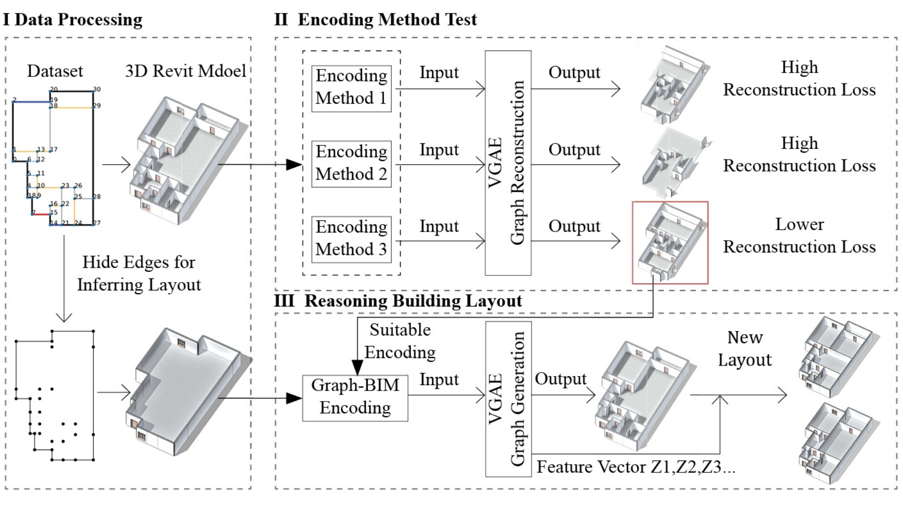


Figure 2. Workflow of Graph-BIM Encoding and VGAE Experiment

3.1. ENCODING METHOD

We encoded the endpoints of walls as nodes and non-overlapping columns were also encoded as nodes, with node features including spatial location and environmental information(Nauata et al., 2020; Zhao et al., 2023; Zhong et al., 2023). Walls were encoded as edges, with edge attributes representing wall types. We combined these three encoding methods and validated them using the VGAE model(Figure 3).

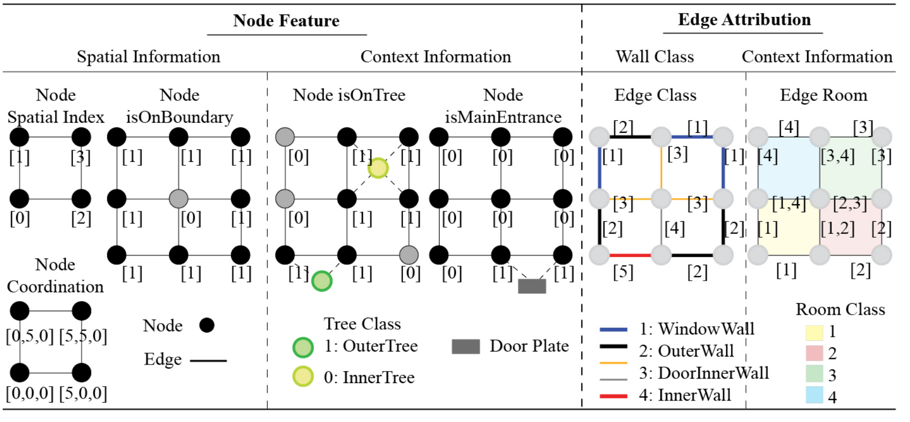


Figure 3. Encoding Methods

3.2. DATA PROCESSING

In this experiment, we utilized two datasets: the House-gan-MainEntrance (House-gan -ME) dataset and the Tree-Grid dataset(Figure 4). The House-gan -ME dataset was chosen for its variety in node numbers and irregular shapes, satisfying the requirements for datasets with diverse node quantities and non-uniform forms(Nauata et al., 2020). In addition to columns and wall endpoints, the Tree-Grid dataset included nodes representing environmental information. This approach facilitated the verification of the Graph-BIM encoding's accuracy in data format transformation for BIM data, unrestricted by the number of data nodes, node types, and layout shapes. We obtained BIM data through Revit to simulate the real workflow of AI learning BIM models.

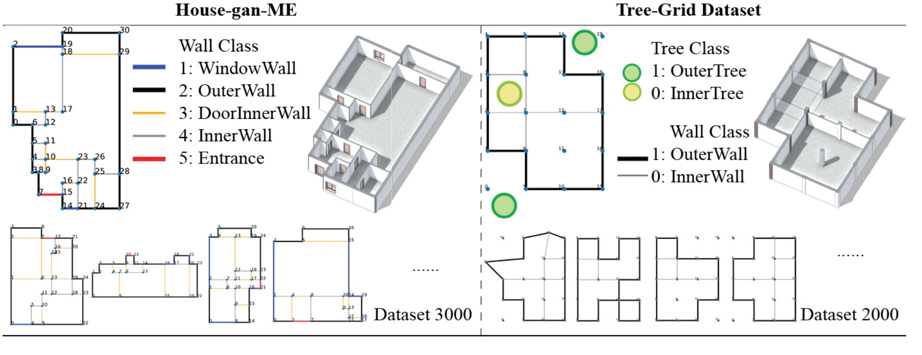


Figure 4. Dataset for VGAE Training

3.3. GRAPH RECONSTRUCTION AND GENERATION USING VGAE

Initially, the VGAE model underwent self-supervised graph reconstruction training. The output, consisting of the reconstructed adjacency matrix (positions of edges) and edge attributes, was compared with the original data's adjacency matrix and edge attributes to calculate the loss(Figure 5(a)). We conducted experiments with three different encoder models on three encoding methods to test suitable Graph-BIM encoding approaches. Then, we used this Graph-BIM encoding method to train the VGAE model's spatial reasoning capabilities. We concealed some of the walls(Edges) in datasets, retaining only the outermost walls of the building. The VGAE model outputted building layouts, which were compared with complete datasets for loss calculation(Figure 5(b)). Finally, we verified the capability to generate new layouts based on the Graph-BIM encoding. By using the VGAE to reconstruct different spatial layout feature vectors and linearly combining them in various proportions to create new feature vectors, the VGAE decoded these vectors into adjacency matrices and edge attributes. This generated result contained layout features proportionate to those in the original architectural space, as represented by the feature vectors.

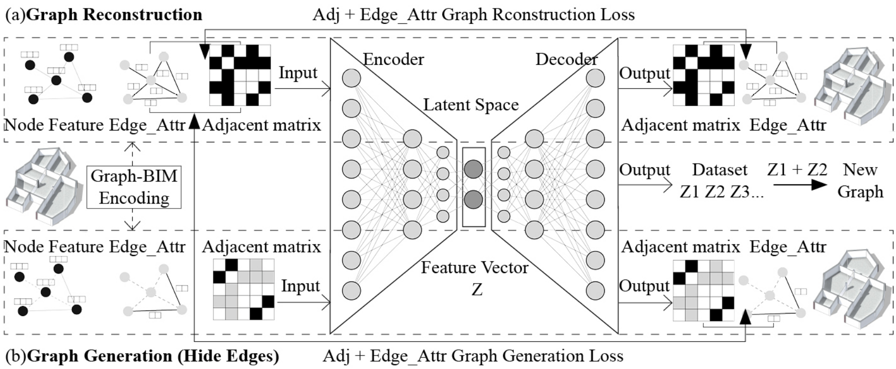


Figure 5. VGAE Graph Reconstruction and Generation

4. Result

4.1. RESULTS OF AD-GAN GENERATION AND DESIGN DECISIONS

The experimental results indicated that the VGAE model using Encoding Method 3 exhibits the best reconstruction capability. In the TransformerGCN model, the reconstructed wall positions and categories were almost identical to the original data. Observing the loss on the test set, Encoding Method 3 demonstrated the lowest reconstruction loss of spatial layouts across all three encoder models, with an average loss on the test set shown in Figure 6 being 0.098. This suggested that Encoding Method 3 enabled the VGAE to accurately learn the relationships between architectural components and spatial layouts. The reconstruction results of Encoding Method 2 were the least effective, either generating many irrelevant diagonal walls or missing most wall information. In the GINE model experiment, the reconstruction loss on the test set was as high as 0.756. This implied that relying solely on environmental information encoding was insufficient for effectively learning the finer details in BIM data. The reconstruction loss of Encoding Method 1 was close to that of Encoding Method 3, capable of reconstructing wall layouts and categories more completely. In the TransformerGCN experiment, the reconstruction loss decreased to 0.184, with only a minimal number of walls missing. In the GINE model experiment, the reconstructed spatial layout was missing some wall information, with a reconstruction loss of 0.574. This indicated that spatial information played an important role in encoding methods, but environmental information was still necessary to learn local features. Specific encoding contents on the Tree-Grid and House-gan-ME datasets were presented in Figure 7.

Encoder Model	TestDataset Mean Loss for Adj + Edge Attr Reconstruction		
	Encoding Method 1 Spatial Information +Edge Attribution	Encoding Method 2 Context Information +Edge Attribution	Encoding Method 3 Spatial+Context Information +Edge Attribution
GCN Kipf & Welling, 2016	0.227	0.613	0.117
GINE Xu et al., 2018	0.574	0.756	0.431
TransformerGCN Shi et al., 2020	0.184	0.412	0.098

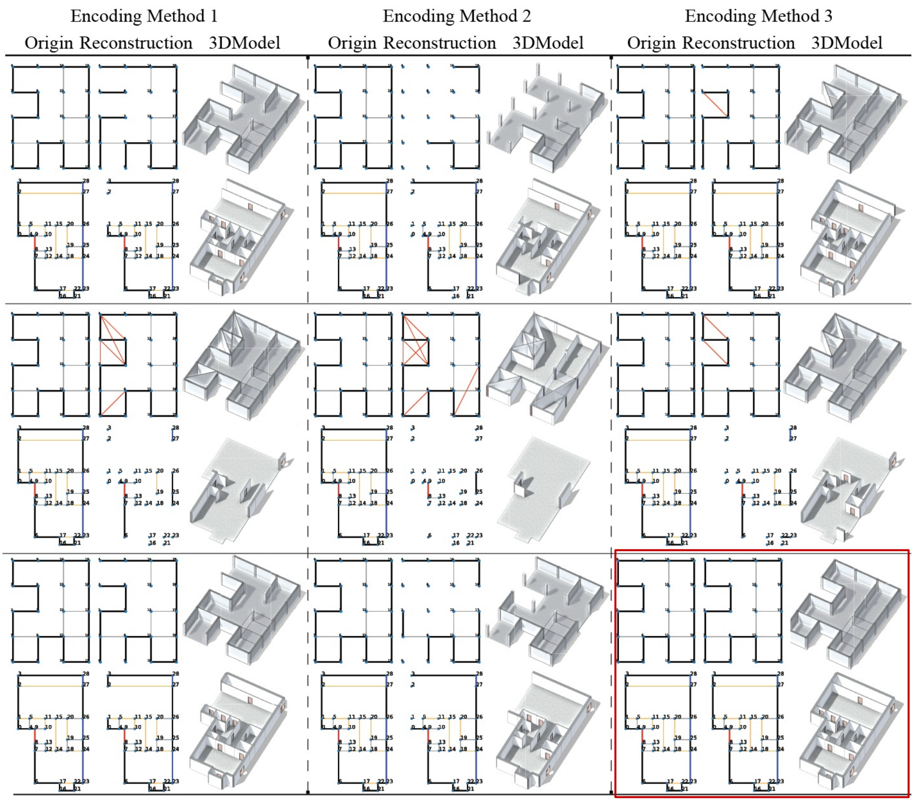


Figure 6. Graph Reconstruction Results for Testing Encoding Method

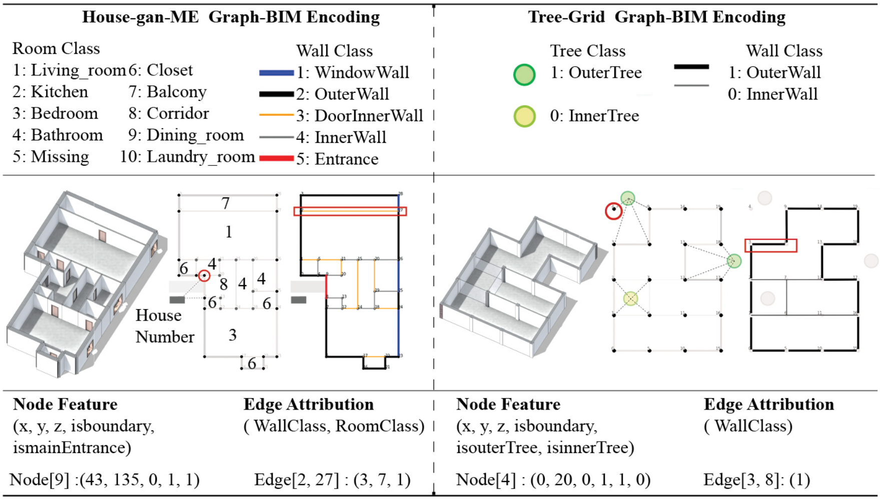


Figure 7. Graph-BIM Encoding Details

4.2. REASONING SPATIAL LAYOUT RESULTS

Figure 8 illustrated that the VGAE model employing the Graph-BIM encoding method was capable of reasoning complete spatial layout results on the test dataset based on partial layout information, and the spatial layouts satisfactorily met environmental constraints. The positions and categories of walls closely resembled the original spatial layout. The walls corresponding to door plates were precisely those of the main entrance category, with the lowest loss compared to the original at 0.067 (Figure 8(a)). Figure 8(d) demonstrated that both the external balconies and internal courtyards in the layout corresponded well with trees, with a reconstruction loss of 0.0743. Although some walls were lost in Figures 8(e) and (f), the impact on the overall spatial layout was not significant. The results substantiated the generalization capability of the Graph-BIM and VGAE could accurately generate spatial layouts that complied with environmental constraints for both regular and irregular spatial arrangements.

4.3. NEW LAYOUTS GENERATION RESULTS

Figure 9 showed the new layouts generated by concatenating feature vectors of different layouts. At a specific vector ratio of 0.3:0.7, the architectural layout tended more toward the feature2 layout, reducing a wall and forming a longitudinal spatial layout(Figure 9(a)). At a 0.5:0.5 ratio, the new spatial layout embodied the spatial characteristics of both original schemes, featuring both horizontal and longitudinal wall layouts, as depicted in Figure 9(b). When the ratio is 0.7:0.3(Figure 9(c)), the layout features leaned more towards feature 1, with an additional wall generated at the bottom of the space due to the influence of feature 2, yet the overall layout remained predominantly horizontal. The experimental results demonstrated that VGAE was capable of generating new building layouts from the combination of different feature vectors. This contributed to data sample augmentation and promoted diversity in the design.

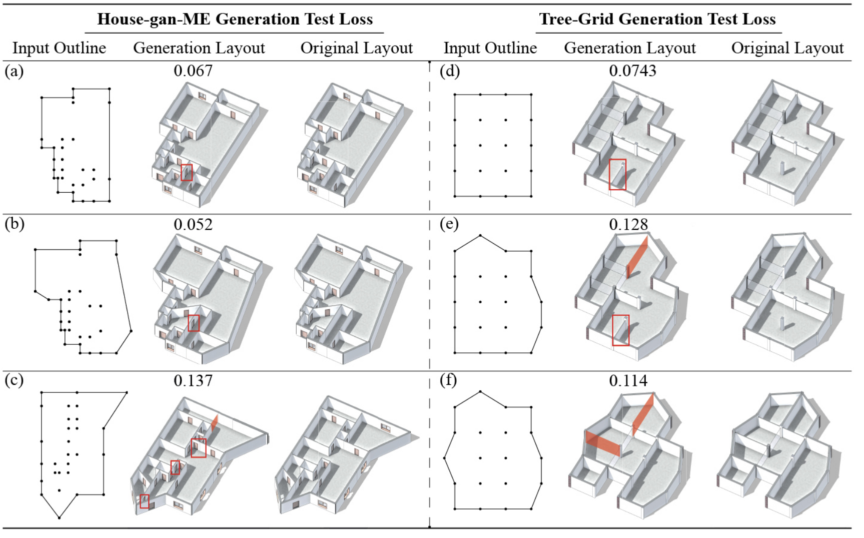


Figure 8. Graph Generation Results of Reasoning Building Layout

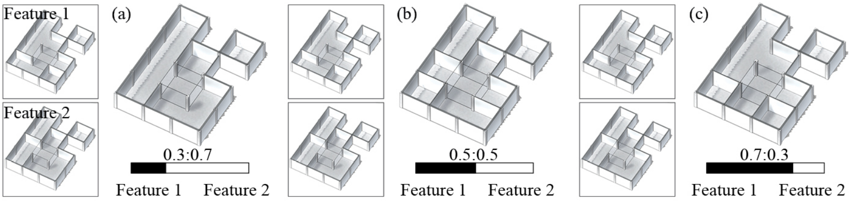


Figure 9. New Building Layouts Generation by Combining Feature Vectors

4.4. LIMITATION

The datasets primarily consist of spatial layouts formed by walls and columns. To enhance AI's performance in learning layouts, it is advisable to incorporate a wider range of architectural components, such as windows, doors, and furniture, into the Graph-BIM encoding method. This expansion aims to facilitate the learning and generation of more comprehensive 3D models. During the experiments, we employed three types of graph convolutional networks, which effectively aggregate node information. However, the integration of edge attributes into node features was suboptimal. A neural network model capable of precisely aggregating edge attributes and node features would be instrumental in deepening the machine's understanding of spatial layouts. Currently, in the graph generation phase, we only reconstructed the adjacency matrix and edge attributes of the dataset, predicting walls based on the original graph's nodes. We will train AI to develop the capability to generate nodes, enabling it to infer and generate overall spatial layouts with minimal information.

5. Conclusion

This paper explores a Graph-BIM encoding approach for converting BIM data into graph-structured data for AI training. The main contribution is to provide a bridge between AI and BIM encoding, enabling self-supervised learning of BIM data, allowing AI to learn multiple types of BIM datasets at a low cost and surpassing previous encoding methods in understanding the complex relationships between architectural elements and generating detailed 3D models. Our work is beneficial for training general-purpose AI models for 3D generation. Utilizing the Graph-BIM, architects can input outlines and interior control points to generate detailed models in Revit. Future developments may enable the training of a universal AI agent capable of addressing design requirements through the Graph-BIM. This offers a paradigm of end-to-end human-machine collaborative design for both architects and machines.

Acknowledgements

We would like to acknowledge author1 and author2 contributed equally to this work.

References

Bassir, D., Lodge, H., Chang, H., Majak, J., & Chen, G. (2023). Application of artificial intelligence and machine learning for BIM: Review. International Journal for Simulation and Multidisciplinary Design Optimization, 14, 5. https://doi.org/10.1051/smdo/2023005

Kipf, T. N., & Welling, M. (2016). Semi-supervised classification with graph convolutional networks. arXiv preprint arXiv:1609.02907.

Kipf, T. N., & Welling, M. (2016). Variational Graph Auto-Encoders (arXiv:1611.07308).

Mildenhall, B., Srinivasan, P. P., Tancik, M., Barron, J. T., Ramamoorthi, R., & Ng, R. (2022). NeRF: Representing scenes as neural radiance fields for view synthesis. Communications of the ACM, 65(1), 99–106. https://doi.org/10.1145/3503250

Nauata, N., Chang, K. H., Cheng, C. Y., Mori, G., & Furukawa, Y. (2020). House-gan: Relational generative adversarial networks for graph-constrained house layout generation. In Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part I 16 (pp. 162-177). Springer International Publishing.

Shi, Y., Huang, Z., Feng, S., Zhong, H., Wang, W., & Sun, Y. (2020). Masked label prediction: Unified message passing model for semi-supervised classification. arXiv preprint arXiv:2009.03509. https://doi.org/10.48550/ARXIV.2009.03509

Simonovsky, M., & Komodakis, N. (2018). Graphvae: Towards generation of small graphs using variational autoencoders. In Artificial Neural Networks and Machine Learning–ICANN 2018: 27th International Conference on Artificial Neural Networks, Rhodes, Greece, October 4-7, 2018, Proceedings, Part I 27 (pp. 412-422).

Tang, S., Li, X., Zheng, X., Wu, B., Wang, W., & Zhang, Y. (2022). BIM generation from 3D point clouds by combining 3D deep learning and improved morphological approach. Automation in Construction, 141, 104422. https://doi.org/10.1016/j.autcon.2022.104422

Winter, R., Noé, F., & Clevert, D. A. (2021). Permutation-invariant variational autoencoder for graph-level representation learning. Advances in Neural Information Processing Systems, 34, 9559-9573. http://arxiv.org/abs/2104.09856

Xu, K., Hu, W., Leskovec, J., & Jegelka, S. (2018). How Powerful are Graph Neural Networks? https://doi.org/10.48550/ARXIV.1810.00826

Zhao, P., Fei, Y., Huang, Y., Feng, Y., Liao, W., & Lu, X. (2023). Design-condition-informed shear wall layout design based on graph neural networks. Advanced Engineering Informatics, 58, 102190. https://doi.org/10.1016/j.aei.2023.102190

Zhong, X., Koh, I., & Fricker, P. (2023). Building-GNN: Exploring a co-design framework for generating controllable 3D building prototypes by graph and recurrent neural networks. In International Conference on Education and Research in Computer Aided Architectural Design in Europe (pp. 431-440). eCAADe.

EXTRACTING SPATIAL STRUCTURE FROM BUILDING INFORMATION MODELS

STYLIANOS DRITSAS[1], KHYSTELLE YEO[2] and CHERYL LEE[3]
[1,2,3]*Singapore University of Technology and Design.*
[1]*dritsas@sutd.edu.sg, 0000-0002-9609-2784*
[2]*khystelle_yeo@sutd.edu.sg 0009-0007-3478-0876*
[3]*cheryl_lee@sutd.edu.sg 0009-0001-0965-9368*

Abstract. The objective of this work is to automatically extract the spatial structure from Building Information Models for subsequently performing building circulatory analysis and regulatory building code compliance checking. This article presents the model comprehension methods employed, the challenges faced, and the solutions developed. We highlight the conceptual and technical limitations of current BIM paradigms and discuss how they may be improved.

Keywords. Building Information Modelling, Automated Extraction, Spatial Structure.

1. Introduction

Building Information Modelling (BIM) gained acceptance during the past few decades within the Architecture, Engineering, and Construction (AEC) industry (Eastman et al, 2011). It offers a natively 3D and semantically rich paradigm of design representation compared with Computer Aided Design (CAD), namely 2D drawings (Howard and Andersen, 2001). Object-based modelling overcomes the limitations associated with conflating geometry, topology, and semantics found in drawings such as plans and sections. Nevertheless, there are still numerous obstacles associated with recovering the spatial structure of a design represented using BIM.

Spatial structure is the logical graph of relationships between a space and its components and among spaces. The difficulty of deriving spatial structure springs from a fundamental problem: space itself is never explicitly modelled but merely implied as the result of its bounding surfaces. Thus, the definition of space geometrically and topologically, in the sense of the relationships among its semantic components, is limited and challenging to reconstruct. If this data was available then building models would allow queries combining geometry, topology, and semantics, such as which elements are associated with a space, which spaces share the common elements, which space contains an element, and which spaces are adjacent or nested within one another.

In theory, and to a certain extent in practice, spaces are captured by such notions as room elements within current BIM, however, these concepts are often poorly implemented and severely limited. Indicatively, the concept of a building's urban context in the sense of its embedding within space is unrepresentable. Additionally,

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 231-240.

determining the extents of spatial volumes that span several levels is problematic, expressing the structure of regions with an open plan space without physical boundaries is unclear, and buildings that conform to complex terrain conditions do not adhere to the conventional notion of a building level.

What would be beneficial for spatial analysis is a graph-theoretic object model where relationships are first-class concepts instead of the result of implicit queries among database tables and computationally expensive solid Boolean operations. This is a conceptual limitation of the BIM object model and its associated data structures (Eastman et al. 2010). Thus, today spaces are manually drawn in the plan, and this opens the door for challenges such as human errors which paired with a lack of industry modelling standards and quality control protocols result in inconsistencies that limit the utility of BIM (Migilinskasa et al. 2013; Wang et al. 2014; Chen and Luo, 2014).

2. Materials and Methods

The objective of the system developed was to automatically recover spaces and their relationships by examining a building information model. The application domain for the data is building circulation (Lee et al. 2010; Firas et al. 2023), feasibility analysis (Sherif and Eastman, 2011), and code compliance (Eastman et al. 2009; Balaban et al. 2012). The data required is semantically labelled regions and graphs conveying the relationship between those. Three basic region types are defined: (a) Spaces, capturing the notion of accessible, in the net internal floor plan area sense, within and around a building; (b) Accesses, expressing thresholds between spaces such as doors, elevators, escalators, and stairs. In addition, the category of virtual access is used for expressing portals where information is either unavailable or irrelevant, such as crossings to adjacent parcels; and (c) Visuals, contain information about BIM elements such as floors, walls, windows, columns, used for visualization (Fig. 1). Regions are expressed as surfaces bounded by piece-wise linear curves containing one exterior ring and optionally interior holes. Attributes associated with a region include a globally unique identifier, a reference to BIM elements, building level number and elevation, and categorical information such as usage information.

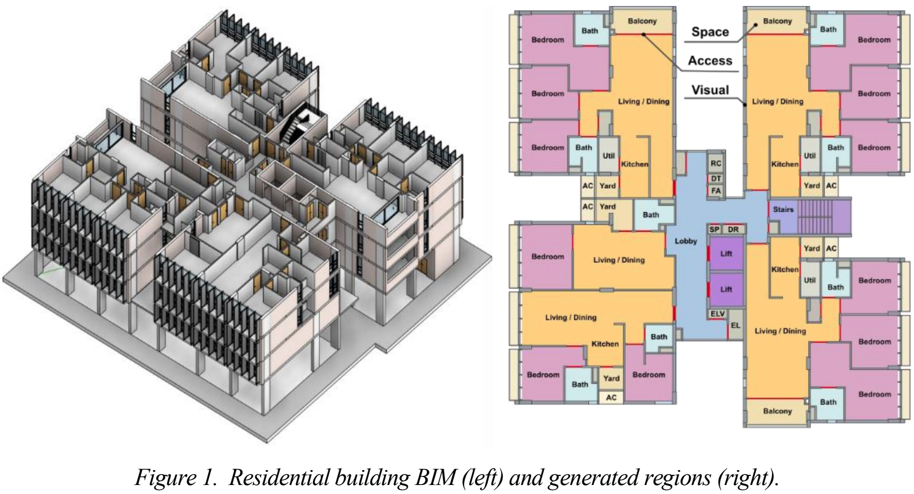


Figure 1. Residential building BIM (left) and generated regions (right).

The process of generating regions from BIM is organized in three parts: (a) Element filtering and translation, where relevant entities are converted to regions; (b) Spatial reconstruction, where Boolean operations are performed among sets of regions to derive spaces; (c) Space identification, where space use characteristics are recovered; and (d) Graph generation, where spatial adjacencies are recovered. The data extraction tool was developed for Revit using its C#/.Net API. Offset and Boolean operations were performed using the Clipper2 library (Johnson, 2003) and space usage recovery was performed using the Quickenshtein library (Turner, 2020).

Overall, the computation reconstructs building plans, by performing intersections and projections between a reference plane and relevant building elements. Shape processing using two-dimensional entities was motivated by their computational efficiency when compared to volume reconstruction methods. In addition, as the resulting analysis visualizations are aimed at reporting for general audiences, the plan representation is also more meaningful. Nevertheless, while the primary output of those operations is planar, vertical information is not discarded but stored as region attributes.

2.1. ELEMENT FILTIERING AND TRANSLATION

BIM entities are strongly typed, in the sense that they follow as rigid object-model hierarchy, with their definition semantically associated with building element categories, such as, floors, walls, columns, doors and windows. This is the case with both the IFC standard and BIM software (Kereshmeh and Eastman, 2014). Nevertheless, due to conceptual limitations (Barekati et al. 2015; Belksy et al. 2016) certain ideas do not always map to concrete entity types (Tomaz and Ziga, 2008). For instance, there are no specialized entities for sidewalks and roads and generally urban features. This motivates the use of untyped generic models or standard floor elements with material properties or text annotations that hint about type. BIM entities are therefore essentially dynamically typed, for extensibility, because they allow arbitrary properties attachment, in a key-value pair sense. Moreover, there is a lack of industry standards and model quality control protocols, manifested by the difficulty of data normalization. Indicatively, there is no simple way to extract the opening width from a door because this is stored as a key-value pair where the key may be “width”, “rough width”, “clear width” or even using keys in various languages based on the manufacturer's origin. Finally, inconsistent use of geometric and semantic information is often encountered, such as elevators and parking lots, which are sometimes modelled as either floors, lines, hatches or bounding box entities. Therefore, extensive use of heuristics is required for filtering and classification.

The system first parses the well-defined entity types. Those are organized in groups: (a) Accessible regions are projections of entities including floors, ramps, flat roofs and topography; (b) Spatial boundary regions include walls, columns, facades, railings and parapets, which may be partially crossing or projected onto the reference plane; (c) Special building components such as elevators, plumbing, furniture, parking lots, planting which are used for semantic analysis and visualization. For each element group we define a set of extraction parameters which express the upper and lower bounding offsets from the base level plane used for clipping solids and deriving crossing and projected regions. For example, vertical circulation elements require clipping above and below the base level to capture transit between floors. Walls are clipped above the finished floor level to extract boundaries and account for openings such as windows and doors. Additionally, curtain walls, apart from the base offset they are also clipped at a small offset above the finished floor level to capture raised seals and the relevant transoms are projected, because otherwise the mullions and glazing produce unrealistic space boundaries.

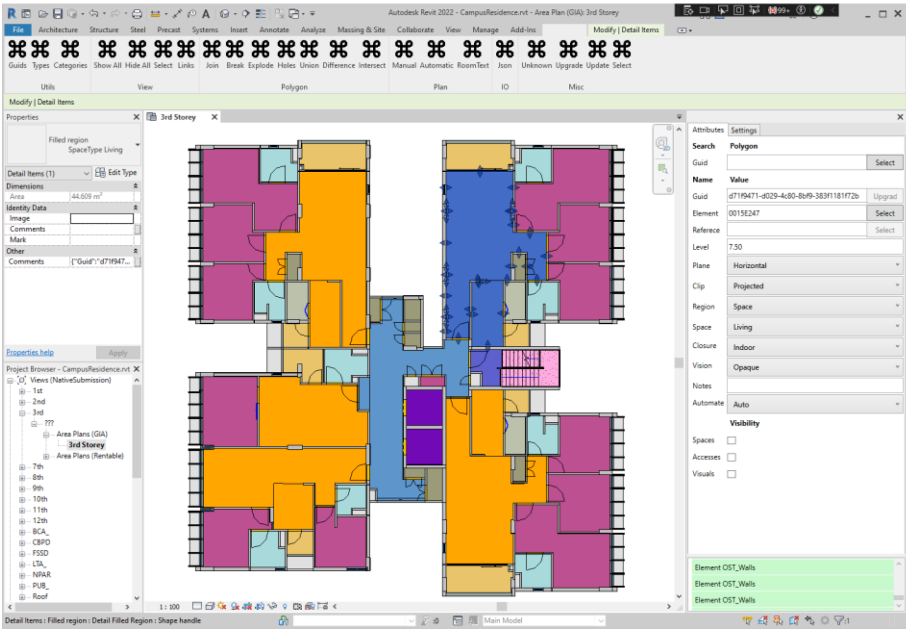


Figure 2. Software application developed for spatial information extraction.

2.2. SPATIAL RECONSTRUCTION

The objective of space recovery is to extract the boundary in the sense of the Net Internal Area (NIA) of building spaces. Unfortunately, while the concept of a room is conceptually supported by BIM standards and applications, it is limiting and often incorrectly modelled. For example, rooms may exclude, partially or completely include walls, often room tags are used to mark entire apartments within a residential building rather than individual rooms, room separation lines are often used inconsistently or to override the BIM logic, and finally the relationships between the space and its bounding elements is inconsistently tracked especially when room separators are used.

From the regions produced earlier we reconstruct spaces by performing a sequence of Boolean operations. An intermediate enclosure representation is computed for space bounding elements. This amounts to projection for walls without their associated doors and windows, the union of mullions, transoms and glazing for curtain walls, the projected union of their solids for parapets, railings, bollards etc. The union of spatial boundaries is computed resulting in a hierarchical structure of polygonal domains. The graph where each node represents a polygon ring, with each optionally containing additional nested rings, is then traversed outside-in. If we assume that all building elements have non-zero thickness, then the Jordan curve theorem allows us to unnest the hierarchy in spaces, should we alternate from exterior to interior at every other level of the tree structure. Thus, we arrive at a set of spatial domains which may contain holes, representing columns, voids, nested rooms, as well as hints as per whether they are indoor or outdoor. However, this is insufficient to completely determine the presence of spaces, especially urban spaces at the ground level, shafts, atria, voids, where the presence of slab or otherwise is important. Accessible regions per level are first united to produce a maximum available region which is Boolean differenced by the spatial boundary union to verify the presence of space in the sense of where people may be physically situated (Fig. 3).

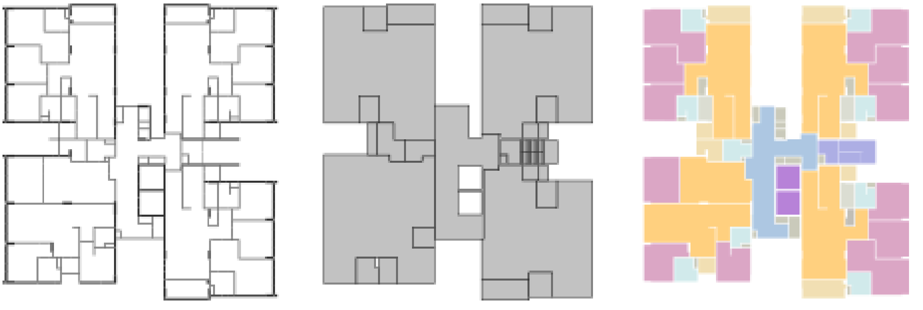


Figure 3. Space recovery using Boolean operations: Union of all spatial boundary regions (left). Union of all accessible regions (middle). Computed spatial regions (right).

Ideally, this procedure would suffice to compute all building spaces. However, we had to consider tolerances and modelling mistakes to improve the system's robustness. The most common problem with the models encountered is with walls that do not properly connect with one another causing gaps preventing enclosure or producing unions among adjacent spaces. This was approached using tolerances and heuristics. For instance, all spatial boundaries are dilated by a few centimetres to ensure overlap and thus spatial closure and separation. The offset is removed from the resulting spaces by erosion with the same tolerance. Accessible regions above the ground are artificially bounded with virtual walls. This ensures for instance that balconies are always accounted for, even if parapets or railings are missing, and such that spaces are always enclosed even if the slab edge does not meet curtain walls.

2.3. SPACE IDENTIFICATION

Spatial analysis, such as regulatory compliance checking, often requires information pertaining to the usage of each space. Therefore, spatial regions must be annotated automatically using BIM information available or inferred from the available semantic hints. While in theory a model's spatial element information may be sufficient, in practice it is rarely the case. The reason for this shortcoming is due to modelling practices, where rooms may be presents as either BIM entities or may appear using plain text annotations or even via linked CAD documents, and lack of standards in terms of a nomenclature and relevant abbreviations for space usage. We developed a list of typical space uses, such as "elevator", their known synonyms, such as "lift", "goods lift", "passenger lift" and abbreviations such as "elv". Space classes are also tagged as indoor or outdoor, primary, or secondary such as service rooms, and accessible by pedestrians, cyclists and/or vehicles (Fig. 4).

SpaceType	C	V	P	A	W	S	O	N	Color	Enumeration,	Wordlist
Atrium	0	0	1	1	0	0	0	0	#236bb0	Atrium = 0x00_0030_236BB0,	Atrium
Balcony	0	0	1	1	0	0	1	0	#e9c46a	Balcony = 0x01_0032_E9C46A,	Balcony,ServiceYard
Banquet	0	0	1	0	0	0	0	0	#d1625c	Banquet = 0x02_0020_D1625C,	Banquet,BanquetRoom
Bathroom	0	0	1	0	0	0	0	0	#a8dadc	Bathroom = 0x03_0020_A8DADC,	Bathroom,Ensuite,Bath,MasterB
Bedroom	0	0	1	0	0	0	0	0	#bc5090	Bedroom = 0x04_0020_BC5090,	Bedroom,Bedroom,Bdrm,Bed,Mast
Cafeteria	0	0	1	0	1	0	0	0	#eac4d5	Cafeteria = 0x05_0028_EAC4D5,	Cafeteria,Cafe,Coffee,CoffeeS
Catering	0	0	1	0	0	0	0	0	#ffcb69	Catering = 0x06_0020_FFCB69,	Catering,CateringRoom
Classroom	0	0	1	0	0	0	0	0	#ffb5a7	Classroom = 0x07_0020_FFB5A7,	Classroom,Class
Closet	0	0	1	0	0	1	0	1	#9b9b7a	Closet = 0x08_0025_9B9B7A,	Closet,Cabinet
Computer	0	0	1	0	0	1	0	0	#9b9b7a	Computer = 0x09_0024_9B9B7A,	Computer,ComputerRoom
Concourse	0	0	1	1	1	0	0	0	#00b4d8	Concourse = 0x0A_0038_00B4D8,	Concourse
Conference	0	0	1	0	1	0	0	0	#f8961e	Conference = 0x0B_0028_F8961E,	Conference,ConferenceRoom,Con
Convention	0	0	1	0	1	0	0	0	#b8e0d4	Convention = 0x0C_0028_B8E0D4,	Convention,ConventionRoom,Con
Corridor	0	0	1	1	0	0	0	0	#90e0ef	Corridor = 0x0D_0030_90E0EF,	Corridor,Corr
Cycling	1	0	1	0	0	0	0	0	#a5ffd6	Cycling = 0x0E_00A0_A5FFD6,	Cycling,Bicycle
Delivery	0	0	1	0	0	1	0	0	#cc5803	Delivery = 0x0F_0024_CC5803,	Delivery,Deliveries,DeliveryR
Dining	0	0	1	0	0	0	0	0	#ff6361	Dining = 0x10_0020_FF6361,	Dining,DiningRoom
Dressing	0	0	1	0	0	0	0	0	#de5a79	Dressing = 0x11_0020_DE5A79,	Dressing,Dress,DressingRoom
Electrical	0	0	1	0	1	1	0	0	#9b9b7a	Electrical = 0x12_002C_9B9B7A,	CableChamberRoom,TransformerR
Elevator	0	0	1	0	1	0	0	1	#7209b7	Elevator = 0x13_0029_7209B7,	Elevator,Elev,PasengerLift,Gc
Entrance	0	0	1	1	1	0	0	0	#0077b6	Entrance = 0x14_0038_0077B6,	Entrance,Entry
Equipment	0	0	1	0	0	1	0	0	#9b9b7a	Equipment = 0x15_0024_9B9B7A,	Equipment,EquipmentRoom,Eqpt

Figure 4. Space typology, attributes, and keywords mapping.

To recover space usage from the model we perform a series of heuristics: (a) The geometric centre of every space is tested for containment against all available rooms, (b) the centre of each room is tested against each space, (c) presence of special features such as stairs, lifts and plumbing equipment is used for inferring stairwells, lift shafts, bathrooms and services, (d) material associations are used to infer roads, planting, water bodies etc. (e) plan area is used to label small spaces such as closets and services, (f) lack of access is used to label shafts and voids. Finally, to decipher arbitrary room names and abbreviations thereof, we employ the Levenshtein editing distance metric against the table of known types. Spaces that fail all of the tests are marked as unknown and require manual labelling using the user interface tools.

2.4. GRAPH GENERATION

Topological information is recovered from the spatial adjacencies by preforming point-in-polygon and point-on-segment queries. The process is computationally expensive therefore a bounding box hierarchy is used to accelerate the reconstruction. The result of this process is a logical graph with nodes representing spaces and access regions and edges their adjacency. Additional data collected during this process is stored within the polygons such that their points and segments are associated with spatial boundaries. The logical graph my be visualized for simple spatial arrangements, where spaces have convex shape, and their nodes are situated in their geometric centres. For complex spatial arrangements this is not possible as the graphs' edges overlap spatial boundaries. A polygon skeletonization procedure was thus employed, based on the Medial Axis Transform (Blum, 1967) shape analysis methodology (Lee, 2004), to generate spatially situated graphs that are contextually meaningful (Fig. 5).

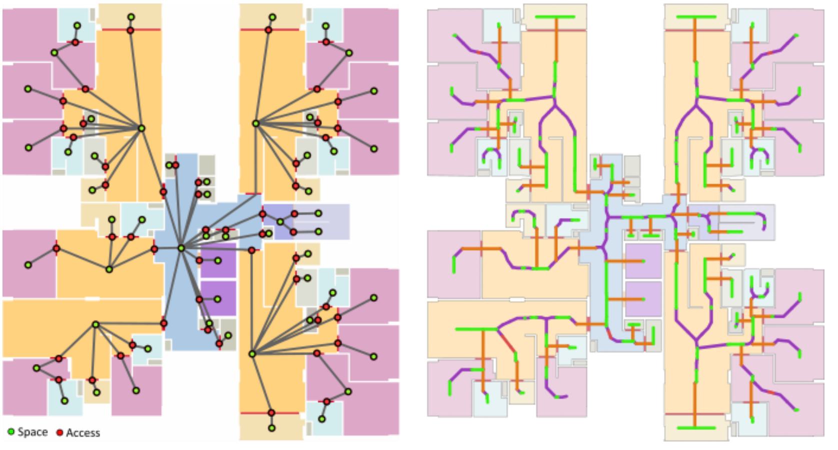


Figure 5. *Logical graph with nodes situated at the geometric middle of spaces and accesses (left) and graph generated using the Medial Axis Transform method (right).*

3. Results

The system was developed using the university campus residential buildings as its initial case study. The campus residences BIM represents a benchmark or prototype of a properly modelled building which can be completely automatically analysed. There were no special procedures employed to build this model other than using the correct BIM entities for each purpose, ensuring that elements were properly hosted, walls met at their end points and rooms were all labelled. Subsequently, we validated the system using ten models submitted by private architecture firms for regulatory assessment to planning authorities. These include residential, commercial, transportation and mixed-use developments. The names and locations of the developments cannot be disclosed for confidentiality purposes and as such labelled as D1 to D10 (Fig. 6).

		D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	
Project Information	Site	28.6	11.6	36.4	14.7	10.9	6.2	9.8	18.6	17.0	39.2	x1000 m²
	GFA	71.5	82.4	64.1	47.9	38.2	22.2	28.4	67.0	4.9	168.7	x1000 m²
	Levels	14	17	12	18	26	10	11	19	3	7	
Polygon Analysis	Generated	526	431	805	184	280	307	251	724	231	1088	
	Relabelled	208	160	73	199	50	13	31	260	96	45	
	Drawn	816	86	655	431	387	125	182	454	258	701	
	Completed	1146	510	1320	687	428	427	499	733	474	1527	

Figure 6. *Project information including site area, gross floor area and number of floors for each of the ten developments analysed. Number of spatial polygons automatically generated and those that had to be manually relabelled and geometrically corrected.*

Overall, the tools developed were able to extract spatial structure from all projects but not without significant challenges requiring manual intervention or creating new heuristics. Corrections were performed for (a) malformed and incomplete geometries and (b) rooms with missing or incorrect tags were relabelled. The number of polygons manually adjusted for geometry corrections was 52% with standard deviation of 21% while relabelling was required for 16% of all polygons with standard deviation circa 12%. The rather pronounced magnitude of the standard deviation highlights the wide range of quality characteristics of the models which for geometry spans from 17% to 90% and for room labels from 3% to 35%. We did not identify correlations between the projects' sizes, in terms of gross floor area (GFA), and the number of geometry and metadata errors, with coefficients of -0.1 and -0.09, respectively.

The specific problems uncovered were explained earlier under the methods section and are presented below (Fig. 7). They are categorized under three semantic groups, namely (a) Urban: related to problems such as absence of project geolocation, properly modelled roads, sidewalks, pedestrian crossings and landscape features; (b) Building: related to proper annotation of rooms, closure of spaces, elements placement, and use of appropriate BIM element categories; (c) Mobility: related to presence of circulatory features such as escalators, elevators and vehicular information such as cyclist and motorist driving and parking related BIM elements. The average across all projects and categories are presented for highlighting the frequency and by association the severity of the problems. Geometric mistakes are more frequently encountered compared to semantic errors with only one in ten projects having properly enclosed spaces. We also note that while BIM is the primary medium of documentation there are still substantial amounts of information contained within associated CAD drawings. In summary, the sources of those problems may be attributed to (a) conceptual limitations of today’s BIM, (b) lack of digital industry standards and (c) lack of methods for model quality control. We observed significant challenges in spatial structure recovery at ground level of all buildings examined in contrast with upper floors. We attribute this to a conceptual hiatus between BIM and GIS, the building itself and its surrounding environment.

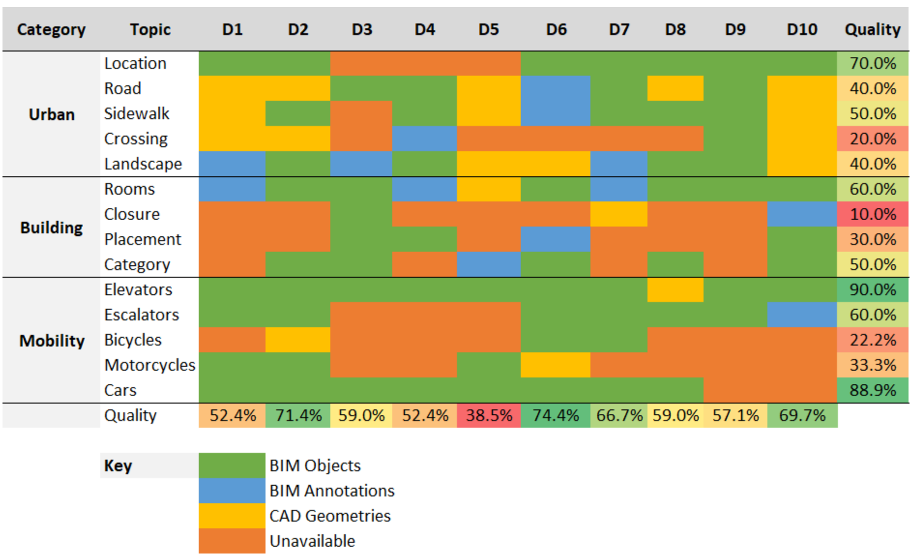


Figure 7. Summary of typical errors and overall model quality assessment per project and category.

4. Discussion

The challenges identified earlier motivated breaking the spatial structure recovery processes into separate tasks, namely (a) fully automated, where the entire model may be processed without user intervention, (b) quasi-automated, where elements are extracted under user-tuned settings and/or when automatically generated information are adjusted, such as updating semantic attributes, via the user-interface, and (c) manual extraction where regions are manually drawn in the CAD sense and their attributes are entered via the user interface.

The university residence model and the analysis tools were developed in parallel. Use of the spatial structure recovery tools during the development of the model in retrospect was instrumental in improving its quality and ensuring that no geometric or semantic issues were present in the final model. This is because periodic execution of the processes highlighted modelling mistakes such as geometric, topological, and semantic omissions and/or ambiguities. Therefore, we foresee that such tools that aim to comprehend BIM models may also be used for quality control as well as education.

A fundamental limitation of the tools presented is in that they primarily operate in two dimensions which may be insufficient for spatial analytical processes such as acoustics and environmental performance evaluation (Alam and Ham, 2014). Additionally, recovering information using heuristics resulted in a substantial number of parameters requiring careful calibration when there is a need for adjustment. Perhaps the use of machine learning may assist in overcoming this problem in the future (Koo et al. 2019; Ahmadpanah et al. 2023).

5. Conclusions

In conclusion, we presented a process for recovering spatial structure from BIM highlighting the challenges identified and the solutions developed. The investigation demonstrated the limitations of current BIM technologies and offers hints about how they may be improved in the future. Applications employing the information extracted from BIM, namely circulatory analysis and building code compliance is the subject of forthcoming work.

Acknowledgements

This research is supported by the National Research Foundation, Prime Minister's Office, Singapore under its Cities of Tomorrow R&D Programme (COT-H1-2020-2). Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not reflect the views of National Research Foundation, Singapore and Ministry of National Development, Singapore.

References

Ahmadpanah H., Haidar A. and Latifi, SM.: 2023. BIM and Machine Learning (ML) Integration in Design Coordination: Using ML to automate object classification for clash detection. ECAADE, Graz, 619-628.

Alam J. and Ham J. J.: 2014, Towards A BIM-based energy rating system: Comparisons between FirstRate5 and ArchiCAD EcoDesigner. CAADRIA, Kyoto, 285-294.

Balaban O., Yağmur Kilimci, ES. Cagdas, G.: 2012. Automated Code Compliance Checking Model for Fire Egress Codes. ECAADE, Prague, 117-125.
Barekati E. Clayton M. J. and Yan, W.: 2015. A BIM-compatible schema for architectural programming information. CAAD Futures, Liege, 311-328.
Belsky M., Sacks R. and Brilakis I.: 2016. Semantic Enrichment for Building Information Modeling. Computer-Aided Civil and Infrastructure Engineering, 31, 261-274.
Blum H.: 1967. A transformation for extracting new descriptors of shape. Models for the Perception of Speech and Visual Form: 362-380.
Chen LJ. and Luo H.: 2014. A BIM-based construction quality management model and its applications. Automation in Construction, 46, 64-73.
Eastman C., Lee JM., Jeong YS., and Lee JK.: 2009. Automatic rule-based checking of building designs, Automation in Construction, 18(8), 1011-1033.
Eastman C. M., Jeong YS., Sacks R. and Kaner, I.: 2010. Exchange model and exchange object concepts for implementation of national BIM standard, Journal of Computing in Civil Engineering, 24(1), 25-34.
Eastman C., Teicholz P., Sacks R., and Liston K.: 2011. BIM handbook: A Guide to Building Information Modeling for Owners, Managers, Designers, Engineers and Contractors. Wiley.
Firas AD., Wei Y. and Edin J.: 2023. Campusim: An Integrated Parametric BIM for Campus Design Simulation and Optimization. CAADRIA, Ahmedabad, 471-480.
Howard R. and Andersen J. L.: 2001. Classification of building information - European and IT systems. ITcon. Mpumalunga, 9:1-14.
Johnson A.: 2003. Clipper2: A Polygon Clipping and Offsetting Library. Available at https://github.com/AngusJohnson/Clipper2
Kereshmeh A. and Eastman C.: 2014, Categorization of building product models in BIM Content Library portals, SIGRADI, Montevideo, 370-374.
Koo B., La S., Cho NW. and Yu Y.: 2019, Using support vector machines to classify building elements for checking the semantic integrity of building information models. Automation in Construction, 98, 183-194.
Lee J.: 2004. A spatial access oriented implementation of a 3-d GIS topological data model for urban entities. GeoInformatica 8 (3): 237-264.
Lee JK. Eastman C. Lee J. Kannala M. and Jeong YS.: 2010. Computing Walking Distances within Buildings Using the Universal Circulation Network. Environment and Planning B. 37(4), 628-645.
Migilinskasa D., Popovb V., Juoceviciusc V., and Ustinovichiusd L.: 2013. The Benefits, Obstacles and Problems of Practical Bim Implementation. Procedia Engineering, 57, 767-774.
Sherif A., Lee J. and Eastman C.: 2011. Automated Cost Analysis of Concept Design BIM Models. CAAD Futures, Liege, 403-418.
Tomaz P. and Ziga T.: 2008: Interoperability in practice: Geometric data exchange using the IFC standard. ITcon, 13, 362-380.
Turner J.: 2020. Quickenshtein: A quick and memory efficient Levenshtein Distance calculator for .NET. Available at: https://github.com/Turnerj/Quickenshtein
Wang J., Wang X., Shou W., Guo J., and Hou L.: 2014. Development of BIM Model Fitness Review System for Modelling Quality Control. Computing in Civil and Building Engineering, 577-584.

MORE WORDS FROM FACILITY OPERATION MANAGERS: TEXT-BASED BUILDING INFORMATION FOR INCLUSIVE ACCESSIBILITY

AI KASAI[1] and TSUKASA ISHIZAWA[2]
[1]*Project Researcher, The Institute of Industrial Science, The University of Tokyo*
[2]*Project Associate Professor, The Institute of Industrial Science, The University of Tokyo*
[1]*kasai-ai@iis.u-tokyo.ac.jp, 0009-0003-7510-1332*
[2]*ishizawa@iis.u-tokyo.ac.jp, 0000-0001-5089-2877*

Abstract. In the domain of Facility Management (FM), the implementation of Building Information Modelling (BIM) poses significant challenges for the Facility Operation Manager (FOMr) in terms of effective integration into their regular workflow. Employing the Soft Systems Methodology (SSM), this study highlights a conducive environment that enhances the usability of BIM for the FOMr. By expanding the application scenarios of BIM within the FM field, we contribute to completing the lifecycle of building models. We introduce 'Text-based Building Information' (TxBI), a system designed to complement Industry Foundation Classes (IFC) while emphasizing a text-centric approach. TxBI strategically focuses on systematically integrating Mechanical, Electrical, and Plumbing (MEP) components, which are pivotal for daily FM operations. It generates a simple yet versatile dataset derived from the as-built BIM deliverable, channeling it to Computerized Maintenance Management Systems (CMMS) and streamlined 3D systems. The utility of TxBI is exemplified through its application in commercial software environments, showcasing the creation of lightweight digital twins of completed buildings. Significantly enhancing FOMrs' access to building information, TxBI establishes a more versatile data environment for all building stakeholders. The accumulation of comprehensive information throughout the Operation and Maintenance (O&M) phase culminates in enhanced asset management and extended longevity of buildings.

Keywords. BIM for FM, operation and management, facility operation managers, text-based building information, long-life building, real time visualization from text, digital twins.

1. Introduction

Building Information Modelling (BIM) has been expected to play an increasingly critical role in the domain of facility management (FM), enhancing the effectiveness of Life Cycle Management (LCM) among other benefits (Wijeratne et al., 2023).

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 241-250.

However, the utilization of BIM in FM is not as active compared to the design and construction phases (Rogage and Greenwood, 2020). This ongoing issue is compounded by challenges such as the limited awareness of building owners and discontinuity of BIM management after the hand-over (Cavka et al., 2017). For most medium to large-scale buildings, Facility Operations Managers (FOMrs) are the individuals who grapple with the daily operational issues of building once it is handed over from the builders. It should be noted that many FOMr lack the adequate BIM operational skills due to various reasons, including a lack of adequate funding, resources, and time for training. (Ashworth et al., 2016).

Presently, BIM for FM software and related research typically assume the exchange of building information via the Industry Foundation Classes (IFC), a format that various BIM software can commonly output (Lai and Deng, 2018). While the creators of these models have easy access to the IFC contents, the same cannot be said for FOMr. A significant drawback for FOMr is typically their limited operational skills with BIM-related software and their unfamiliarity with the model contents (Ashworth et al., 2016). Consequently, improving the means for non-BIM stakeholders to access informational contents of building asset is a meaningful step toward realizing the active utilization of the BIM-based data sets in the day-to-day operations by the FOMr.

2. Research Objective

Given the paramount importance of extending the building service life in today's context, the digital transformation of operations exerts a profound influence on the entire building life cycle (Wijeratne et al., 2023). Throughout the Operation and Management (O&M) phase, which constitutes the majority of a building's lifespan, it is crucial that building information remains readily accessible (Carvalho et al., 2023). More importantly, operational records should be systematically accumulated in a database and form an integral part of the building information asset. The research objective of this paper is predicated on a system proposal that addresses the aforementioned challenges and facilitates high accessibility for a wider range of building stakeholders throughout the extended building lifespan.

Numerous prior studies have addressed this issue. For example, Altohami et al. (2021) have utilized IFC data within a Semantic Web environment to facilitate seamless integration of BIM with the Internet of Things (IoT) devices. Kim et al. (2018) have enhanced the interoperability and accessibility of FM data through the Semantic Web. Katsigarakis et al. (2022) have proposed an environment that converts IFC to OBJ files and leverages the Resource Description Framework (RDF) format to promote diverse data integration. However, as previously mentioned, the challenges associated with handling the model itself remain unresolved. One of the most accessible, lean, and human-readable datasets is structured text, such as comma-separated values (CSV). This research specifically addresses use cases by FOMr, proposing a concept of text-based information enriched with relevant facility-related details and its case studies.

3. Method

This study employs the Soft Systems Methodology (SSM), a framework designed for the analysis of complex and ambiguous problems (Checkland, P and Poulter, 2006). The subsequent sections adhere to the structure of SSM as follows. In Section 4, we will identify system issues and propose a conceptual model. This model will then be validated through case studies in Subsection 4.4. The ensuing sections will encompass a discussion, limitations of this paper, and the conclusion. Figure 1 illustrates the correspondence between this workflow and the SSM framework.

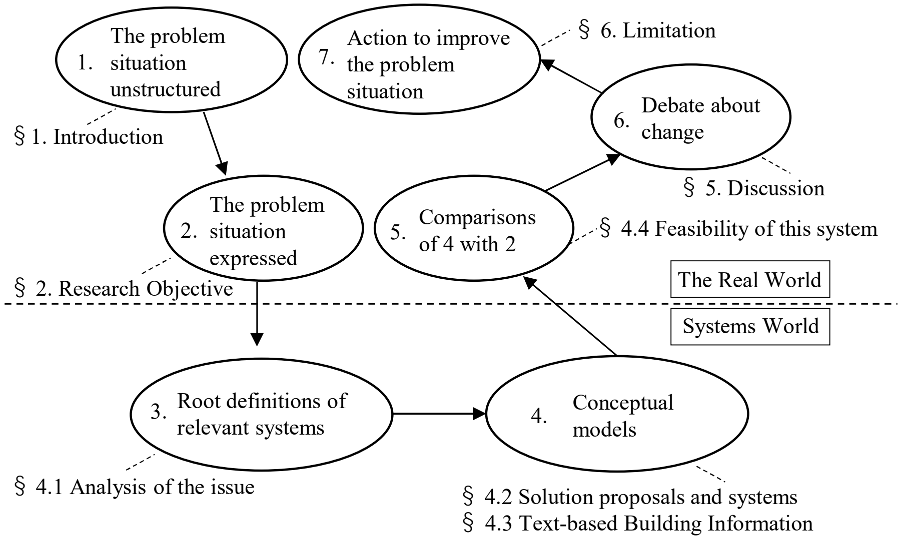


Figure 1. The research methodology based on the seven stages of Soft Systems Methodology.

4. Analysis

4.1. ANALYSIS OF THE ISSUE

In this section, we address the deficiencies of the current system and establish a fundamental definition for the ideal system. When transmitting BIM information as text data, converting BIM to IFC format has the advantage of allowing data sharing in the consistent format from various BIM software, thereby facilitating system standardization (Dhillon et al., 2014). Furthermore, BIM geometry often encapsulates building-specific shapes such as floors and walls, which can sometimes be of irregular complex shapes, making IFC a beneficial format for preserving such shapes. However, IFC primarily serves as an intermediary format for BIM and Construction Information Modeling (CIM), and the software capable of processing it still encounters numerous challenges. Custom software that can process IFC format necessitates users to learn specific operational methods (Pärn et al., 2017), thereby prolonging the learning period for BIM creators and FOMr, consequently lengthening post-construction preparation tasks.

Moreover, whether BIM information is correctly output in IFC format depends on the applications. Although IFC is a unified format convertible from various BIM software, some applications cause information conversion losses in IFC format (Lai and Deng, 2018). While BIM creators can reconcile these potential discrepancies by cross-referencing BIM with IFC, FOMrs, who often face challenges in utilizing BIM software as highlighted by Lin et al. (2022), may find it difficult to verify the accuracy of IFC format against the as-built structure. Consequently, FOMrs face challenges in determining whether losses occurred during the conversion of BIM to IFC, or if they were due to elements originally not included in the BIM.

Conversely, what if we could directly output structured text data from the original BIM data without going through IFC? Text data, being lightweight and versatile, not only facilitates easier integration with existing systems but also enables the visualization of such data in formats familiar to FOMrs, such as spreadsheets, which are openly accessible. This adaptation significantly reduces the verification burden on FOMrs. This shift from a model-centric to a text-centric approach increases the likeliness that the system could be utilized throughout the building lifecycle.

4.2. SOLUTION PROPOSALS AND SYSTEMS

Here, we explore a text-based system with a focus on the high-priority areas of O&M, particularly Mechanical, Electrical, and Plumbing (MEP). Unlike architectural and structural components, MEP elements exist as parts of a continuous system. Each component is simple enough in form to estimate its shape from text information alone. Figure 2 illustrates the representation of the BIM geometry's shape using attributes. For instance, consider piping: given that pipes are cylindrical, knowing the total length and cross-sectional area allows for an approximate understanding of their external shape. However, due to the characteristic of MEP components being segmented and connected, the coordinates indicating their position become critical. It is also possible to obtain height information from reference points or centers. For instance, if a reference point on a site has three-dimensional coordinates (x, y, z), then the geometry's coordinates can be transformed into relative coordinates from this reference point $(x+\alpha, y+\beta, z+\gamma)$. Clarification of the coordinate information pertaining to the central or reference points of each geometry facilitates the accurate acquisition of positional data in relation to other components within a system.

Notably, while MEP systems encompass components with complex shapes, such as various equipment, converting most of these components to simpler structures does not significantly impact the actual use of FM. Any irregularly shaped components can be effectively accommodated by incorporating approximate dimension information as additional attributes. Consequently, this allows for a direct conversion of MEP components from BIM software to text data, without necessarily going through IFC conversion. Leveraging these MEP characteristics, we propose Text-based Building Information (TxBI) that facilitates the linkage of BIM to FM software without IFC intermediation, utilizing text information.

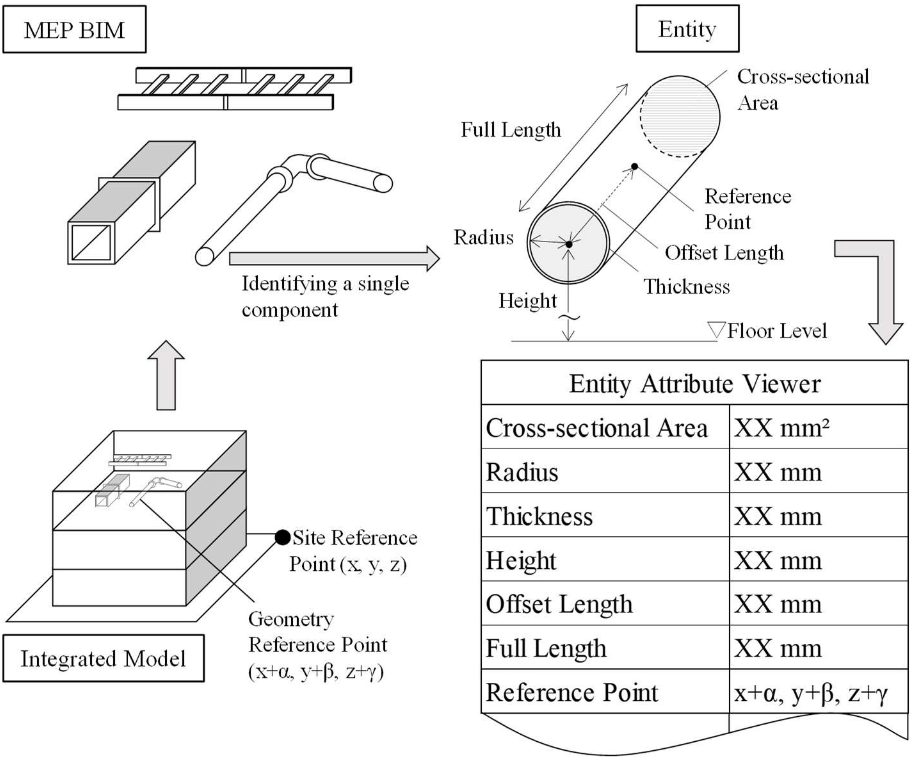


Figure 2. The representation of the BIM geometry shape using attributes.

4.3. TEXT-BASED BUILDING INFORMATION (TXBI)

As a concrete solution, we propose a conceptual model: TxBI system, which is predicated on text data as the primary means of information conversion. The detail of concept is illustrated in Figure 3. The assumed situation is that the as-built project models, mostly created by the project BIM specialists during the design and construction phases, are fully succeeded to the owner and the FOMr in the post-construction and O&M/FM phases.

A. Separating BIM into MEP BIM

When architectural, structural and MEP are initially coordinated within a federated model, it becomes necessary to segregate the MEP components from the rest. This critical task of separation should be undertaken by the project BIM personnel to accurately distinguish and categorize these distinct components within the model.

B. Outputting Text from MEP BIM Model

Text data, which includes the shape information and coordinates linked to each attribute of the MEP BIM geometry, is exported using the attribute export functionality. Numerous modeling software incorporated capabilities for exporting structured text data in universal data formats such as CSV. Even in cases where BIM software does not possess a native text export feature, the add-on program or tailor-made script will be able to fulfill this need.

C. Transfer of Text Data from BIM to TxBI: Before completion

Prior to construction completion, project BIM manager hands over the text data extracted from BIM to the FOMr. The FOMr then conducts a thorough verification of this data on the TxBI platform, ensuring that no critical information pertinent to O&M or FM is omitted. Post-completion, in the absence of a BIM operator, the FOMr assumes the responsibility of adding or modifying information directly on TxBI. This strategy effectively eliminates the necessity for the FOMr to engage with the BIM system both before and after the project's completion. The TxBI interface will become familiar to many project stakeholders because it behaves as if a standard spreadsheet, facilitating straightforward data input and manipulation for the FOMr. Moreover, it is anticipated that TxBI will enable the import of information into the BIM system at any stage post-completion. Additionally, unique identifiers derived from BIM geometry, potentially linked to attributes via a Globally Unique Identifier (GUID), will be managed (Kang and Hong, 2015). To enhance the precision in identifying individual components, multiple identifiers, such as the BIM creation date and the creator's details, might be utilized.

D. Integration with CMMS

TxBI will be then seamlessly integrated into a computerized maintenance management system (CMMS). This integration facilitates the utilization of BIM-derived information during the FM phase. The FOMr predetermine the specific information that is required within the CMMS, ensuring that only this designated information is exported from the TxBI system. A reciprocal interaction between TxBI and CMMS is established, allowing data input into CMMS to be correspondingly reflected back into TxBI.

E. Simple Visualization System

Now we delineate the real-time visualization from the text. A process of simple re-modeling can be undertaken by utilizing the exported information. The re-modeling process involves calculating the shape representation based on geometric coordinate points, including dimensions like diameters and cross-sections. This approach addresses the inherent limitations of textual descriptions, such as the challenges in intuitively grasping shapes, by enabling shape reconstruction from text without the necessity for IFC.

Previous studies have documented the use of BIM in web browsers for 3D visualization and FM application (Wijeratne et al., 2023). However, in our focus on MEP systems, further simplifications are achievable in comparison to other scenarios. For instance, a 2D representation with additional height information or the use of pins to mark specific locations can be employed for a more simplified visualization. Such simplifications not only facilitate digital twin applications on the web but also enable their practical implementation onsite through Augmented Reality (AR) technologies.

F. Transferring Text Data from TxBI back to BIM

In the context of renovation projects, BIM models developed during the initial construction phase are repurposed to generate detailed renovation plans. Consequently, the extensive data accumulated in TxBI during O&M phases can seamlessly integrated back into the BIM environment. This integration effectively incorporates crucial FM information into the existing BIM model, thereby enhancing its utility and accuracy for future renovation activities.

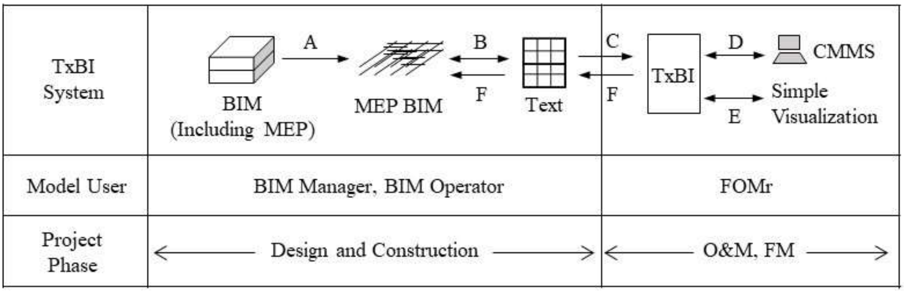


Figure 3. The workflow and task allocation of TxBI's system.

4.4. FEASIBILITY OF THIS SYSTEM

We validate the feasibility and plausibility of the proposed system by presenting a practical implementation. A commercial BIM environment named Pinspect is designed as an Augmented Reality application (ARapp) which integrates BIM with the AR system in the mobile environment as the building digital twins. Figure 4 illustrates the concept diagram of the Pinspect system.

The software efficiently uploads key points of geometry coordinates and essential management attributes from BIM software to the cloud, circumventing the need for IFC. This data is then accessible via the mobile ARapp, where pins representing the coordinates are displayed. These pins, embodying both coordinate information and attributes, facilitate the onsite verification of BIM data within the actual building structure. By superimposing these pins onto their real-world counterparts, inspectors can validate the BIM input's precision directly at the construction site. The pins also support the addition of attributes, such as the project's progress status, enabling real-time feedback to the BIM system through cloud technology.

Empirical trials, conducted with the involvement of the first author who was a member of the development team, have shown that this app potentially reduces time spent on construction site management by approximately 70% (M.SOFT Co., Ltd., 2021). In the context of O&M, the pre-emptive identification and AR visualization of critical MEP components via the app can significantly mitigate the risk of oversight during onsite inspections. These pins can serve as the representation of issues recorded during the on-site inspection, which can be easily transferred to CMMS after the hand-over. This case study underscores that even a simplified digital twin, derived from textual data, can be effectively leveraged for building management.

Moreover, the integration of TxBI with CMMS is further streamlined by the conventional practice of importing text data, such as CSV files, into CMMS. This connection establishes a seamless flow of data between TxBI and CMMS.

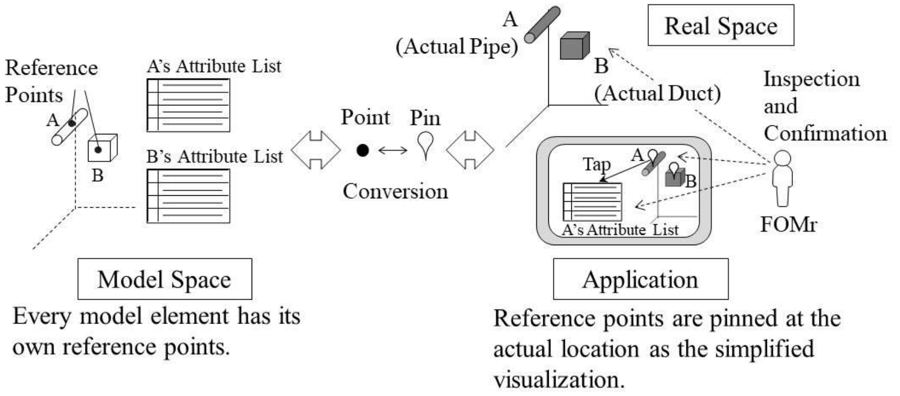


Figure 4. Pinspect system concept diagram.

5. Discussion

TxBI represents a significant advancement in the field of BIM for FM, highlighting the potential for direct text exportation from BIM for use with tools such as CMMS and Realtime Visualization from Text in O&M applications. The system, intentionally simplified to a text-based format, is specifically designed to facilitate frequent usage by the FOMr with minimal technical expertise requirements.

The primary objective of TxBI is to challenge and reevaluate the conventional assumptions set by traditional BIM approaches for FM, thereby promoting a more widespread adoption and utilization of BIM data by the FOMr. This initiative bridges the gap between their O&M expertise and BIM, transitioning from a design and construction-centric model to an information model-focused approach during the maintenance phase. The emphasis here is not on an extensive, comprehensive information source, but rather on providing lighter, more accessible alternatives.

The centralization of BIM attribute management and O&M records within TxBI significantly amplifies the value derived from BIM in FM, particularly through the application of Artificial Intelligence (AI) analysis. This technology has the potential to facilitate comparative assessments of challenges and best practices encountered in various facilities globally against the actual buildings under management. A key application involves linking BIM's coordinate data and component quantities with FOMr's inspection records and response history to enable enhanced O&M strategies. For instance, given a scenario involving critical server equipment adjacent to piping and HVAC units. When a leakage occurs on piping and endangers the server, the detailed actions should be taken by FOMrs in response along with the components' specific coordinate data and quantities, are ultimately recorded and linked in TxBI. AI analysis of this data can then promptly suggest remedial actions for similar occurrences and provide advance warnings of potential O&M issues during the integration of BIM into TxBI, based on coordinates and quantities. Additionally, leveraging insights from facilities where exemplary management practices have optimized Life Cycle Costs (LCC) could facilitate recommendations for optimal equipment placement and management strategies, informed by similar coordinate and quantity data. This approach not only fosters improvements and training but also could appropriately recognize and motivates FOMrs who demonstrate exceptional performance, thus encouraging superior outcomes.

When the utilization of BIM in numerous buildings is further increased, even marginally, and its economic benefits are widely recognized, appropriate incentives for BIM creation and its application in FM may subsequently emerge. Therefore, it is imperative to develop systems tailored to the needs of FOMrs who actively engage with BIM data in FM. Our advocacy should extend beyond promoting the utility of BIM; it should also involve actively incorporating the perspectives of FOMrs, fostering collaborative dialogues about the practical applications of BIM in FM.

6. Limitation

It is acknowledged that the scope of information required may vary depending on specific cases. Thus, the development of case studies tailored to these varying needs will be essential in the future.

Furthermore, we posit that the BIM for FM in architectural and structural realm necessitates distinct approaches compared to MEP. Simplifying major components such as walls, columns, floors, and beams into text form can present a significant challenge in architectural BIM context. While MEP is composed of a series of simple elements forming a system, the architectural and structural ones show complexity due to their substantial size, the diversity of shapes, and the custom adjustments made onsite, such as components cut to specific dimensions or shapes freely formed with concrete. For detailed polygons, accurately replicating their shape is contingent on knowledge of the number of vertices and the vectors from a central point.

Therefore, this task poses a greater challenge than the simpler modeling typically associated with facilities management. Recent studies have explored the use of voxels for the simplified modeling of BIM (Khan et al., 2023). While voxels, as aggregates of cubes, offer a potentially more straightforward path to text conversion, they also highlight the possibility that architectural and facility information might be represented in disparate formats.

7. Conclusion

This study introduces TxBI, a novel approach that facilitates the direct utilization of BIM for FM without the reliance on intermediary formats. TxBI strategically focuses on connecting with the most critical information necessary for O&M tasks, adopting a text-centric methodology. This approach embodies a concept where, BIM that reflects more words from FOMrs, thus greatly contributes to the longevity and sustainable management of buildings.

References

Altohami, A. B. A., Haron, N. A., & Law, T. H. (2021). Investigating Approaches of Integrating BIM, IoT, and Facility Management for Renovating Existing Buildings: A

Review. *Sustainability: Science Practice and Policy*, 13(7), 3930. https://doi.org/10.3390/su13073930

Ashworth, S., Tucker, M. P., Druhmann, C., & Kassem, M. (2016). Integration of FM Expertise and End User Needs in the BIM Process Using the Employer's Information Requirements (EIR). *Proceedings of CIB World Building Congress 2016*, 942–953. http://researchonline.ljmu.ac.uk/id/eprint/16359/

Carvalho, A. T., Calejo, R., Santos, J., & Maia, L. (2023). BIM model to support O&M tasks. *ITcon*, 28, 322–331. https://doi.org/10.36680/j.itcon.2023.016

Cavka, H. B., Staub-French, S., & Poirier, E. A. (2017). Developing owner information requirements for BIM-enabled project delivery and asset management. *Automation in Construction*, 83, 169–183. https://doi.org/10.1016/j.autcon.2017.08.006

Checkland, P., & Poulter, J. (2006). Learning for Action: A Short Definitive Account of Soft Systems Methodology and Its Use for Practitioners, Teachers and Students. Wiley.

Dhillon, R. K., Jethwa, M., & Rai, H. S. (2014). Extracting building data from BIM with IFC. *International Journal on Recent Trends in Engineering & Technology*, 11(2), 202. http://msysmhsptkazqgnbycne.s3.amazonaws.com/user/WV_1500747159_6570_1.pdf

Kang, T. W., & Hong, C. H. (2015). A study on software architecture for effective BIM/GIS-based facility management data integration. *Automation in Construction*, 54, 25–38. https://doi.org/10.1016/j.autcon.2015.03.019

Katsigarakis, K., Lilis, G., Rovas, D., González-Gerpe, S., Bernardos, S., Cimmino, A., Poveda-Villalón, M., & García-Castro, R. (2022). A Digital Twin Platform generating Knowledge Graphs for construction projects. In *Third International Workshop on Semantic Digital Twins (SeDiT 2022), co-located with the 19th European Semantic Web Conference (ESWC 2022)*, Hersonissos, Greece. CEUR Workshop Proceedings. https://discovery.ucl.ac.uk/id/eprint/10159228/

Khan, M. S., Kim, I. S., & Seo, J. (2023). A boundary and voxel-based 3D geological data management system leveraging BIM and GIS. *International Journal of Applied Earth Observation and Geoinformation*, 118, 103277. https://doi.org/10.1016/j.jag.2023.103277

Kim, K., Kim, H., Kim, W., Kim, C., Kim, J., & Yu, J. (2018). Integration of ifc objects and facility management work information using Semantic Web. *Automation in Construction*, 87, 173–187. https://doi.org/10.1016/j.autcon.2017.12.019

Lai, H., & Deng, X. (2018). Interoperability analysis of IFC-based data exchange between heterogeneous BIM software. *Journal of Civil Engineering and Management*, 24(7), 537–555. https://doi.org/10.3846/jcem.2018.6132

Lin, Y. C., Hsu, Y. T., & Hu, H. T. (2022). BIM model management for BIM-based facility management in buildings. *Advances in Civil Engineering*, 2022, Article ID 1901201, 13 pages. https://doi.org/10.1155/2022/1901201

M.SOFT Co., Ltd. (2021, February 24). Addition of BIM collaboration function to inspection and inspection app 'Pinspect' using AR [in Japanese]. Retrieved from https://www.msoft.co.jp/news/detail-64.html

Pärn, E. A., Edwards, D. J., & Sing, M. C. P. (2017). The building information modelling trajectory in facilities management: A review. *Automation in Construction*, 75, 45–55. https://doi.org/10.1016/j.autcon.2016.12.003

Rogage, K., & Greenwood, D. (2020). Data transfer between digital models of built assets and their operation & maintenance systems. *Journal of Information Technology in Construction*, 25, 469–481. https://doi.org/10.36680/j.itcon.2020.027

Wijeratne, P. U., Gunarathna, C., Yang, R. J., Wu, P., Hampson, K., & Shemery, A. (2023). BIM enabler for facilities management: a review of 33 cases. *International Journal of Construction Management*, 1–10. https://doi.org/10.1080/15623599.2023.2222962

PREDICTING PEDESTRIAN TRAJECTORIES IN ARCHITECTURAL SPACES: A GRAPH NEURAL NETWORK APPROACH

RUNYU YANG[1], WEILI WANG[2] and PENG GUI[3]
[1,2,3]*Glodon Company Limited.*
[1]*691302245@qq.com, 0009-0002-7646-5376*
[2]*2280918070@qq.com, 0009-0005-5059-7903*
[3]*gui.ruipeng@outlook.com, 0009-0001-5820-4423*

Abstract. This paper introduces a graph neural network-based model for predicting pedestrian trajectories in architectural spaces. Compared to traditional simulations based on physics-based models, this data-driven model has a stronger ability to learn and predict pedestrian behaviour patterns from real-world data. The model is pre-trained based on Hongqiao Railway Station Dataset, then trained and tested based on the ETH Dataset and the Stanford Drone Dataset, enabling comparisons with other AI models. By creating a more intelligent model, we can establish a digital replica of the real world that can predict pedestrian flow with higher accuracy in daily life or extreme situations such as sudden fires. Our results underscore the critical role of such models in comprehending how architectural spaces are utilized, and thus in improving architectural design and urban planning.

Keywords. Multi-agent Simulation, Trajectory Prediction, Graph Neural Network, Conditional Variational Autoencoder, Path-finding.

1. Introduction

The prediction of pedestrian trajectories involves using pedestrian features or historical trajectories to anticipate future pedestrian movements. This is a pivotal topic in computer science and social science, and has gained widespread attention and application in areas like industrial robots, autonomous driving, spatial analysis, and design (Bendali-Braham et al., 2021).

Pedestrians are intelligent agents who can naturally navigate and avoid obstacles in complex environments. They possess an inherent theory of mind, which allows them to reason about other people's actions in terms of their mental states (Gweon et al., 2013). Therefore, an effective pedestrian trajectory prediction model must not only perceive the environment but also perceive and anticipate the behaviors of neighboring agents.

There are numerous multi-agent behavior prediction methods, ranging from traditional cellular automata models and social force models to neural network-based regression models, generative models, and more. However, many of these methods mechanistically simplify agent behavior, failing to model agents as entities capable of

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 251-260. © 2024 and published by the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA), Hong Kong.

perception, prediction, and decision-making. Our goal is to implement a model where each agent can perceive its surrounding environment, including obstacles and neighboring agents, and combine temporal features to predict the behavior of other agents before making its own decisions. The model parameters should be learnable, allowing for continuous optimization using real-world data.

2. Related Work

2.1. PHYSICS-BASED APPROACHES

The Social Force Model is a classic physics-based model that describes pedestrian movement. Based on Newtonian mechanics, this model assumes that pedestrians move under the influence of social forces (Helbing et al., 1995). Pedestrians are subjected to three types of forces: self-driven force, interpersonal force, and the force between individuals and obstacles.

Physics-based models align with the intuitive understanding of individual movement, with straightforward modelling and computation processes. However, they simplify pedestrians into particles, ignoring the individual's comprehensive perception capabilities, decision-making abilities, and randomness. Additionally, because the formulas involve many empirical parameters such as the avoidance range, these models lack generalization ability. The simulation of specific scenarios depends on subjective parameter adjustments or experiments, thus the practical application and validation present considerable challenges due to their higher usage threshold.

2.2. CLASSIC MACHINE LEARNING-BASED MODELS

In contrast to physics-based models, classic machine learning-based models analyze historical data to predict future trajectories. Specific methods include Kalman filtering, support vector regression models, hidden Markov models, and more.

The Kalman filtering method can account for the randomness of the motion process. The individual's uncertainty can be modelled by the Gaussian distribution (Narayanan et al., 2021). In each time step, the hidden states are first calculated and updated, and the mean and covariance matrix of the states in the next time step are predicted to represent the uncertain pedestrian movement.

Support Vector Regression (SVR) can learn the behaviour of agents in complex environments. It uses a kernel function to project input data into high-dimensional space, conducts regression prediction, finds the most probable movement strategy, and thus predicts future trajectories.

2.3. DEEP LEARNING-BASED MODELS

Classic machine learning-based methods are suitable for simple prediction scenarios and short-term prediction tasks. However, in recent years, deep learning-based trajectory prediction methods have become increasingly popular. These models consider various factors, such as physical factors, environmental factors, and the influence of neighbouring agents, and can be continually optimized through training. By using transfer learning, they can adapt to diverse scenarios.

Sequential models are capable of extracting historical trajectory features. It stores and updates information in hidden layers to predict the output. Alahi et al. (2016) proposed the Social LSTM for predicting pedestrian trajectories in crowded spaces. The attention mechanism is increasingly being employed in trajectory prediction tasks. Messaoud et al. (2020) used a multi-head attention mechanism to emphasize the impact of adjacent vehicles, achieving competitive trajectory prediction performance.

Graph Neural Networks (GNNs) rely on message aggregation processes for prediction. Agents in the environment are represented as nodes, and relationships between agents are represented as edges. GNNs excel at aggregating spatial information and handling dynamic data, making them well-suited for complex trajectory scenarios. Shi et al. (2021) introduced SGCN for pedestrian trajectory prediction. The results demonstrate that it effectively captures adaptive interactions between pedestrians and their motion tendencies.

Generative models can be utilized for generating diverse trajectories. Notably, some of these methods have shifted their focus from predicting trajectories to predicting the distribution of trajectories. This shift better supports downstream tasks such as trajectory matching. Dendorfer et al. (2020) presented Goal-GAN, an end-to-end trajectory prediction model that generates diverse trajectories consistent with physical constraints. Xu et al. (2022) introduced SocialVAE to capture the uncertainty and multi-modality of human navigation decision-making.

3. Methodology

Our objective is to predict pedestrian trajectories based on their features and historical trajectories within built environments.

- Pedestrian trajectories represent temporal data which can be efficiently encoded by sequential models.
- Architectural spaces often contain various obstacles, such as walls and atriums. Pedestrians have the ability to perceive their surroundings, so an environment encoder can be used to gather spatial information.
- Pedestrians also have the ability to perceive and anticipate the actions of other individuals. Therefore, neighbour information should be aggregated through message aggregators in graph network layers.
- To learn the distribution of pedestrian features from real-world data, we introduce a feature generation model.

In summary, we present a hybrid trajectory prediction and generation model that consists of Temporal Convolutional Networks (TCN) for trajectory encoding, CNNs for environmental encoding, graph message aggregators for compiling neighbour information, and Conditional Variational Autoencoder (CVAE) for synthesizing pedestrian features. In the output section, we employ a Multilayer Perceptron (MLP) to merge the features obtained earlier and to predict the mean and variance of the trajectories. This model is named the Graph Temporal Convolutional Network (GTCN). We optimize the predicted values to match the ground truth by minimizing the negative log likelihood of the Gaussian likelihood function.

3.1. TRAJECTORY ENCODER

Trajectory is a type of temporal sequence data. We represent the pedestrian location at a given moment with two-dimensional coordinates. The sampling interval for the trajectory is denoted as t_interval, and the interval length is denoted as interval_length. Note that as time progresses, the interval_length gradually increases. Therefore, we capture the data up to a maximum time duration of max_length. Consequently, for n pedestrians, the shape of the trajectory data is (n, max_length, 2). It is important to recognize that the historical trajectory lengths of pedestrians do not always equate to max_length, necessitating the use of a variable-length temporal encoder.

Temporal models in deep learning include LSTM, TCN, Transformer, and others. LSTM introduces gated units that can store and forget memories, making it well-suited for processing temporal information. However, it calculates by the sequence order, which is not easily parallelizable, making it slow for long sequence data. The Transformer does not use recursive structures, allowing for parallelization to speed up computations and effectively managing long-distance dependency information. However, it has a large number of model parameters, requiring extensive training data and computational resources, and it is prone to overfitting on short sequences.

TCN is a variant of CNN that processes temporal data through convolution. It can compute in parallel, flexibly adjust its receptive field based on the size of the convolutional kernel to acquire information from different temporal scales, and with fewer model parameters, it is easier to train and less likely to overfit, making it suitable for encoding trajectory information with lengths from single to tens of digits. Therefore, we choose TCN as our trajectory encoder.

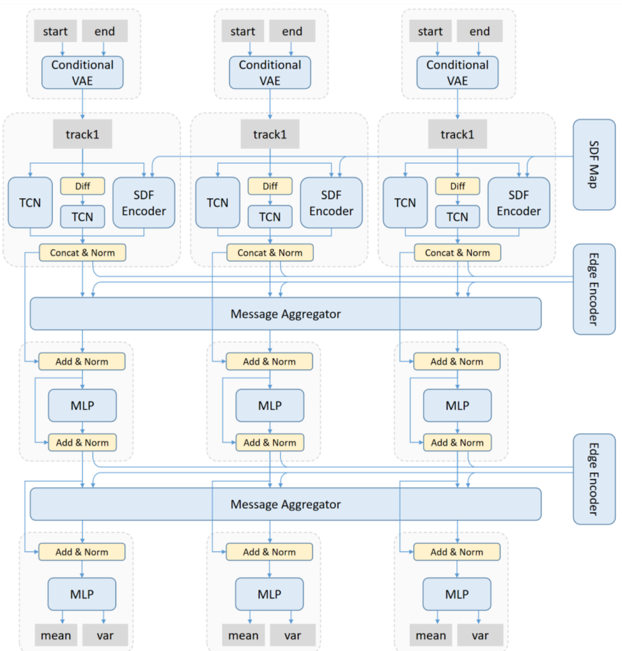


Figure 1. Model structure of GTCN.

3.2. MAP ENCODER

The architectural spaces are complex and diverse, encompassing walls, atria, ramps, furniture, and more. Pedestrians can perceive these obstacles and adjust their movement strategies accordingly. Therefore, it's crucial to encode obstacle information.

The Signed Distance Field (SDF) is commonly used in computer graphics. It represents a distance field where each point stores the distance to the nearest object. In our case, we consider all inaccessible areas within the space as obstacles and compute the Signed Distance Field of that space, denoted as SDF_space, to represent the global environmental information. For the k-th pedestrian, we extract a surrounding fixed-size SDF_k to represent the local environmental information. We then input SDF_k into a map encoder. We use a CNN as the encoder, utilizing a five-layer structure of Conv2d - ReLU - MaxPool2d for convolution and pooling. This is followed by a fully connected layer, which outputs a 64-dimensional vector to serve as the environmental encoding.

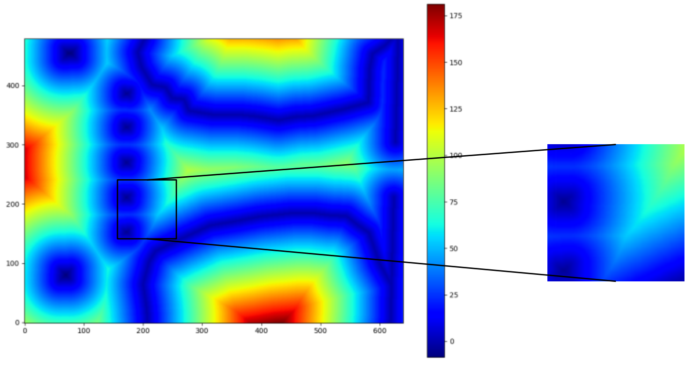


Figure 2. Left: global Signed Distance Field. Right: local Signed Distance Field.

3.3. GRAPH NEURAL NETWORK

Pedestrians as sentient beings, are capable of perceiving and predicting the movements of nearby pedestrians. To avoid collisions or catch up with companions, they adjust their movements accordingly. Therefore, we introduce Message Aggregators to gather neighbourhood information.

The perception range of a pedestrian is relatively fixed, but the number of other pedestrians within this range constantly changes. This requires a model that can handle complex non-Euclidean data. Graph Neural Networks (GNNs) are particularly suitable for this task. We represent the features of agents as nodes and compute the distance between each pair of nodes. We then assign an edge to every pair of nodes that are within a certain distance threshold. Nodes and edges are fed into a Graph Convolutional Network layer for aggregation. The number of layers in the graph is a significant hyper-parameter. Through experimentation, we found that 1 to 2 layers of GCN can effectively aggregate neighbour information. In contrast, more than 3 layers, although capable of collecting long-distance information, may lead to over-smoothing. Therefore, we set the number of layers to 2.

3.4. CONDITIONAL VARIATIONAL AUTOENCODER

When applying a trajectory prediction model to the pedestrian flow simulation within architectural spaces, a key issue is initializing pedestrian features. In common trajectory prediction tasks, the initial features of an object, namely its historical trajectory, serve as inputs to the model. However, in the task of predicting flow within a building space, we typically use path-finding algorithms to determine the starts and goals of the pedestrians. Additional features, such as historical trajectory, body gait and stride, need to be manually set or generated out of thin air. By learning the distribution of pedestrian features in dataset, we can generate reasonable feature while avoiding complex manual inputs. The starts and goals can be used as conditions to guide the feature initialization for the generative model.

Common generative models include Variational Autoencoders (VAE), Generative Adversarial Networks (GAN), and Diffusion Models. GAN consists of a generator (G) and a discriminator (D), which iteratively optimize during the training process to generate highly convincing data. However, GANs are difficult to train and can fall into producing overly limited samples. Diffusion Models have recently become a popular generative model, especially in the field of image generation. They can produce high-quality samples but require a considerable number of time steps in the diffusion process, leading to high computational time and resource requirements.

VAE compresses data into latent vectors through an encoder and then uses a decoder to generate data from the latent vector. It optimizes the reconstruction accuracy by comparing the original data with the reconstructed data to minimize the Mean Squared Error (MSE) loss and generate data from noise by pushing the latent vector towards a standard normal distribution to minimize the Kullback–Leibler divergence (KL) loss. VAE has a low computational cost and are suitable for learning continuous and smooth features, thus is able to generate structurally simple pedestrian features. Further, by concatenating generation conditions to the latent vector, we can achieve a conditional generation model.

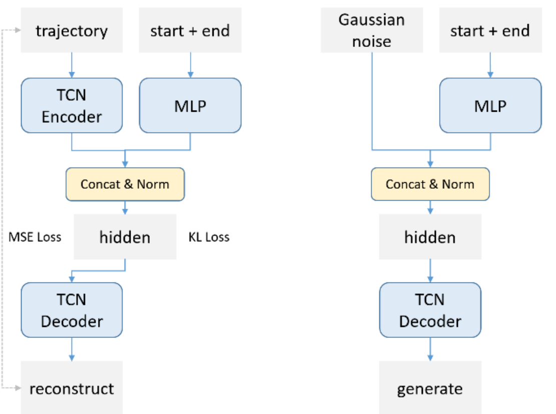


Figure 3. Left: training of Conditional Variational Autoencoder. MSE loss will be calculated between label trajectory and reconstructed trajectory. KL loss will be calculated in hidden states. Right: generate trajectory data from start location, end location and Gaussian noise.

4. Experiment

4.1. SHORT-TERM PREDICTION ON VIDEOS

To visually assess the efficacy of our model, we conducted an experiment using videos. We captured a video of a pedestrian intersection at Shanghai Hongqiao Railway Station, spanning a duration of 90 minutes with a frame rate of 30 and containing a total of 162,000 frames. We employed YOLOX (Ge et al., 2021) as the object detector to detect pedestrians and ByteTrack (Zhang et al., 2022) as the object tracker to obtain continuous pedestrian trajectories.

For each pedestrian at each moment, the object detector provided a four-dimensional dataset containing the top-left corner coordinates, as well as the width and height of the detection box. In this case, we not only predicted the coordinates but also the width and height of the detection box. We considered a maximum historical length of 8 frames, which served as input to predict the trajectory for the next 12 frames. We divided all 162,000 frame data into 8100 groups with 20 frames per group, further dividing them into 80% training set, 10% validation set, and 10% test set. We employed Average Displacement Error (ADE) to calculate the prediction error (Yue et al., 2022). ADE was calculated as the L2 error between a predicted trajectory and the ground truth. We used the mean of the predictions as the trajectory prediction value and visualized the predicted variance as the thickness of the detection box. We used a batch size of 128, a learning rate of 0.005, and an Adam optimizer. After about 50 epochs, ADE was reduced to about 4.1 pixels. As a comparison, when using Kalman filtering as prediction model, the ADE is approximately 7 pixels, which shows that the AI model performs much better than the traditional model.

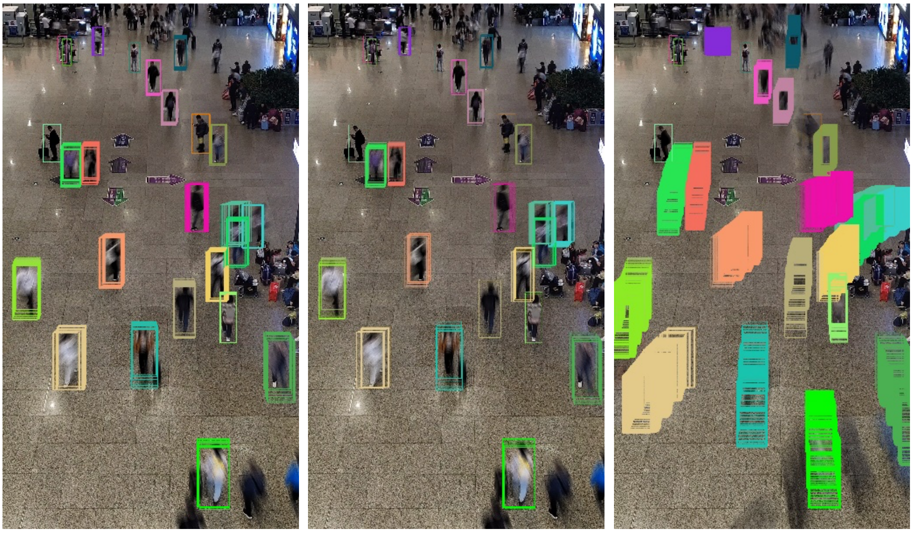


Figure 4. Left: the ground truth of detection box trajectory, with a length of 12 frames, corresponding to 0.4 seconds. Middle: the predicted detection box trajectory with a length of 12 frames. The predicted values are very close to the ground truth. Right: the long-term predicted trajectory obtained by iterating the prediction results, with a length of 60 frames, corresponding to 2 seconds.

4.2. COMPARISON ON OPEN SOURCE DATASETS

To ensure the versatility of our trajectory prediction model, we pretrained it on the Hongqiao Railway Station Dataset and fine-tuned it on the ETH Dataset (Pellegrini et al., 2009) and the Stanford Drone Dataset (Robicquet et al., 2016). This approach enabled the model to adapt to different scenarios.

We compared our model with S-GAN (Gupta et al., 2018) , Sophie (Sadeghian et al., 2019) , Trajectron++ (Salzmann et al., 2020) , SGCN (Shi et al., 2021) on the ETH and SDD datasets. Note that the Social Force Model is a traditional physics-based model that requires manual parameter settings and destination settings, which are not included in the dataset. Therefore, it cannot be directly compared with other trajectory models. The test metrics were ADE and FDE, and the test results are summarized in the following table.

Table1. Results on ETH/UCY and SDD. Our model outperforms baseline methods in both ADE and FDE.

Methods	Metrics	ETH	Hotel	ZARA1	ZARA2	SDD
S-GAN	ADE	0.82	0.70	0.33	0.42	27.23
	FDE	1.52	1.63	0.69	0.81	42.60
Sophie	ADE	0.70	0.76	0.31	0.38	17.21
	FDE	1.40	1.68	0.63	0.80	29.34
Trajectron++	ADE	0.71	0.22	0.30	0.23	17.90
	FDE	1.68	0.46	0.78	0.60	30.83
SGCN	ADE	0.87	0.68	0.33	0.42	30.31
	FDE	1.62	1.38	0.69	0.83	41.76
GTCN (Ours)	ADE	0.68	0.23	0.29	0.21	22.81
	FDE	1.49	0.44	0.63	0.69	34.43

4.3. MODEL APPLICATION IN ARCHITECTURAL SPACES

The simulation of pedestrian movement in architectural spaces typically involves two main components: global path-finding for each agent to reach their destination, and local motion modeling to simulate agent movement. Path-finding typically utilizes heuristic algorithms such as A*, while the motion modeling aspect uses a motion model at the local level (Mashhadawi, 2016). The building traffic system encompasses both horizontal and vertical traffic elements. Horizontal traffic includes floors, carpets, and other walking areas, while vertical traffic includes escalators, elevators, and evacuation stairs.

To model the reachable building space for pedestrian simulations, we represent it as a graph. Each floor that is reachable is divided into grids, and each grid serves as a node in the graph. Edges are established between adjacent nodes. Then, we randomly distribute or manually input the initial positions and destinations of pedestrians, use path-finding algorithms to determine the path for each pedestrian. We utilize the trajectory prediction model as a motion model to simulate the local movement of the crowd. For each newly generated agent, we consider their starting point and destination, generate initial features using the CVAE model, and input them into the trajectory prediction model to predict their movement. The trajectory prediction model predicts the movement of each pedestrian for 12 frames, approximately 0.4 seconds, and then updates their position and trajectory in a loop until they reach the exit. After calculation, approximately 780 updates and 312 seconds later, all pedestrians have exited the building. We accumulate the positions of all pedestrians over time to create a spatial distribution heatmap. This heatmap allows us to quantitatively identify crowded areas and optimize the design accordingly.

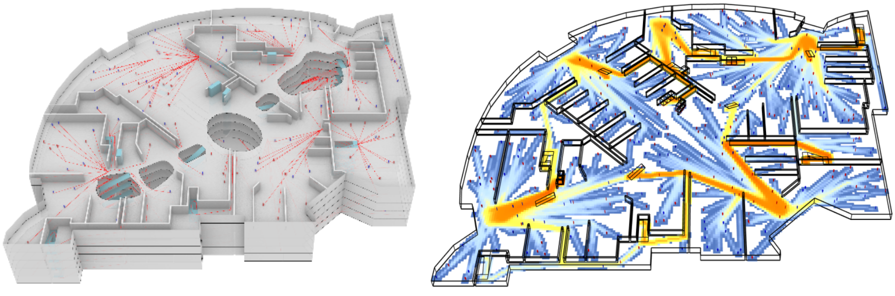


Figure 5. Left: Path-finding algorithms. The red lines represents the route. Right: Representation of the simulation results using a heatmap. The red areas indicate potential congestion zones.

5. Conclusion

Pedestrians are intelligent individuals, with the ability to perceive crowds and environments and predict the behavior of others. Pedestrians can move regularly, but may also change their intentions to make unpredictable movements. Therefore, physics-based models, such as Social Force Model, are difficult to simulate diverse pedestrian behaviors. We have built a novel AI model, called GTCN, to predict pedestrian trajectories in architectural spaces. The model comprehensively considers the perception and prediction abilities of pedestrians, as well as the diversity of pedestrian characteristics.

Video experiments have shown that GTCN can simulate trajectories with higher accuracy than traditional Kalman filters. Experiments on open-source dataset have shown that GTCN outperforms other AI models. We applied the model in architectural spaces to predict pedestrian distribution and movement. Through quantitative analysis, we obtained indicators such as spatial distribution heatmaps. Based on visual results, architects can identify design vulnerabilities such as congestion or accident-prone areas, to optimize building efficiency or improve safety.

References

Bendali-Braham, M., Weber, J., Forestier, G., Idoumghar, L., & Muller, P. A. (2021). Recent trends in crowd analysis: A review. Machine Learning with Applications, 4, 100023.

Gweon, H., & Saxe, R. (2013). Developmental cognitive neuroscience of theory of mind. Neural circuit development and function in the brain, 3, 367-377.

Helbing, D., & Molnar, P. (1995). Social force model for pedestrian dynamics. Physical review E, 51(5), 4282.

Narayanan, S., Moslemi, R., Pittaluga, F., Liu, B., & Chandraker, M. (2021). Divide-and-conquer for lane-aware diverse trajectory prediction. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 15799-15808).

Alahi, A., Goel, K., Ramanathan, V., Robicquet, A., Fei-Fei, L., & Savarese, S. (2016). Social lstm: Human trajectory prediction in crowded spaces. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 961-971).

Messaoud, K., Yahiaoui, I., Verroust-Blondet, A., & Nashashibi, F. (2020). Attention based vehicle trajectory prediction. IEEE Transactions on Intelligent Vehicles, 6(1), 175-185.

Shi, L., Wang, L., Long, C., Zhou, S., Zhou, M., Niu, Z., & Hua, G. (2021). SGCN: Sparse graph convolution network for pedestrian trajectory prediction. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 8994-9003).

Dendorfer, P., Osep, A., & Leal-Taixé, L. (2020). Goal-gan: Multimodal trajectory prediction based on goal position estimation. In Proceedings of the Asian Conference on Computer Vision.

Xu, P., Hayet, J. B., & Karamouzas, I. (2022, October). Socialvae: Human trajectory prediction using timewise latents. In European Conference on Computer Vision (pp. 511-528). Cham: Springer Nature Switzerland.

Ge, Z., Liu, S., Wang, F., Li, Z., & Sun, J. (2021). Yolox: Exceeding yolo series in 2021. arXiv preprint arXiv:2107.08430.

Zhang, Y., Sun, P., Jiang, Y., Yu, D., Weng, F., Yuan, Z., ... & Wang, X. (2022, October). Bytetrack: Multi-object tracking by associating every detection box. In European Conference on Computer Vision (pp. 1-21). Cham: Springer Nature Switzerland.

Yue, J., Manocha, D., & Wang, H. (2022, October). Human trajectory prediction via neural social physics. In European Conference on Computer Vision (pp. 376-394). Cham: Springer Nature Switzerland.

Mashhadawi, Mohammad. "MassMotion evacuation model validation." LUTVDG/TVBB (2016).

Pellegrini, S., Ess, A., Schindler, K., & Van Gool, L. (2009, September). You'll never walk alone: Modeling social behavior for multi-target tracking. In 2009 IEEE 12th international conference on computer vision (pp. 261-268). IEEE.

Robicquet, A., Sadeghian, A., Alahi, A., & Savarese, S. (2016). Learning social etiquette: Human trajectory understanding in crowded scenes. In Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part VIII 14 (pp. 549-565). Springer International Publishing.

Gupta, A., Johnson, J., Fei-Fei, L., Savarese, S., & Alahi, A. (2018). Social gan: Socially acceptable trajectories with generative adversarial networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2255-2264).

Sadeghian, A., Kosaraju, V., Sadeghian, A., Hirose, N., Rezatofighi, H., & Savarese, S. (2019). Sophie: An attentive gan for predicting paths compliant to social and physical constraints. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 1349-1358).

Salzmann, T., Ivanovic, B., Chakravarty, P., & Pavone, M. (2020). Trajectron++: Dynamically-feasible trajectory forecasting with heterogeneous data. In Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XVIII 16 (pp. 683-700). Springer International Publishing.

REVIEW ON THE USE OF CONVERSATIONAL AI NPC AVATARS IN TEACHING AND LEARNING BIM

A preliminary observation of its introduction in a built environment related course in Singapore

JUNIARTO HADIATMADJA[1]
[1]Temasek Polytechnic, Singapore
[1]hadiatmadja_juniarto@tp.edu.sg, 0009-0002-3124-3788

Abstract. Abilities in the use of BIM are critically needed in many industries but there are major challenges to current BIM training. It is unrealistic to assume that the current predominantly teacher directive mode of BIM training is sufficient or responsive enough to tackle rapidly changing challenges and cater to individual pursuits. This article reviews the findings of a research deploying conversational AI NPC avatars and BIM models in a game engine environment as a complementary learning tool that is non-directive and more enquiry based in nature. Enabling learners to autonomously converse and spatially direct the avatar movements to parts of the BIM model they wish to focus on. This article answers some ways the use of AI NPC avatars could benefit the learning of students that are newly introduced to BIM. The research compares tangible results as well as the learner's perceptions toward the use AI NPC avatars. The findings shed light on the technology's current utility and limitations in various aspects of the current topic. Some directions for development of future related research will be also be discussed.

Keywords. BIM Training, Conversational AI NPC Avatars, Game Engine Environment, Individual Enquiry, Learning Tools.

Introduction

1.1. PROBLEMS OF CURRENT BIM TRAINING AND MAIN QUESTIONS, OBJECTIVES AND APPROACHES OF THE RESEARCH PAPER

Knowledge and Abilities in the use of Building Information Modelling (BIM) is in high demand in various industries (Yakami et al., 2018). The fields in which the skills can be deployed have grown beyond its traditional use in AEC (Architecture, Engineering & Construction) related fields. It now ranges even to its possible utility in more explorative sectors such as for creation of virtual simulation environments in the computer gaming industry (Cousins, 2019).

The great potential above is in contrast with the challenges faced in BIM training.

This can range from the perennial problem of not enough teaching time to the efficacy of the predominant one-way directive approach to training. The increasing complexity of BIM would make attainment of sufficient BIM knowledge and abilities through the limited subject teaching time of formal education increasingly unattainable. The challenges are compounded by the limited appeal of the visual styles available to show the models in current BIM software. Visual interfaces that generally do not sufficiently show in real time the true rendered state of the model that could have increased the viewer's comprehension (Huang et al., 2017).

At the same time, the predominantly directive teaching approach to BIM could be further improved. To some extent, this has already started with the emerging popularity of online video tutorials that enable individual pace of learning. At the same time videos visually focuses attention on the learning topic up to a step-by-step basis. This enables videos to be more easily followed and visualised compared to static medium such as books or slides (Tsai et al., 2019). However, the directive teaching is still dependent on the teaching topic delivered by the teacher and not the specific area of interest of the individual learner.

Another limitation of directive teaching is that the main feedback given to models created by learners in their training is the teacher's observation of the model. There is a need for more mechanisms that enable opportunity for the students to actively analyse the efficacy of specific parts of their model on their own. An enquiry-based approach.

At the same time there is growing evidence of the emerging research on the combined utility of BIM and AI (Ozturk et al., 2021). This research studies the possible use of a particular introduction of AI technology in the teaching and learning of BIM. One that can potentially help overcome some of the current problems described above. This research is on the possible benefits of conversational AI NPC avatars in teaching and learning BIM.

The current research wishes to specifically address two possible important questions to gauge the possible benefit of deploying the technology. Firstly, the research questions the conversational AI NPC avatar's potential benefit as a tool to assist students in the making of a BIM model. Secondly, the research questions the conversational AI NPC avatar's potential benefit as a tool to help students evaluate and correct their model.

This early research on introduction of the avatar was done with an objective of attaining some modest benefit in which to build further from and avoid becoming an unwanted hindrance. Therefore, the students were invited to interact with the avatar voluntarily and on their preferred time. The focus of the two questions is meant to cover the role of the avatar in the initial as well as latter stages of the student's learning.

1.2. EMERGENCE OF GAME ENGINE ENVIRONMENTS AND THE RISE OF CONVERSATIONAL AI NPC AVATARS

Research has been done that shows that the viewer's comprehension is improved by enabling a more realistic view of the model to aid the viewer's comprehension. This can be seen in examples such as the use of visualisation of BIM in specific facilities Emerging synchronised integration features between BIM and architectural visualisation software has greatly supported this (Huang et al., 2017).

Within the range of digital platforms available that enables better visualisation, there is even higher utility potential when game engines are used. Beyond visualisation, game engines create simulation environments that enable the learner to analyse the BIM model in greater depth and realism (Cousins, 2019).

A basic example of expanded simulation ability comes from the game engines feature of enabling the player collision settings of the BIM model's building elements. These settings enable the game engine users (formally termed as players) to simulate more realistic situations. These can range from inhibiting movement through solid walls, falling through holes due to gravity or enabling movement to different floor levels by stepping on stairs and ramps.

The development of Non-Playing Characters (NPCs) enables another interactive element of the game engine environment. The NPC avatar has the basic ability to manoeuvre horizontally as well as through stairs or ramps to different levels of a building similar game player. In contrast to the player avatars, the NPCs manoeuvre autonomously and is not controlled by the player.

The appearance of the avatar could also be customised by retargeting the game engine environment's default robotic mannequin avatar with an avatar whose appearance are more human, with clothing and animation that the learner prefers. Thus, providing a more realistic human like avatar that the player of the game environment can interact with.

The potential abilities of NPCs have more recently incorporated the additional abilities to converse in an autonomous (non-scripted) manner through the use of generative AI technology (Moltenbrey, 2023). Thus, becoming a conversational AI NPC avatar. The common knowledge bank of this avatar can be expanded with new additional information provided by the game developer and a customised voice type. Therefore, a unique customised NPC can be created to suit different needs.

Another important ability of conversational AI NPC avatars is the ability to move in the game engine environment in response to the player's verbal command. The ability of the AI NPC avatar to move towards different parts of the building will be leveraged as a key capability that differentiates it from other technologies. Autonomous manoeuvrability enables the avatar to assist in a way which is far different from a static chatbot that does not interact with the digital spatial environment.

2. Methodology

2.1. RESEARCH CONTEXT

The research was conducted on students who were learning BIM in the second year of a built environment related course in Singapore. The main mode of study of the subject in which the student learned BIM is termed as online self-paced learning. As part of the current BIM curriculum developed over the years, at the beginning of the subject students are given access to online learning videos developed by teaching staff that fits the needs of the course. The videos are sequentially accessible to assist students to do specific assignments that incrementally develop skills and knowledge. The student's further developments after the initial assignment are encouraged to be achieved by self-sourcing of external materials such as software documentation and online videos.

At the time of research, 68 students divided into three classes were in attendance. For this research, two classes of the cohort were introduced to the additional use of conversational AI NPC avatars. This consists of 45 students. The students can converse with the avatar and ask the avatar to move to different parts of a BIM model that is set up in a game engine environment. The students can ask the avatar the dimensions, material and other characteristics of the model that the avatar has in its knowledge bank. The students can go to any area of the building and enquire on demand. A departure from the sequential flow of information provided by the video.

In contrast to this, a class that will not be introduced to the avatars will be used as a control group. This comprised of 23 students. These students will only learn using tutorial videos for the same assignments.

This research looks specifically into two different deployment scenarios in which the potential use of the AI NPC avatars could be observed. The scenarios cover the first and second assignment that students do in the process of learning BIM.

Research on the first assignment is related to the first research question. This focuses on the utility of AI NPC Avatars to assist in the making of BIM models. The assignment requires students to follow steps in the video to learn basic commands and features of BIM as they make a designated basic model. The AI NPC avatar is used as an additional learning tool that provides information that complements the main tutorial video used to learn basic BIM software commands.

Correspondingly, research on the second assignment is related to the second research question. It relates to the utility of AI NPC Avatars in the evaluation of BIM models. In the second assignment, the students further explore BIM while designing their own customised BIM model according to their interest. The students embark on an assignment to create a model they design themselves with focus on more advanced forms that they aspire to model. This requires them to independently source for learning material to seek procedures of the BIM software that could help them achieve the forms they wish to achieve. The lecturer provides support by discussing learning strategies and possible resources that students could pursue.

2.2. DATA COLLECTION AND METHODS OF OBSERVATIONS

Potential effects of deploying the AI NPC avatars were studied through tangible outcomes as well as input from the student's perception. For assignment one, a simple tangible comparison of the average marks of students from classes where the avatar was deployed will be compared to the results of students from the control class. For the second assignment, tangible results are gained by observing the movements of the avatar to the intended locations that students had instructed.

For both assignments, information was collected on the perception of students from classes where the avatar was deployed. This was done through an online survey. Google Forms was chosen as a simple platform that all students could easily access. The students were briefed that the survey would be conducted anonymously. This was done to enable the students to be at ease and encourage them to freely express their feedback. The students were requested to provide information through open-ended questions. The students were asked to explain the specific enquiry topics they asked the avatar, their positive as well as negative perception of their interaction.

The students were also asked if they could provide suggestions for improvement. This was seen as an important input to gain knowledge of the user's main aspirations. The students input are seen as valuable for ongoing development of the research.

2.3. GAME ENGINE ENVIRONMENT SET UP

In line with the curriculum of the students surveyed in the research, Autodesk Revit 2023 and 2024 was the BIM software used. The BIM models used for the research would then be imported into a game engine environment. The game engine used for this research is Unreal Engine version 5.2. (UE5.2).

The avatar creation software used was Character Creator 4 (CC4). This software enables a high degree of customisation of the avatar's form and movement. Once imported into UE5.2 the avatars are used to create either third person player characters or non-playing characters.

UE5.2. and CC4 are also compatible with the platform that can enable NPC avatars to be live linked with conversational AI technology. Care was also given to select a conversational AI technology that enables speech, knowledge and movement of the avatar throughout the game engine environment. In this case, Convai web platform was selected. The overall combination of components integrated into the UE5.2. environment and their roles are summarised in Figure 1.

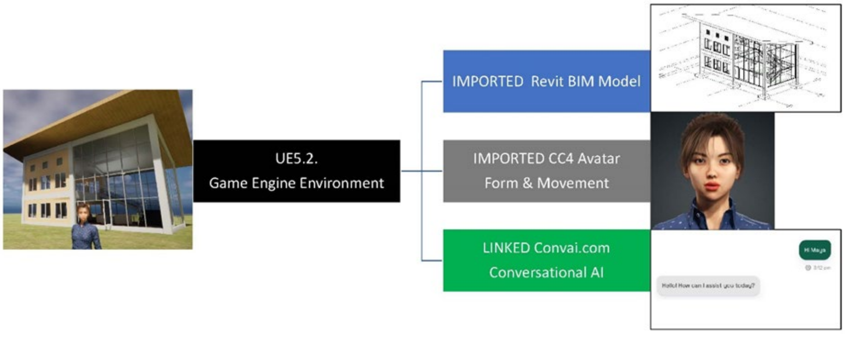


Figure 1. Diagram of components needed to be imported or linked to the game engine environment.

As part of a first-person computer game in UE5.2, the student takes the role of the player and verbally commands the Convai AI NPC avatar to move to parts of the building. The students can ask the avatar to 'lead to', 'move to' or 'go to' a specific location of an object. For example, instructing movement to the 'target' object that has been referenced to the NPC avatar's settings. Alternatively, the avatar can be commanded to 'follow' the player or 'come closer to' the player's position.

The student's commands direct the avatar's movement to parts of the model that the student is keen to investigate. Once the avatar reaches a specific location of the BIM model, the student could also ask the avatar to provide specific information on BIM that it has in its knowledge bank. The lecturer can continually modify and expand the avatar's knowledge bank with new information of the BIM model according to the student's needs.

2.4. DEVELOPMENT OF THE SPECIFIC AVATAR USED IN THE RESEARCH

For this research the avatar is given the name Maya, a young adult Asian female character. The intention was to have an avatar that had a background which students from Singapore could well relate to. The voice, movement, appearance and knowledge bank of the avatar was developed in line with this character. The term 'Maya' in various Asian cultures relates to a 'virtual' element. It is in keeping with the current digital game engine environment context. The Asian character chosen also enables this research to give a more Asian outlook to the perception of students learning BIM. This is in comparison with existing research that studies other regions such as United states, United Kingdom (Shelbourn et al., 2017) or Australia (Casasayas et al., 2021).

2.5. SPECIFIC INFORMATION ON THE DIFFERENT COMPUTER INTERFACES USED

For assignment one, the students in the classes which are assisted by the avatar will need screens to show the avatar's virtual environment, the tutorial video and a screen that shows the BIM model they are making. The avatar's environment and the tutorial video can be hosted as two windows within one computer. This is shown in Figure 2. Alternatively, students can alternate between one window to another if they prefer. Commonly, the student gradually makes the model in another computer.

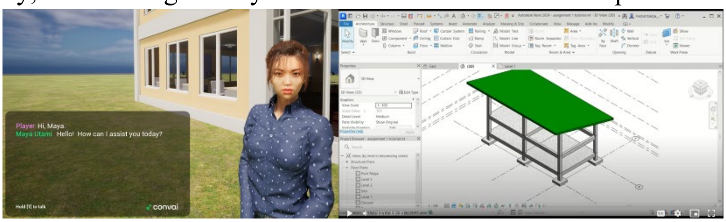


Figure 2. First interface used for assignment one, with two separate windows for the virtual environment and the tutorial video.

As part of the research, an additional sub observation was done during the first assignment to provide an alternative interface. An interface that attaches the video as an object that moves with the avatar as seen in Figure 3. The game environment is arranged so that the student can play, pause and fast forward or backward to parts of the video they wish to focus on. The video object is not given any player collision to avoid obstruction of the avatar's movement in the building. This additional sub element to the research is used to test the student's feedback on the different interfaces offered.

Figure 3. Alternative interface for assignment one with the tutorial video attached to the avatar.

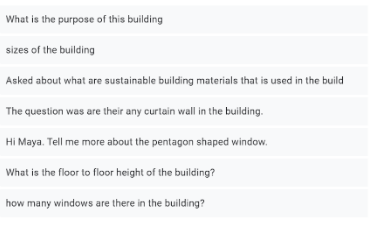
For the second assignment of using the avatar's movement as an evaluation tool, there will be only one screen needed at any one time. The single window is sequentially used from the creation of the BIM model till its set up in the virtual game environment.

3. Results

3.1. ANSWERING THE FIRST RESEARCH QUESTION: RESULT OF THE USE OF THE AVATARS FOR ASSISTING STUDENTS IN MAKING OF A COMMON BASIC BIM MODEL

The class of students whose learning was additionally assisted by the AI NPC avatar on average had a mark that was approximately 6 % above the results of the students in the control class. From this result, currently at least the introduction of the avatar was not detrimental to the student's learning. Specific benefits as well as limitations, are expanded in the following paragraphs based on the observations and feedback.

The student's questions to the avatar mostly focused on clarification of specific parts to be made by in the BIM model. Some samples are shown in Figure 4 and 5.



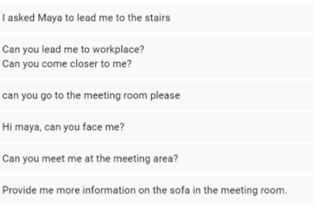


Figure 4 & 5 samples of discussion topics of questions to the avatar.

The students on average stated that the positive interaction they had was that the avatar was able to answer their questions and move to the requested area of the building. They also noted that the main negative interaction they had focused on limitations on verbal communication rather than the movement of the avatar. This involves having to ask more than once, slower than natural response time and overly long answers. When asked to provide suggestions, the students also mainly cited the need for improved verbal communication, more in depth information but concisely stated and a better 'human like' voice.

The students gave feedback that the alternative interface with the video attached to the avatar had visual constraints. Parts of the video were blocked by the building elements. The video becomes overly dark or bright in accordance with the lighting of the space that the avatar is in. These can be seen Figure 6 and 7 correspondingly.


Figure 6 & 7 constraints of assignment one's alternative interface.

3.2. ANSWERING THE SECOND RESEARCH QUESTION: RESULT OF THE USING THE AVATAR FOR ASSISTING STUDENTS TO EVALUATE THE CUSTOM BIM MODEL THAT THEY INDIVIDUALLY DESIGNED.

The avatar's response to the student's verbal command has some inconsistencies. At times the avatar can respond to move upon a request to 'move to' or 'go to' the reference object. On other occasions the avatars move upon receiving a request to 'lead to'. However, the avatar was able to respond to at least one of the three commands.

Once the avatar responds to the request to move, there were no problems for the avatar to horizontally move to a reference object that is located on the same floor level. The avatar was able to identify openings in the building as part of the route to reach the target location. This was achieved autonomously without a specific path instructed.

For movements requiring the avatar to move from one floor level to another through a staircase there were mixed results. For some student models, the avatar was able to climb the stair and reach the intended reference object location on the other floor on the first attempt. On other student models the avatar either only moved to a position on the same floor or went up the stairs until a part that the avatar could not move further towards.

For some models where the avatar initially could not climb up the stairs successfully, some incorrectly sized or placed building parts that obstructed the avatar's movement were identified. Once corrected, the avatars could successfully climb the stairs and progress to the target object. Figure 8, 9,10,11 visually shows this sequence.

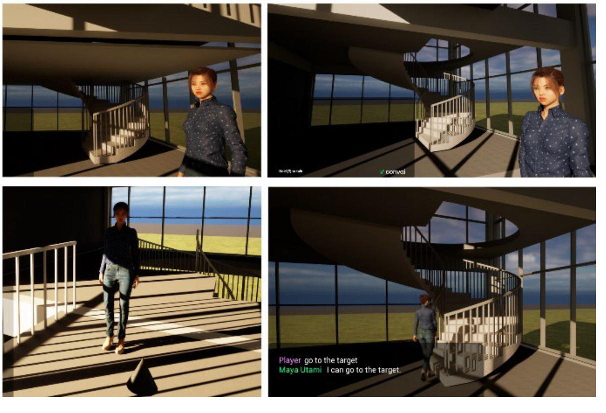


Figure 8 till 11. Images clockwise starting from top left showing removal of obstructing element resulting in successful movement of avatar to the target cone destination.

In the end there remained some models, that even after correction, the avatar could still not climb the stairs completely. It was observed that that these stairs were complex spiral stairs that required the avatar to climb more than one floor level at any one time (e.g., from level one to level 3).

The students gave feedback that the use of the avatar as an evaluation tool was generally useful. They also gave suggestion about possible future development of the avatar beyond just walking to a location. Students suggested development of the avatar's animation to enable it to sit and lie down so that interaction with the furniture (e.g., seats or beds) of the BIM model could be done.

In some models there were more than one route was possible that the avatar could go through to reach the target location. The avatar was able to move using a consistent path of travel to the target on repeated trials. It was observed the avatar travelled paths that were the shorter alternatives. When the starting position of the avatar or the target object was adjusted, the avatar was responsive to use an altered path to reach the target.

4. Discussions

The current research done on the use of conversational AI NPC avatars was conducted on a small sample size. At the same time, the current preliminary research has only investigated some capabilities in which the avatar could be deployed for the teaching and learning of BIM. The discussion of the research findings below should consider these limitations.

4.1. DISCUSSIONS AND SUGGESTIONS FOR USE OF THE AVATAR IN ASSISTING STUDENTS TO MAKE A COMMON BASIC BIM MODEL.

From this initial deployment in assignment one, the students who were assisted by the avatar in general could perceive some benefits of it use. The students showed interest to give feedback suggestions for improvement of the avatar as an additional learning tool.

Improvement of the 'speech comprehension' capability of the conversational AI platform is significantly needed. This is the current major obstacle that students face in communicating more effectively with the avatar. At the same time, the avatars facial movement are not completely in sync with the words spoken. There is also limitation that the knowledge bank of the avatar can only be inputted as text only files. This limits the range of information that could be supplied to the avatar.

The current Convai web platform used also has a substantial lag time. The lag is both in fully capturing the player's command as well as responding to it. Sometimes the students had to repeat the command a few times before it is fully registered. In other times the student had to rephrase their command. The researcher observed that the quality of the communication varied from day to day. This situation also potentially highlights the current difficulty of achieving a consistently strong live link between the game engine and the AI web platform.

4.2. DISCUSSION AND SUGGESTIONS OF THE USE OF THE AVATAR FOR ASSISTING STUDENTS TO EVALUATE THE BIM MODEL THEY INDIVIDUALLY DESIGNED.

The successful use of the avatar to help alert some students of the efficacy of their model is an important finding of this research. The enthusiasm of students who volunteered their model to be tested is also encouraging. This potentially could be used as motivation to self-evaluate and continually improve their BIM skills in the process.

At the same time, existence of some cases where the avatar could not pass through some of the students designs, even after correction were done, informs that there are some limitations of the avatar's movement that needs to be further studied. This also gives a glimpse of the potential for further revelations that could be revealed if this research were to be expanded. Some thoughts are presented in the final section below.

4.3. POTENTIAL FOR FURTHER RESEARCH.

It would be interesting to conduct research to observe if the avatar had a different body form and agility (e.g., of an elderly person, a gown up male or even a small child) would the results be different. Would the movements of these avatar be more limited or less restricted? This also suggests possibilities of expanded scenarios of avatar movements and conversations. Scenarios where there could be interaction between different avatars, different building spaces and different players. More detailed research could also be done on parts of the avatar form or settings that currently inhibit movement or how to expand the range of animated movements possible.

The considerations behind the avatar's ability to use a path leading to the target location even without a clear line of sight to the target location is also worthy to be further studied. This could be done in such terms as complexity of the route to navigate, consistency of movement patterns and the pace of travel. Therefore, the research could combine investigation of both a spatial as well as temporal dimension.

Research is currently ongoing to answer some of these expanded topics. These are planned to be completed in the near future. The expanded topics seem to be very timely to be developed and a very fertile ground for further study.

Reference:

Casasayas O., Reza Hosseini M., D. Edwards D., Shuchi S., Chowdhury M (2021) Integrating BIM in higher education programs: Barriers and remedial solutions in Australia. *Journal of Architectural Engineering.* Volume 27, Issue 1. https://doi.org/10.1061/(ASCE)AE.1943-5568.0000444

Cousins S. (2019, April 14) *Extended reality: it's more than just a game for BIM*. Retrieved December 1, 2023, from https://www.bimplus.co.uk/extended-reality-its-more-just-game-bim/ https://www.archdrawoutsourcing.com/blogs/importance-of-bim-in-the-aec-industry-in-2020/

Huang C.Y., Yien H.W, Chen Y. P., Su Y.C., Lin Y.C. (2017) Developing a BIM-based Visualization and Interactive System for Healthcare Design. In *34th International Symposium on Automation and Robotics in Construction (ISARC 2017)*

Moltenbrey K. (2023. June 1) *Expanding the Role of Non-Player Game Characters*. Retrieved November 30, 2023, from https://gfxspeak.com/featured/my-npc-is-smarter-than-your-npc/

Ozturk G., Tunka M. (2021) Artificial Intelligence in Building Information Modelling Research: Country and Document-Based Citation and Bibliographic Coupling Analysis. *Celal Bayar University Journal of Science.* Volume 16, Issue 3. 269 - 279. DOI: 10.18466/cbayarfbe.770565

Shelbourn M., Macdonald J., McCuen T., and Lee S. (2017). Students' perceptions of BIM education in the higher education sector: A UK and US perspective. Industry and Higher Education. Volume 31, Issue 5. 293–304. https://doi.org/10.1177/0950422217725962

Tsai M., Chen K., 1 Chang Y. (2019). Development of a Project-Based Online Course for BIM Learning. *Sustainability.* Volume 11, 5772; https://doi.org/10.3390/su11205772

Yakami M., Singh V,, Suwal S.(2018) What Do Students and Professionals Think of BIM Competence? *14th IFIP International Conference on Product Lifecycle Management (PLM),*

URBANSIGHT

Application of Digital Twin to Augmented Reality

TZU-HSIN HSIEH [1] and TAY-SHENG JENG [2]
[1]*Department of Computer Science and Information Engineering, National Cheng Kung University*
[1,2]*Department of Architecture, National Cheng Kung University*
[1]*celinehsieh68@gmail.com, 0009-0007-2031-7928*
[2]*tsjeng@mail.ncku.edu.tw*

Abstract. BIM (Building Information Modelling) is currently used to construct building components, which does not provide good assistance for environmental sensing data and user feedback for management and maintenance. There are often problems with timely synchronization of building model information. In order to extend the BIM technology from the building design phase to the facility management phase, a 3D visualization platform is needed to optimize the efficiency of building management by introducing Digital Twin. This study therefore aims to combine Digital Twin developed through building information modelling with augmented reality (AR) headset to achieve automated control of building operation management for future urban governance. In this paper, we use Unity with Autodesk Forge API to build a BIM model with building information in the real site and display it in Microsoft HoloLens 2, which is a wearable device. First, Autodesk Forge API is used to build the monitoring platform of the field model. The data collected by the sensors are transferred to the IoT base station built by Raspberry pi to build the database. Then the digital twin dashboard interface on HoloLens 2 is displayed to control the sensors in the environment. This system breaks the 2D interface of the existing urban management interface and visualizes it in 3D in augmented reality (AR), creating a more intuitive and efficient way to manage smart cities.

Keywords. Digital Twin, Augmented Reality (AR), Internet of Things (IoT), Smart City, Management Platform

1. Introduction

In recent years, the application of Building Information Modelling (BIM) into design and engineering has become increasingly popular. With the aim of incorporating data in building life cycle into 3D computer-aided design (CAD) models of buildings, thereby reducing costs and preventing errors. However, BIM still falls short in updating

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 271-280.

the current state of the building in real time, making it difficult to meet the need for real-time information transformation throughout the building life cycle, from design to construction, operation, and maintenance.

To solve this problem, Digital Twin has emerged as a solution. It creates a dynamic digital model of the physical environment, allowing users to visualize information in a virtual environment. This study integrates digital twins into augmented reality (AR) using head-mounted devices for real-time synchronization. The goal is to provide intuitive management platforms for building administrators and urban decision makers. Digital twins ensure up-to-date data availability throughout the building life cycle. The study analyses the feasibility and challenges of integrating digital twins and augmented reality (AR) in smart cities. Its objective is to explore the strategic development of digital twins in augmented reality (AR) for future smart city management.

2. Related Work

2.1. DIGITAL TWIN CONCEPT AND DEVELOPMENT

Building design and management tools have been developed for decades, starting with computer-aided design (CAD). The term "Building Information Modelling (BIM)" was first coined in the early 1990s (van Nederveen and Tolman, 1992) and has attracted widespread attention and application in the past. Each of these phases can be divided into overlapping information layers that require an effective information exchange strategy to achieve interoperability across all lifecycle phases (Deng et al., 2021).

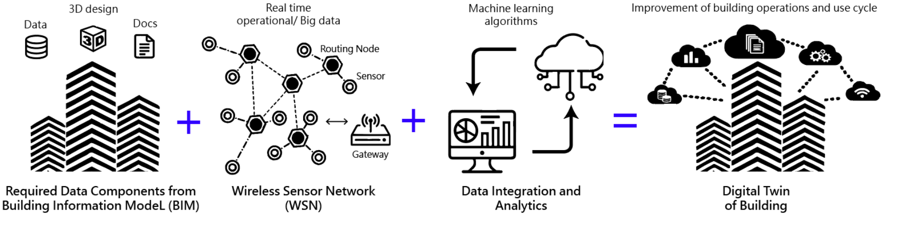


Figure 1. Basic components for creating a digital twin (revised from Khajavi et al., 2019)

However, BIM itself usually provides only static data of the built environment and cannot automatically update the real-time information in the model. With the emergence of the Internet of Things (IoT), defined as the interconnection of sensing devices capable of providing information exchange across different platforms (Gubbi et al., 2013), the integration of sensing data deployed in the environment and data provided by BIM models in real time has led to the concept of Digital Twin. Digital Twin aims to synchronize real-world and virtual platforms to control facility management, environmental monitoring, and other lifecycle processes in the built environment, connecting data to bring the virtual and real worlds together. Digital twins in buildings can use a variety of sensors and the Internet of Things (IoT) to create a real-time view (Figure 1). This dynamic view enables real-time analysis, informed decision making, building efficiency and comfort enhancement.

2.2. AUGMENTED REALITY APPLICATIONS

Digital twins replicate the physical world in a virtual environment, while augmented reality (AR) merges virtual elements with the real-world using display screens like cell phones or glasses. Augmented reality (AR) enhances the user's perception of the real environment by incorporating computer-generated displays, sounds, text, and effects (Dunleavy and Dede, 2014). It provides information that may not be directly perceivable in the physical world. In Figure 2, four rapidly evolving technologies are summarized that correspond to the four stages required to develop augmented reality applications (Loijens et al., 2017). AR integrates the virtual and real worlds through technologies such as cloud computing, location tracking, mobile devices, and user interfaces. Augmented reality also finds applications in the AEC industry, where it is used for tasks like construction site monitoring. For instance, Zaher et al. developed a system that combines BIM and AR on smartphones to monitor construction progress, leveraging the portability and visualization capabilities of mobile devices in construction projects (Zaher et al., 2018).

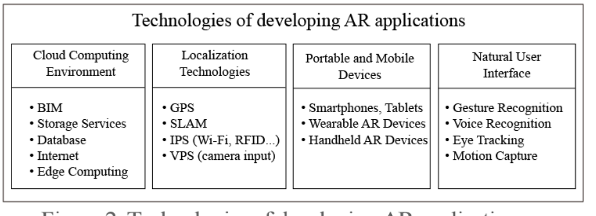


Figure 2. Technologies of developing AR applications

2.3. RELEVANCE

Digital Twin solves the problem of updating model information and status in real time for BIM. However, the lack of intuitive and interactive user interface has been described in some literature as difficult to understand and use. On the other hand, Augmented Reality (AR), while not creating a virtual world in sync with the real world like Digital Twin, enhances the real world by visualizing data and information. For example, to provide information to assist in construction, or maintenance management, allowing for a smoother flow through all phases of the building lifecycle. Therefore, this study combines digital twins with augmented reality. Digital twins enhance augmented reality systems by accurately linking to remote sensors, helping to pinpoint the location of users and nearby objects. Augmented reality, in turn, adds an interactive and user-friendly layer to Digital Twin, making it easier for users to comprehend and utilize the information effectively for better decision-making and maintenance.

3. UrbanSight

3.1. SYSTEM DESIGN

To validate the digital twin application for building management, this study focuses on the system's development architecture and experimental construction method. As depicted in Figure 3, the system comprises two primary parts: the back-end (web platform) and the front-end (Unity platform). The front-end uses the collected data from the back-end to build a managed user interface, which is deployed in the Microsoft HoloLens 2 headset for managers to monitor the site more intuitively.

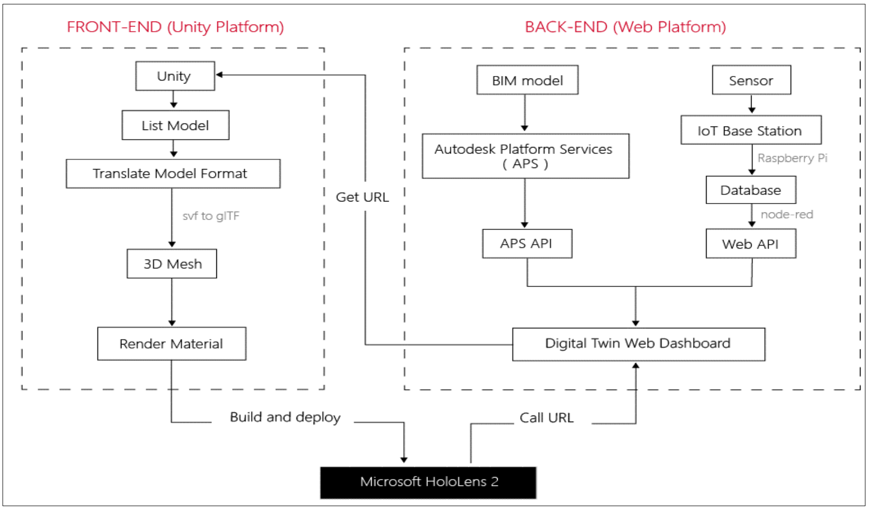


Figure 3. System Framework

3.1.1. Development Environment and Tools

Autodesk Platform Services (formerly Autodesk Forge) provides APIs to convert model files into a more usable format for web and mobile viewing and analysis. It is a cloud-based platform for developers to create and manage applications for the architecture, engineering, construction, and manufacturing industries, empowering innovation for cloud applications and system integration, and enhancing the efficiency of information transformation. Through Autodesk Platform Services, this study can access 2D and 3D design, engineering, and construction data uploaded to Autodesk Platform, build a database of sensor data, and build an IoT base station by Raspberry pi Linux system to combine BIM modelling and sensor data. The data is transferred to the real-time Digital Twin Web Dashboard, which provides an initial interface integration with the front-end (Unity platform).

For the Augmented Reality (AR) user management interface, this study uses the free personal version of the Unity Engine game development engine, version 2021.3.11f1, for prototyping, and all development scripts are built in C# programming language and use the Microsoft Mixed Reality Toolkit to help automate the setup of the mixed reality project. MRTK (Mixed Reality Toolkit) is a developer's mixed reality toolkit for Microsoft HoloLens that helps accelerate cross-platform mixed reality development in Unity by providing cross-platform input systems and building blocks for spatial interactions and UI, allowing users to quickly build interfaces through in-editor emulation, allowing them to view changes immediately. MRTK provides an extensible architecture that allows developers to exchange core components and supports a variety of devices in addition to Microsoft HoloLens 2.

3.1.2. System and Users

The system planning of UbranSight is divided into a back-end (web platform) web page builder and a front-end (Unity platform) interface display. Since this project uses two different platforms, the system operation relationship diagram is drawn in Figure 4. Microsoft HoloLens 2 will return the corresponding model and sensor data for display on the user interface. On the other hand, users (administrators) of the web platform (Digital Twin Web Dashboard) can constantly update the new BIM models through the Autodesk Platform Services cloud service to store the models, and the Digital Twin Web Dashboard provides users with a maintenance update viewing platform. The Microsoft HoloLens 2 and Digital Twin Web Dashboard are connected via the network, and when Microsoft HoloLens 2 calls the Digital Twin Web Dashboard's url, the Digital Twin Web Dashboard returns the corresponding model with the transmission of the model. Dashboard returns the matching model and sensor data.

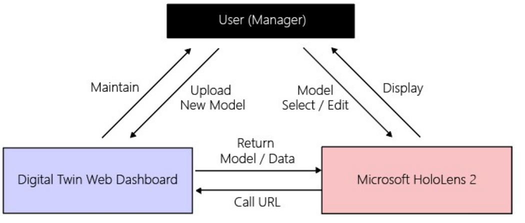


Figure 4. System Operation Relationships

In this study, the system implementation is divided into back-end (web platform) and front-end (Unity platform). The following section describes the back-end (web platform) that deals with the data collection and interfacing of models and sensors, and the front-end (Unity platform) that uses the collected data to construct a user interface.

3.2. BACK-END SETUP (WEB PLATFORM)

To experiment the practicality and effectiveness of this system, the research office of National Cheng Kung University was used as a test site for digital twin model construction and sensor deployment connection. The back-end setup (Web Platform) is planned to be divided into two parts as follows.

1. Digital twin modelling: Implementation of BIM model building and cloud publishing corresponding to the real space.

2. IoT sensor connection: Realize the deployment of IoT sensing and database construction in the real space to establish real-time data connection to the cloud BIM.

Finally, the two parts will be integrated and attached to the Digital Twin Web Dashboard to provide a maintenance platform for users to update their models.

3.2.1. Digital Twin Modelling

This study uses digital twin technology developed by Building Information Modeling (BIM) to create a platform for model change and maintenance for administrators to control the actual state of a spatial area. The technology enables real-time 3D visualization of sensor data in the most intuitive way. The webpage displays the historical environmental parameters of the space and its environment, such as air parameters, crowd status, etc., to assist administrators or users in decision making. The experiment starts with using Autodesk Revit to build the BIM of the experimental field. The basic elements of Revit's building information model can be broadly divided into (1) Model Elements, (2) Datum Elements, and (3) View-Specific Elements.

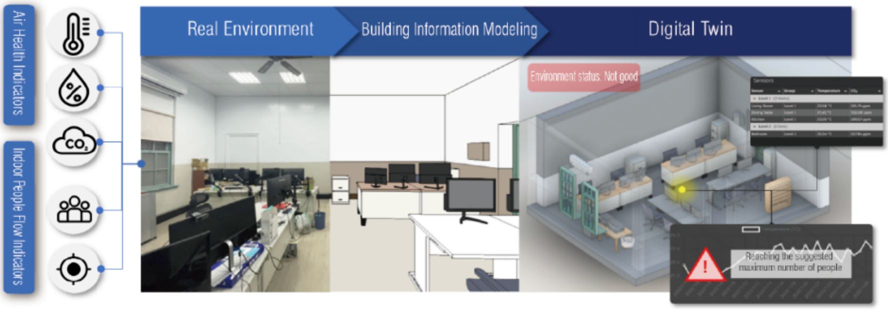


Figure 5. Digital Twin Modelling

Model Elements in a digital model include fixed objects and furniture like walls, roofs, windows, and chairs. Datum Elements encompass structural guidelines such as grids and floors. View-Specific Elements are used for annotations and dimensional marks. Attributes of each component, like usage status and environmental conditions (temperature, humidity), are monitored and adjusted using sensor data.

The model information is then uploaded and stored through Autodesk Platform Services (APS), utilizing four open API functions: Authentication, OAuth, Data Management API, Model Derivative API, and Viewer API. The overall flow of Autodesk Platform Services (APS) platform modelling is shown in Figure 6. Developers create an account and an application on Autodesk Platform Services, receiving a client ID and password. OAuth, a standard for authorization, allows third-party app access to resources on the platform via an Access Token. The Data Management API, mainly used for HTTP Get requests, handles data access and management, including folder and object management in the BIM 360 database.

The Model Derivative API supports over 60 industry model formats and is used for file conversion, including 2D and 3D models, into web-viewable formats like SVF/SVF2. This API also captures model data and assists in listing files for rendering in the viewer SDK. Autodesk Forge Tools, a plugin for Visual Studio Code, is used in this study for efficient interaction with these APIs, enabling easy upload, download, viewing, and deletion of models.

Lastly, the Viewer API allows web-based viewing of models. This JavaScript library, based on three.js, facilitates browsing,

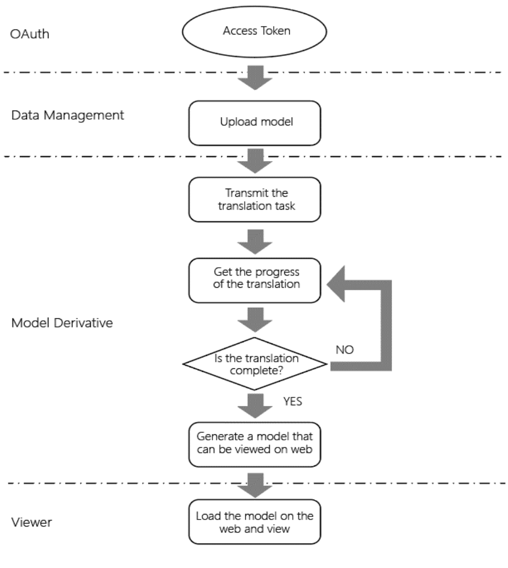


Figure 6. Autodesk Platform Services Modelling Process

viewing, and collaboration on 2D and 3D models within a webpage. In this study, it is used to set up the application interface and add functions such as profiling view, explosion view, measurement, and accessing component attributes, enhancing the application's capabilities.

3.2.2. IoT Sensor Connection

The sensors in this study on the Internet of Things use ADI's Smart Air Detector (SD) and Eagle Eye sensors. The SD is used for real-time sensing of gas composition (temperature, humidity, CO2, PM2.5, and other impact factors) and reporting data for back-end data collection and analysis. Eagle Eye is also used to detect the distribution of pedestrian flow in aisles and areas of the building to understand the usage of the building. In terms of sensor components, Eagle Eye was deployed in the ceiling of the research office. Based on the results of the BIM-assisted calculations, the smart air detector SD was placed in the most suitable position to obtain air health indicators and indoor occupant flow indicators for the site.

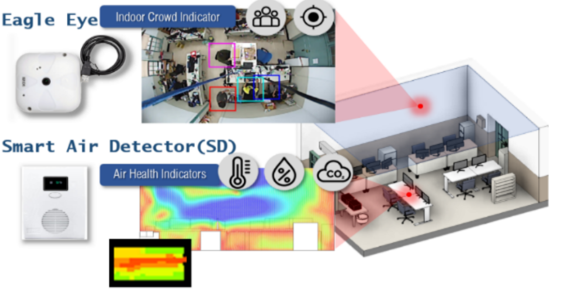


Figure 7. Deployment of sensors

This project integrates two environmental indicators using sensor data, which is then transmitted to an IoT base station built with Raspberry Pi, creating a database. Control of the sensor is managed through a Digital Twin Web Dashboard, allowing adjustments based on detected improvements in environmental quality. For the database, we use Node-RED and MongoDB to construct a NoSQL database, as illustrated in Figure 8. MongoDB, a NoSQL database, stores data in files rather than in a relational format, offering flexibility to suit the specific needs of the building space environment. It utilizes JSON/BSON for data storage and access. Leveraging the JOIN function and the Mongoose library, a popular NPM module, MongoDB enables the development of complex relationships and efficient database designs, free from the constraints of traditional queries. This adaptability makes MongoDB the chosen platform for building our project's database.

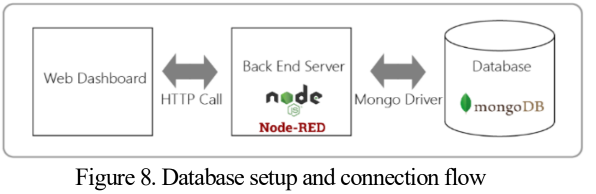


Figure 8. Database setup and connection flow

3.3. FRONT-END SETUP (UNITY PLATFORM)

After the completion of the back-end setup (Web Platform) in the previous section, this section proceeds to the front-end setup (Unity Platform) to import the models and data integrated in the Digital Twin Web Dashboard into the augmented reality (Microsoft

HoloLens 2). The model files are converted to a format that can be rendered by Microsoft HoloLens 2. To achieve a more intuitive, replaceable, and editable model user management interface.

3.3.1. Interface Integration

For the Augmented Reality (AR) user management interface, this study was developed using the Unity Engine. The C# programming language was written to allow Unity to communicate with the Node.js server to load models on the web platform (Digital Twin Web Dashboard). The glTFast-generated glb files are loaded and previewed to convert the model format on different platforms. The model type and rendered material are then created through 3D Mesh, and the rendered glTF model material is converted to a standard material that can be rendered by Microsoft HoloLens 2.

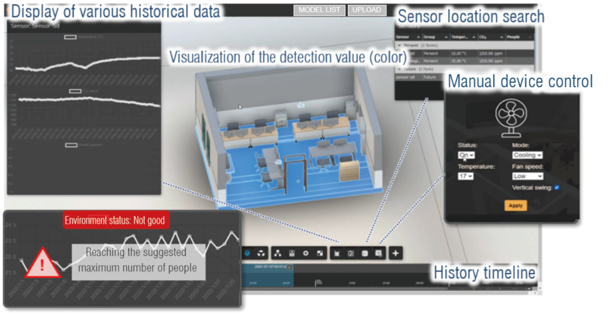


Figure 9. Interface connect to Digital Twin Web Dashboard

For model placement, the study creates a pointer script in Unity that interacts with the Node.js server. This script can identify objects and their locations by decoding the object ID in the color channel. Key variables include the collider (interaction point of the ray), distance (from ray start to collider), triangle index (specific triangle in the mesh), and point (coordinates of ray and collider intersection, defined as Vector3 objects). Vector3 origin specifies the ray's origin point, and a Draw Ray function is used to render the Ray in the virtual world.

3.3.2. Augmented Reality Components

This study uses the Microsoft Mixed Reality Toolkit (MRTK) to set up a mixed reality application in Unity to be compatible with the Microsoft HoloLens 2 headset. MRTK provides components and functionality to accelerate the development of cross-platform mixed reality applications. It was built in Unity using Universal Windows Platform and then deployed to HoloLens 2 using Microsoft Visual Studio using Wi-Fi or USB. In this case, the Wi-Fi deployment project was chosen to connect to Microsoft HoloLens 2 by entering the IP address of the device for matching and remote connection.

4. Experimental Results & Analysis

4.1. DATA CONNECTIVITY

In this study, we use a Raspberry Pi as a cost-effective IoT base station for connecting sensor data, and MongoDB for its scalable data storage capabilities. The database effectively handles model building data and sensor data, but the Raspberry Pi's limited processing power and memory make it unsuitable for larger IoT systems needing extensive data processing. Currently, the system manages two weeks of historical data smoothly. However, scaling this model to a smart city level may overwhelm the Raspberry Pi's capacity, especially for processing large-scale urban data like traffic and weather.

The study also utilizes a Node-RED database, which is user-friendly but can become slow and resource-heavy with large data sets or complex operations, particularly on smaller devices like the Raspberry Pi. Expanding the experiment and increasing the BIM model size could make Node-RED's functions more complex and challenging to manage, as they depend on external libraries.

4.2. USER INTERFACE

Figure 10 presents the HoloLens 2 user interface. The left image shows the startup screen where the user selects our app to enter the augmented reality management interface. The middle image shows the initial screen of the app, still loading the model from the Digital Twin Web Dashboard, with a real environment, a virtual menu, and two virtual cubes for orientation. The right image displays the BIM model fully loaded, allowing the user to switch between models using the virtual menu.

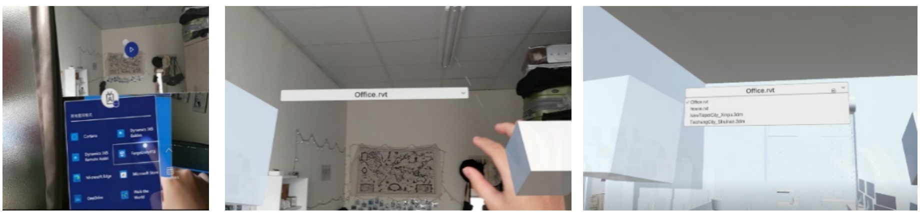


Figure 10. HoloLens 2 user interface
(Left: HoloLens 2 home screen/ Middle: Application Start/ Right: Display BIM model and menu)

Figure 11. Model Object Display (Left: PC device in unity / Right: HoloLens 2)

After testing with HoloLens 2, we discovered that the Unity-designed virtual menu doesn't work well with the HoloLens 2's augmented reality interface. Users have to use a virtual line from their fingertip instead of directly touching the menu. To improve this, the HoloLens 2's user interface should be updated with the Mixed Reality Toolkit (MRTK) modules for a smoother and easier-to-use admin interface.

In our application, model objects are dynamically loaded into the AR environment to keep the user experience updated. However, we've faced challenges with the HoloLens 2 device in accurately rendering textures and colors of these model objects. Sometimes, the objects appear white even with the Mixed Reality Toolkit/Standard shader. This issue could be due to the device's limited processing time, affecting its ability to render during loading. Further investigation and optimization are needed to enhance the visual quality and user experience in our mixed reality application.

5. Conclusion

This study creates a management system combining digital twins with AR headsets, creating an integrated visual management platform. It focuses on using digital twins for effective building lifecycle management and maintenance, facilitated by an AR-enhanced interface for easy use by administrators. The Digital Twin Web Dashboard, developed with Node.js, is integrated into Microsoft HoloLens 2 using the Unity Engine. The system features a back-end (web platform) and a front-end (Unity platform), enabling administrators to intuitively monitor and control the interface.

To enable real-time data transmission, sensors are set up in the set field space. This sensor data is stored in a database and linked to a cloud-based Building Information Modeling (BIM) system for live display. For the Augmented Reality, the Microsoft Mixed Reality Toolkit was used to help automate the setup of the mixed reality project to be compatible with the Microsoft HoloLens 2 headset. By integrating Digital Twin technology with AR, the system provides up-to-date information on urban buildings and infrastructure. The system integrates Internet of Things (IoT) and Building Information Modeling (BIM) data in an easy-to-use Augmented Reality (AR) interface, enhancing smart city management and decision-making.

Acknowledgement

This work was supported by the Taiwan Ministry of Science and Technology, grant No. MOST 111-2221-E-006 -049 -MY2. We sincerely thank Pei-Chi Tsai, Chien-Kai Kuo, Yen-Ching Chang, and Sheng-Heng Liao for their contributions to this project.

References

van Nederveen, G. A., & Tolman, F. P. (1992). Modelling multiple views on buildings. *Automation in Construction, 1*(3), 215-224.

Deng, M., Menassa, C. C., & Kamat, V. R. (2021). From BIM to digital twins: A systematic review of the evolution of intelligent building representations in the AEC-FM industry. *Journal of Information Technology in Construction, 26.*

Khajavi, S. H., Motlagh, N. H., Jaribion, A., Werner, L. C., & Holmström, J. (2019). Digital twin: vision, benefits, boundaries, and creation for buildings. *IEEE access, 7,* 147406-147419.

Gubbi, J., Buyya, R., Marusic, S., & Palaniswami, M. (2013). Internet of Things (IoT): A vision, architectural elements, and future directions. *Future generation computer systems, 29*(7), 1645-1660.

Dunleavy, M., & Dede, C. (2014). Augmented reality teaching and learning. *Handbook of research on educational communications and technology*, 735-745.

Loijens, L. W., Brohm, D., & Domurath, N. (2017). What is augmented reality? In *Augmented reality for food marketers and consumers* (p. 356). Wageningen Academic Publishers.

Zaher, M., Greenwood, D., & Marzouk, M. (2018). Mobile augmented reality applications for construction projects. Construction Innovation, 18(2), 152-166.

Chi, H. L., Kang, S. C., & Wang, X. (2013). Research trends and opportunities of augmented reality applications in architecture, engineering, and construction. *Automation in Construction, 33*, 116-122.

Generative, Algorithmic, & Evolutionary Design

A HYBRID MULTI-OBJECTIVE MODEL FOR MULTI-STORY WAREHOUSE DESIGN

A case study in Shenzhen

YI WEN[1], WEI YE[2] AND GANG YU[3]
[1,3]School of Architecture, Harbin Institute of Technology, Shenzhen,
[2]School of Engineering Technology, Purdue University.
[1]23b356003@stu.hit.edu.cn, 0009-0005-9694-5627
[2]ye375@purdue.edu, 0000-0001-9341-2212
[3]yuarch@hit.edu.cn, 0000-0002-9352-4385

Abstract. The thriving on-demand delivery economy and the increasing focus on addressing the global environmental crisis have spurred the need for efficient and sustainable logistics facilities in urban central areas. This paper delves into the optimization of multi-story logistics warehouse design in Shenzhen, China. Based on comprehensive investigations of existing multi-story warehouses in Shenzhen, the study proposes a hybrid computational model of integer programming and NSGA-II tailored for the generation and optimization of multi-story warehouse general layout design. The prototype, aimed at enhancing efficiency and sustainability, translates these concepts into attainable goals of optimizing land utilization, construction cost, and transportation distance. By addressing multi-objective challenges in the design process, the prototype's effectiveness is validated through a real-world case study. This paper seeks to offer a pragmatic approach to designing cost-effective and resilient multi-story logistics warehouses in the long term, applicable in both Shenzhen and other densely populated urban centers. The insights derived from this study may contribute to the ongoing discourse on optimizing logistics in dynamically evolving industrial landscapes.

Keywords. Logistics warehouse design, multi-story warehouse, generative design, multi-objective optimization, efficiency, sustainability, integer programming, NSGA-II.

1. Introduction

In the past decade, the evolution of supply chain and automation industry dynamics has catalyzed a "vertical shift" in the logistics landscape(Aderneck, 2020; Lim & Park, 2020; Xiao et al., 2021). The surge in e-commerce and the increasing demand for last-mile delivery underscore the vital role of high-capacity logistics facilities in metropolitan areas. In this context, the strategic design of efficient and sustainable

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 283-292.

multi-story logistics warehouses has emerged as a significant topic. Given the fact that salient constraints on land availability(Yuan & Zhu, 2019; Xiao et al., 2021) in monocentric metro areas has less impact in previous studies(Sakai et al., 2015; Giuliano & Kang, 2018; Heitz et al., 2020), this paper delves into the critical domain of multi-story logistics warehouse design in Shenzhen, China.

Researchers have proposed a wide range of multi-story warehouse layout optimization strategies. At the urban scale, current studies have presented the large-scale, multi-tenant and mixed-use logistics solutions in dense urban areas to address space constraints(Xiao et al., 2021; Buldeo Rai et al., 2022). Concerning the need for sustainability, Xiao et al. (2021) has highlighted the sustainable logistics land use patterns in Shenzhen, whereas other studies have also suggested green logistics transportation(Colicchia et al., 2013; Ploos van Amstel et al., 2021). At the facility scale, new heuristics based on GA(Zhang et al., 2002) or its hybrids(Zhang & Lai, 2006) have been developed to tackle multiple-level warehouse layout problems and meet the need for logistics automation. However, the perspective of architectural design of multi-story warehousing buildings has not been adequately studied yet.

Integer programming (IP) is a widely applied mathematical optimization technique in the field of architecture. Previous studies have been conducted to solve the problems including plan layout(Keatruangkamala & Sinapiromsaran, 2005) and transportation network design(Hua et al., 2019). On the other hand, non-dominated sorting genetic algorithm II (NSGA-II) was tailored to cope with multi-objective optimization more effectively(Coello Coello, 2006), which has been used to explore high-performance layout possibilities(R. Wang et al., 2021; Wu et al., 2023). While prior applications of either approach are ample, there are relatively limited researches on combining IP and NSGA-II in layout optimization(Lv & Wang, 2020).

In response to the shortage, this paper proposes a generation and optimization prototype model that integrates IP and NSGA-II. Aiming to provide multi-objective solutions for optimizing multi-story warehouse general layouts, this paper addresses the requirements of efficiency and sustainability from the perspective of architectural design. By presenting a real-world case study in Shenzhen, this research tries to offer insights applicable to regions with similar land use patterns.

2. Methodology

The study employs a synergistic approach, integrating IP and NSGA-II, to optimize layout configurations. IP is known for its efficiency in handling large-scale problems with reduced computational time complexity. On the other hand, NSGA-II, excels in addressing multi-objective optimization problems. By combining the precision and efficiency of IP with the multi-objective optimization prowess of NSGA-II, the study aims to achieve an optimal layout configuration that balances the benefits of both methodologies.

The methodology of this study can be divided into four parts: 1) Input and transformation of real-world information; 2) Establishment of a mathematical model for optimization objectives; 3) Generation of most occupied layouts with IP; 4) Muti-objective optimization to find the best actual layout. The methodology is shown in Figure 1 and further explained in the following.

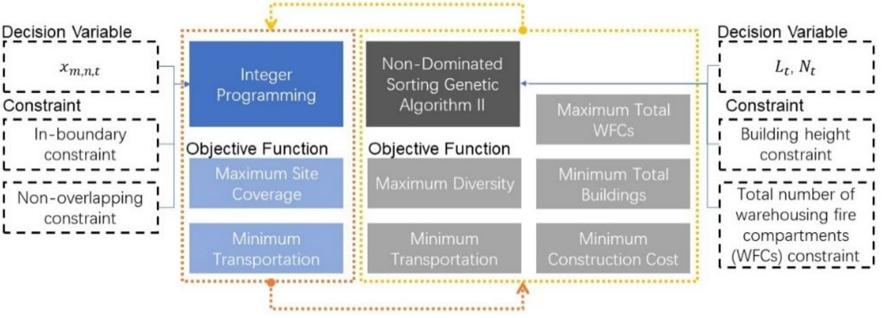


Figure 1. Methodology

2.1. INPUT AND TRANSFORMATION

2.1.1. Processing of Given Data

The processing of basic data involves a three-step procedure. First, a 2D grid coordinate system is established, with the subdivision precision aligned with the standard column spacing of the warehouse (12m). Second, the geometric parameters of the site are input and gridded as a computational domain. Third, warehouse plans are transformed into a series of warehouse templates (WTs) comprising with warehousing fire compartments (WFCs), ramp units (RUs), loading bay and spacing (elaborated further in Section 3.1). Given data and decision variables are declared as follows.

Given data:

- G_h, G_w are the length and width of the grid.
- h_0, w_0, l_0 are the length, width and single storey height of a single fire compartment
- S is the set of the site grid $S = \{i, j | 0 < i \leq G_h, 0 < j \leq G_w\}$.
- H_t, W_t are the length and width of template t, respectively.
- s_i, b_j are additional constraints representing the index of available grids for rows and columns, respectively.
- T is the set of n types of WTs, $T = \{t_1, t_2, \dots, t_n\}$.
- (i_0, j_0) is the fixed reference point representing the site entrance.
- L_0 is the maximum number of building storeys.
- M_t is the number of WFCs of each template t.
- $\text{Max\_}total$ is the maximum number of total warehousing fire compartments.
- $Price$ represents the given construction cost data.

Decision Variables:

- $x_{m,n,t}$ is a binary variable that equals 1 if the top-left corner of template t is placed at cell (m, n) in the grid, and 0 otherwise.
- L_t the number of storeys of actual placed template t.
- N_t the total number of actual placed template t.

2.1.2. Space Constraints

The constraints should be added to the model for placing templates on the grid, considering the dimensions of both the grid and the templates, as well as specific constraints defined by arrays s and b.

1. In-boundary constraint

Restrict placement in last rows and columns:

$$x_{i,j,t} = 0,\ \forall i \in \{G_h - H_t, \dots, G_h\}, j \in \{G_w - W_t, \dots, G_w\} \tag{1}$$

Row-based restrictions using s:

$$x_{i,j,t} = 0, \forall i \in \{0, \dots, G_h - H_t\}, j \in \{\max(0, \min(s_i, s_{i-1+H_t}) - W_t), \dots, G_w\} \tag{2}$$

Column-based restrictions using b:

$$x_{i,j,t} = 0, \forall j \in \{0, \dots, G_w - W_t\}, i \in \{\max(0, b_{j-1+W_t} - H_t), \dots, G_h\} \tag{3}$$

2. Non-overlapping constraint

The key is to ensure that each cell (i, j) in the grid is covered by at most one part of any template, and this is determined by the position of the top-left corner of the templates and their dimensions.

$x_{m,n,t}$ is indicating whether the top-left corner of template t is placed at cell (m, n). m, n iterate over the potential top-left corner positions of the templates that could cover cell (i, j).

$$\sum_{t \in T} \sum_{m=\max(0,i-H_t)}^{\min(i,G_h-H_t)} \sum_{n=\max(0,j-W_t)}^{\min(i,G_w-W_t)} x_{m,n,t} \leq 1,\ \forall i \in \{0, \dots, G_h\}, j \in \{0, \dots, G_w\} \tag{4}$$

3. Building height constraint

According to the prescribed warehousing storey height (8m) and the baseline height (24m) for multi-story buildings, a multi-story warehouse has a minimum of 3 storeys.

$$3 \leq L_t \leq L_0 \tag{5}$$

4. Total number of warehousing fire compartments constraint

The total number of WFCs is constrained by the site area and the prescribed floor area ratio.

$$\sum_{t \in T} L_t M_t N_t \leq \max\_total \tag{6}$$

2.2. MATHEMATICAL MODEL ESTABLISHMENT

Multi-story warehouses, as opposed to their traditional single-story counterparts, are characterized by their occupation of denser urban spaces, elevated floor-to-ceiling heights, and enhanced spatial efficiency, promoting sustainable land usage. These structures facilitate optimized transportation networks, potentially diminishing transit distances, curtailing logistical expenditures, and contributing to a reduction in carbon emissions. This research quantitatively translates these efficiency and sustainability imperatives into measurable objectives: land utilization, construction cost, and transportation proximity. Land utilization is further disaggregated into planar density and diversity, as well as vertical intensification metrics. Accordingly, a total of six distinct objective functions are established.

1.Max Site Coverage: The strategy optimizes planar density by employing a stochastic template filling method to determine the maximum feasible WFCs on the ground plan, ensuring optimal long-term land use efficiency.

2.Max Diversity: Based on the candidates derived in the preceding step, this objective ensures diverse warehouse scales within the site by selecting the results featuring the richest types of WTs. This approach is designed to enhance the long-term profitability of warehouse operations and the overall land utilization value.

3.Max Total WFCs: The aggregate count of WFCs, distributed across all levels, directly correlates with the storage capacity of the land, where an increase in WFCs within predetermined bounds augments the overall storage rate.

4.Min Total Buildings: The vertical intensification is evaluated through randomized allocations of storeys to WTs, conforming to regulatory height and floor area ratio limitations, aiming to minimize the count of buildings.

5.Min Construction Cost: Estimated construction expenses are derived from predefined pricing datasets, providing a cost framework for the evaluation of design alternatives.

6.Min Transportation Distance: An ideal logistic layout is characterized by its succinct transportation network, with the aggregate transportation distance minimized across all levels, excluding pedestrian traffic considerations. The overall transportation distance comprises 1) the distance from the site entrance to the buildings and 2) the distance within the buildings, with the former minimized using linear programming, and the latter minimized with NSGA-II.

2.3. IP LAYOUT GENERATION

We first employ IP to address the objectives of site coverage and transportation distance, involving the variable $x_{m,n,t}$ and the constraints expressed by Eqs. (1)-(4).

2.3.1. Maximum Site Coverage

The generation of general layout candidates starts from maximizing site coverage, where the variables under consideration are the types and quantities of WTs. The primary objectives include 1) maximizing the number of WFCs on the site and 2) increasing the quantity of WTs containing more WFCs.

While IP would yield a single optimal solution for these objectives, the prototype should provide a comprehensive set of planar arrangements for screening during the early stages to better align with real-world scenarios. Therefore, we transform these two objectives into constraints based on practical considerations and generate all solutions meeting these criteria. The result contains a set of layouts that meets the maximum site coverage requirement.

2.3.2. Minimal Transportation Distance from the Site Entrance

The transportation distance from the site entrance to the buildings is minimized in the IP model. The distance $Dis_{m,n,t}$, for a template t placed at position (m,n) is calculated as:

$$Dis_{m,n,t} = |m - (i_0 - H_t)| + |n - (j_0 - W_t)| \quad (7)$$

This represents the distance from the bottom-right corner of the placed template t to the reference point (i_0, j_0).The objective is to minimize the total distance of all placed templates from the entrance:

$$min \sum_{t \in T} \sum_{m=0}^{G_h} \sum_{n=0}^{G_w} Dis_{m,n,t} * x_{m,n,t} \quad (8)$$

The result is an integrated solution that optimally achieves maximum site utilization while also minimizing the transportation distance between the site and the constructed facilities.

2.4. MULTI-OBJECTIVE OPTIMIZATION

In the optimization phase, NSGA-II is employed to search for the Pareto Fronts. Based on the principles of NSGA-II, the multi-objective optimization process can be described as follows.

To enhance computational efficacy and maintain the integrity of integer constraints post-genetic operations, we refined the genetic algorithm component within NSGA-II. This refinement ensures that the resultant offsprings, following crossover and mutation events, remain confined to the predefined integer bounds. The decision variables in NSGA-II, are L_t, N_t, which are both integer variables with a given range. the range of L_t is from building height constraint. The result from Max Site Coverage can be used as an indicator for the range of N_t. The constraints in NSGA-II are expressed by Eqs. (5) and (6).

The objectives in NSGA-II are:

1.Max Diversity: $\max \sum_{t \in T} [N_t = 0]$ (9)

2.Max Total WFCs: $\max \sum_{t \in T} L_t M_t N_t$ (10)

3.Min Total Buildings: $min \sum_{t \in T} N_t$ (11)

4.Min Construction Cost: $\min \sum_{t \in T} N_t * Price(L_t)$ (12)

5.Min Transportation distance within the building:

$$\min \sum_{t \in T} \left(\sqrt{(2h_0 + 2w_0)^2 + {l_0}^2} (L_t - 1) + max\,(H_t, W_t) \right) \times 2N_t \qquad (13)$$

3. Case Study

The application of the prototype is illustrated in the following via a comparative case study of a multi-story warehouse project located in the Yantian Port (Figure 2a), the eastern part of central Shenzhen, China.

The current general layout (Figure 2b) features two five-story warehouses positioned on the southeast side of the site. By re-designing and optimizing the general layout with the proposed prototype, the existing project is used as a benchmark for evaluating the efficacy of the design approach.

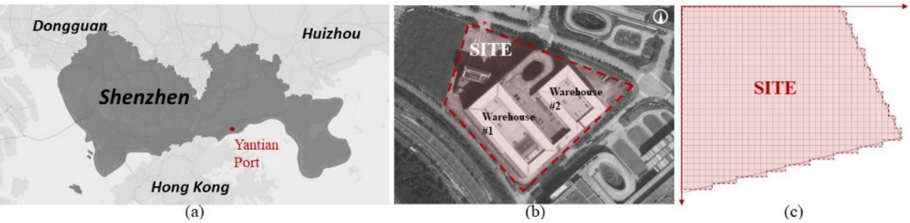


Figure 2. Site location (a), current general layout (b), and gridification of the site(c)

3.1. PRELIMINARY DEFINITION

After gridifying of the site (Figure 2c), we defined fundamental assumptions related to WTs, WFCs, and RUs. 1) They are all rectangular, with the wall width being negligible. 2) A WFC adhering to fire safety regulations consists of 25 grid cells. 3) A WT consists of at least two WFCs, at most four WFCs, and one RU. WTs can be arranged with a 0-

degree orientation or rotated by 90 degrees, resulting in a set of 6 possible configurations, denoted as $T = \{t_1, t_2, t_3, t_4, t_5, t_6\}$. 4) The prescribed spacing between WTs is one grid and is contained in the templates. An illustrative diagram based on the assumptions is shown in Figure 3.

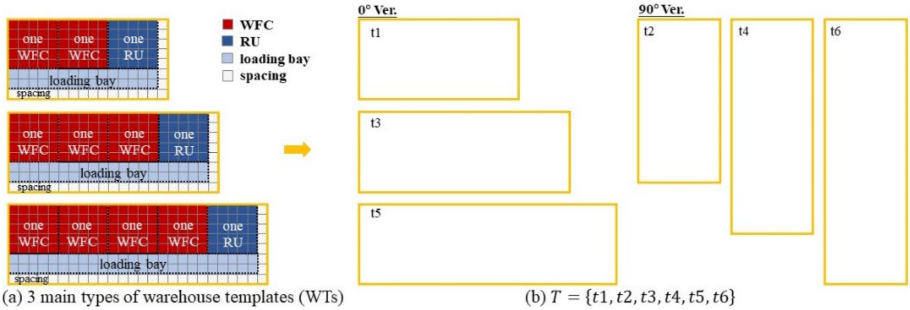


Figure 3. Preliminary definition of templates

The area of the site is approximately 55,000 ㎡. Considering the local regulations on floor area ratio (3.0) and building height (50m), as well as the practical requirements of decision-makers for this real-world project, we have calculated that the total number of WFCs placed on the site should not exceed 40, with a maximum number of building storeys $L_0 = 6$. The fixed reference point representing the site entrance is defined as $(24, 20)$. Additionally, based on the given construction cost data, we have compiled a rough cost estimate, as shown in Table 1.

Table 1. Given average construction cost (CNY)

Storeys \ WT	Per ㎡	t1, t2	t3, t4	t5, t6
[redacted]	2,200	33,264,000	44,352,000	55,440,000
[redacted]	2,800	84,672,000	112,896,000	141,120,000
[redacted]	3,200	217,728,000	290,304,000	362,880,000
[redacted]	3,600	302,400,000	403,200,000	504,000,000
[redacted]	4,000	399,168,000	532,224,000	665,280,000
[redacted]	4,400	217,728,000	290,304,000	362,880,000

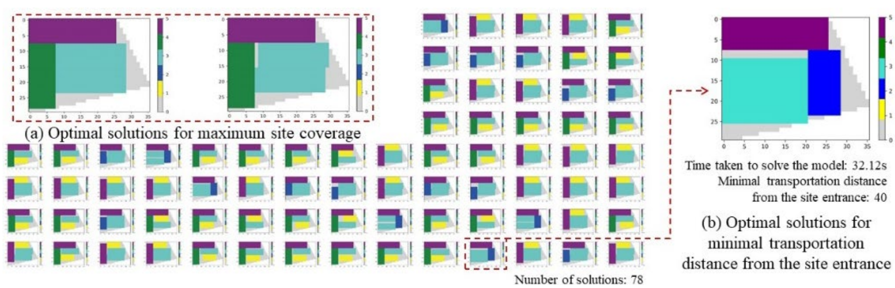


Figure 4. Optimal solutions generated through IP

3.2. ALTERNATIVE GENERAL LAYOUT GENERATION

This section demonstrates a set of optimal solutions generated through IP for maximum site coverage (Figure 4a) and minimal transportation distance from the site entrance (Figure 4b). Each of these solutions serves as an alternative proposal for the existing general layout.

3.3. OPTIMIZATION RESULTS

NSGA-II is applied here to solve the multi-objective problems. The numbers of initial individuals and generations are both set to be 100. The Pareto Fronts obtained are visualized in a three-dimensional chart. A comparative analysis of two sets of solutions is conducted and further explained as follows.

3.3.1. Optimal Solutions for Land Intensification

In this section, we fix the value of diversity at 3 or 4 and use it as a constraint, with land intensification as the variable. The obtained Pareto Fronts (Figure 5) spans from the top-left to the bottom-right of the plot, which illustrates the gradual change of layout from "density and intensification" to "low cost and transportation".

In the analysis, a total of five solutions grouped into three sets were examined, focusing on better intensification. The comparison involved cost, total number, and transportation metrics. Set 1 exhibited the highest intensification but the lowest WFC quantity. Between Set 2 and Set 3, Set 2 outperformed in both intensification and total numbers, while Set 3 was less intensive but more cost-effective. Notably, Solution 1(a) and Solution 3(c) share the same configuration in the previous IP results, as do Solution 2(b) and Solution 3(d). A closer look at the original configurations revealed better transportation in the former one. Thus, for those pursuing the shortest transportation distance, the former layout configuration is recommended, while those emphasizing WFC total number and intensification may prefer the latter.

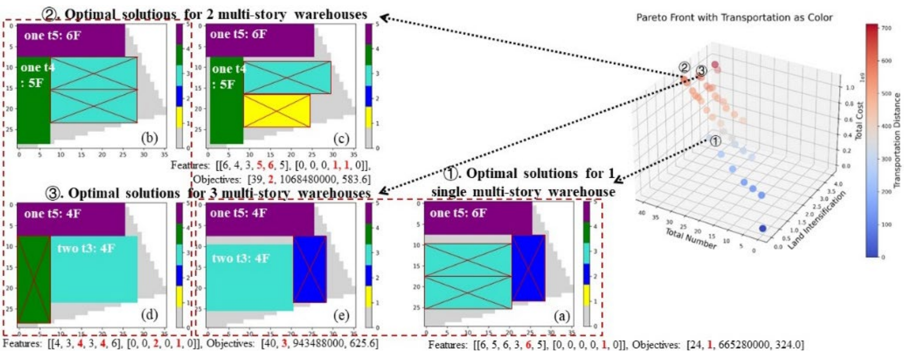


Figure 5. Optimal solutions for land intensification

3.3.2. Optimal Solutions for Diversity

In this section, we set diversity as the variable, with the value of intensification fixed at 4 and treated as a constraint. This approach enables us to derive solutions for different levels of diversity. As shown in Figure 6, the majority of the solutions are clustered towards the middle of the "Total Number" and "Land Intensification" range, with a moderate "Total Cost" and "Transportation Distance", representing a balanced trade-off among the considered objectives.

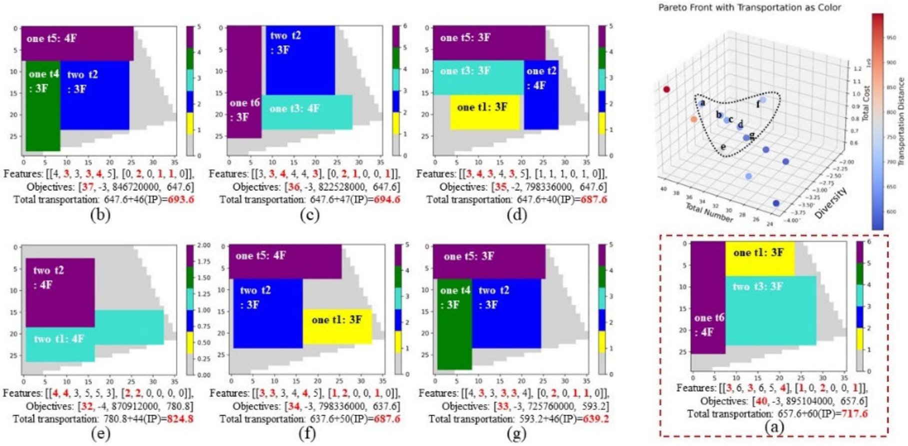


Figure 6. Optimal solutions for diversity

In the analysis, 7 solutions are selected from the cluster with costs under 1,000,000,000, transportation under 800, and a total number greater than or equal to 32. Notably, Solution (a) stands out from a long-term sustainability perspective due to its highest number of WFCs. This result is consistent with the objective of maximizing site coverage, confirming alignment between the outcomes of IP and NSGA-II.

4. Conclusion

This research proposed a generation and optimization prototype model that supports the multi-story logistics warehouse general layout design. The prototype exhibits a significant advantage in providing considerable decision-making flexibility for navigating trade-offs between efficiency and sustainability. The future refinement of the model will involve rotations of the grid, particularly in sites with non-orthogonal contours. Additionally, enhancing the diversity of warehouse templates (WTs) will be considered, taking into account practical scenarios such as two warehouses sharing a circular ramp. While the prototype is developed using Shenzhen as a sample, its potential applicability extends to other densely populated or concentrated metropolitan areas confronted with comparable challenges of land scarcity and intensive development, which requires further research.

References

Aderneck, E. (2020). From Horizontal to Vertical: Industrial Intensification Grows Up. *Development*, *Volumn 51, No.1*(Spring 2020), 74.

Buldeo Rai, H., Kang, S., Sakai, T., Tejada, C., Yuan, Q. (Jack), Conway, A., & Dablanc, L. (2022). 'Proximity logistics': Characterizing the development of logistics facilities in dense, mixed-use urban areas around the world. *Transportation Research Part A: Policy and Practice*, *166*, 41–61. https://doi.org/10.1016/j.tra.2022.10.007

Coello Coello, C. A. (2006). Evolutionary multi-objective optimization: A historical view of the field. *IEEE Computational Intelligence Magazine*, *1*(1), 28–36. https://doi.org/10.1109/MCI.2006.1597059
Colicchia, C., Marchet, G., Melacini, M., & Perotti, S. (2013). Building environmental sustainability: Empirical evidence from Logistics Service Providers. *Journal of Cleaner Production*, *59*, 197–209. https://doi.org/10.1016/j.jclepro.2013.06.057
Giuliano, G., & Kang, S. (2018). Spatial dynamics of the logistics industry: Evidence from California. *Journal of Transport Geography*, *66*, 248–258. https://doi.org/10.1016/j.jtrangeo.2017.11.013
Heitz, A., Dablanc, L., Olsson, J., Sanchez-Diaz, I., & Woxenius, J. (2020). Spatial patterns of logistics facilities in Gothenburg, Sweden. *Journal of Transport Geography*, *88*, 102191. https://doi.org/10.1016/j.jtrangeo.2018.03.005
Hua, H., Hovestadt, L., Tang, P., & Li, B. (2019). Integer programming for urban design. *European Journal of Operational Research*, *274*(3), 1125–1137. https://doi.org/10.1016/j.ejor.2018.10.055
Keatruangkamala, K., & Sinapiromsaran, K. (2005). Optimizing Architectural Layout Design via Mixed Integer Programming. In B. Martens & A. Brown (Eds.), *Computer Aided Architectural Design Futures 2005* (pp. 175–184). Springer Netherlands. https://doi.org/10.1007/1-4020-3698-1_16
Lim, H., & Park, M. (2020). Modeling the Spatial Dimensions of Warehouse Rent Determinants: A Case Study of Seoul Metropolitan Area, South Korea. *Sustainability*, *12*(1), Article 1. https://doi.org/10.3390/su12010259
Lv, H., & Wang, T.-K. (2020). Development of a Nonlinear Integer Optimization Model for Tenant Mix Layout in a Shopping Centre. *Advances in Civil Engineering*, *2020*, e2787351. https://doi.org/10.1155/2020/2787351
Ploos van Amstel, W., Balm, S., Tamis, M., Dieker, M., Smit, M., Nijhuis, W., & Englebert, T. (2021). *Go electric: Zero-emission service logistics in cities*. Hogeschool van Amsterdam.
Sakai, T., Kawamura, K., & Hyodo, T. (2015). Locational dynamics of logistics facilities: Evidence from Tokyo. *Journal of Transport Geography*, *46*, 10–19. https://doi.org/10.1016/j.jtrangeo.2015.05.003
Wang, R., Wang, Y., Gundersen, T., Wu, Y., Feng, X., & Liu, M. (2021). A multi-objective optimization method for industrial park layout design: The trade-off between economy and safety. *Chemical Engineering Science*, *235*, 116471. https://doi.org/10.1016/j.ces.2021.116471
Wu, Z., Ran, K., Lv, H., & Wang, T.-K. (2023). Generative design: Integrating rent and retail compatibility goals for automated tenant mix layout. *Journal of Building Engineering*, *79*, 107845. https://doi.org/10.1016/j.jobe.2023.107845
Xiao, Z., Yuan, Q., Sun, Y., & Sun, X. (2021). New paradigm of logistics space reorganization: E-commerce, land use, and supply chain management. *Transportation Research Interdisciplinary Perspectives*, *9*, 100300. https://doi.org/10.1016/j.trip.2021.100300
Yuan, Q., & Zhu, J. (2019). Logistics sprawl in Chinese metropolises: Evidence from Wuhan. *Journal of Transport Geography*, *74*, 242–252. https://doi.org/10.1016/j.jtrangeo.2018.11.019
Zhang, G. Q., & Lai, K. K. (2006). Combining path relinking and genetic algorithms for the multiple-level warehouse layout problem. *European Journal of Operational Research*, *169*(2), 413–425. https://doi.org/10.1016/j.ejor.2004.08.007
Zhang, G. Q., Xue, J., & Lai, K. K. (2002). A class of genetic algorithms for multiple-level warehouse layout problems. *International Journal of Production Research*, *40*(3), 731–744. https://doi.org/10.1080/00207540110093909

ARCHITECTURAL TOPOLOGICAL FORM-FINDING INTEGRATING SOLID AND FLUID STRUCTURAL PERFORMANCES

XIN YAN[1], KUN DI SHU[2] and DING WEN BAO[3,*]
[1]*Future Laboratory, Tsinghua University | Academy of Arts & Design, Tsinghua University*
[2]*School of Architecture, Tianjin University*
[3]*School of Architecture and Urban Design, RMIT University | Centre for Innovative Structures and Materials, RMIT University*
[1]*yanxin2022@tsinghua.edu.cn,0000-0002-5033-3597*
[2]*arch kundishu@tju.edu.cn, 0009-0001-5202-7243*
[3]*nic.bao@rmit.edu.au, 0000-0003-1395-8747*

Abstract. With the recent developments of digital architecture techniques, performance-based optimisation has been an essential topic in architecture design. Using Finite Element Analysis (FEA) and structural topology optimisation algorithms, designers can easily generate architectural forms with high mechanical performances and unique elegant shapes. Comfortable and pleasant architectural microenvironments can also be designed with Computational Fluid Dynamics (CFD) techniques. However, the architectural form-finding method integrating the above two aspects remains a current research hotspot with room for further exploration. This paper presents an innovative Fluid-Structure-Interaction (FSI) topological optimisation workflow for optimising architectural forms based on both inner solid and surrounding fluid mechanics. This framework consists of three basic parts: (1) fluid-structure interaction (FSI) analysis of buildings and their surroundings, (2) automatic modelling of building forms & surrounding environments, and (3) architectural evolutions referred to gradient-based theory. The research aims to construct an innovative architectural morphological topology optimisation algorithm based on the integration of solid and fluid structural performances. The method also shares the potential to coordinate the diverse architectural physical requirements in the form-finding process for complex building contexts, which holds significant practical potential in architectural and urban design.

Keywords. Topology Optimisation, Solid Structural Performance, Fluid Structural Performance, Fluid-structure Interaction, Form-finding.

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 293-302. © 2024 and published by the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA), Hong Kong.

1. Introduction

The advancement of computational technology has provided architects and engineers with numerous digital performance analysis tools. However, normal architectural designs usually require analysis and simulation based on existing architectural form drafts. Given correct decisions during the early stages of design significantly impact future energy consumption and construction costs, the simulation should shift from "posterior" analytical tools to "anterior" generative tools (Lin, 2019).

Currently, significant advancements in computational capabilities have led to substantial developments in performance-based architecture form generation. But most of them are always developed for several specific issues based on some certain digital platforms, rather than the synchronous comprehensive considerations of the diverse complex physical architectural form requirements. As a complex project, architectural forms should be a negotiated result that is generated with the comprehensive consideration of multi-physical fields and these fields can always be divided into building inner solid field and surrounding fluid field. Thus, the classical working flows with diverse professional analysis platforms fail to effectively achieve accurate form-finding results with synchronous and collaborative analysis of multi-fields.

As the two famous general numerical analysis techniques for static solid objects and dynamic fluid fields respectively, Finite Element Analysis (FEA) and Computational Fluid Dynamics (CFD) are simultaneously introduced into an architectural topological form-finding method using the Fluid Structure Interaction (FSI) method in this paper. This innovative automatic digital architectural topological form-finding method can effectively perform fluid-structure coupling operation, analyse the reciprocal influences between a building and its environment. This research bridges the gap that has historically existed in architectural form-finding methods, which were constrained to single building and solid structural performance. The profound significance of this study is evident in its capacity to analyse and optimize the intricate interactions not only among building clusters but also between architectural entities and their environments within urban contexts.

2. Literature Review

2.1. ARCHITECTURAL TOPOLOGY OPTIMISATION

Based on FEA platforms, structural topology optimisation algorithms recently gain widespread attention due to their abilities to effortlessly generate architectural forms characterized by both high mechanical performance and distinctive elegant shapes. The overarching objective of structural topology optimisation is to determine an optimal primary structural layout while adhering to constraints on material consumption, leveraging principles from computational mechanics. There have been several notable methods applied in architecture design. (Yan et al., 2022; Yan, Bao, Xiong, et al., 2023) propose several detail control strategies to pre-design architectural forms in topology optimisation. (Li & Xie, 2021) introduces multi-material constraint into BESO method to separate tension and compression structures in architecture design. The works of (Ohmori, 2011) and (Sasaki et al., 2007) have been instrumental in assisting numerous Japanese architects in designing various building forms through their extended evolutionary structural optimisation (EESO) method. Other notable contributions include the introduction of the BESO method in (Bao et al., 2022; Duan et al., 2023; Ma et al., 2023; Yan et al., 2019; Yan, Bao, et al., 2023), aimed at achieving diverse building structures and fabrications. (Xie, 2022) has also proposed generalized versatile multi-directional control approaches for architectural design.

However, most of the above researchers aim to find the optimal architectural form design with high solid mechanical performance using unchangeable mesh models and static mechanical analysis only referred to the physical building bodies. This limitation significantly impacts the applicability in collaborative research involving architectural and environmental considerations.

2.2. ENVIRONMENTAL PERFORMANCE BASED FORM FINDING

The building's outer environments are intangible and variable fields that require CFD techniques that can effectively deal with fluid analysis and thermal radiation. In these fields, various scholars from architecture-related disciplines have employed diverse techniques for studying different design objects. The wind environment analysis techniques and low-speed wind tunnel experiment platform are considered a crucial tool for analysing architectural wind environments (Yuan et al., 2021). (Song & Yuan, 2021) has integrated CFD technology with parametric modelling tools, machine learning, and other digital techniques to develop architectural design methods focused on wind environments.

However, constrained by specific architectural simulation platforms, existing research predominantly revolves around interconnecting data from various platforms for diverse building performances to establish workflows, which still encounters certain limitations in terms of computational efficiency and diversity of generated solutions. Although (Feng et al., 2022) develops an evolutionary working loop necessitating manual data transfer across various platforms, the manual processes and rudimentary calculation models have proven to be inefficient and imprecise. These inherent limitations have impeded their practical application in real architectural design contexts.

2.3. COMPUTATIONAL GEOMETRIC MODELLING

In this study, a distinctive difference compared to conventional topology optimisation methods is the dynamic adaptation of the FEA mesh of the architectural form during the optimisation process. This adaptation allows the mesh to evolve with the building morphological changes, enabling the synchronous update of the surrounding fluid field at each iteration. Therefore, the additional computational geometric modelling approaches, particularly mesh smoothing and reconstruction, become crucial research directions.

Remeshing involves improving a triangulation from a potentially noisy triangulation or sampled (scanned) data, and this method is useful for polish the surfaces of topologically optimised architecture forms. After FEA and topology optimisation, it is convenient to generate the point cloud on the optimised design, and thus there are mainly four fundamental methods to construct the smooth surface mesh, which are Alpha shapes (Edelsbrunner & Mücke, 1994), Ball pivoting (Bernardini et al., 1999), Poisson surface reconstruction (Kazhdan et al., 2006), and voxel grid (Lorensen & Cline, 1998). The 3D FEA/CFD mesh should be generated based on the 3D mesh filling algorithms (Hang, 2015).

3. FSI based Architectural Form Optimisation Methodology

3.1. ALGORITHM FRAMEWORK

The FSI based Architectural Topology Optimisation (FSI-ATO) method in this paper is composed of three basic parts: (1) FSI analysis of buildings and their surroundings, (2) automatic modelling of building forms & surrounding environments, and (3) topological evolutions referred to gradient-based theory. Fig. 1 illustrates the basic algorithm workflow in this paper. The evolutionary iteration procedure of FSI-ATO is given as follows.

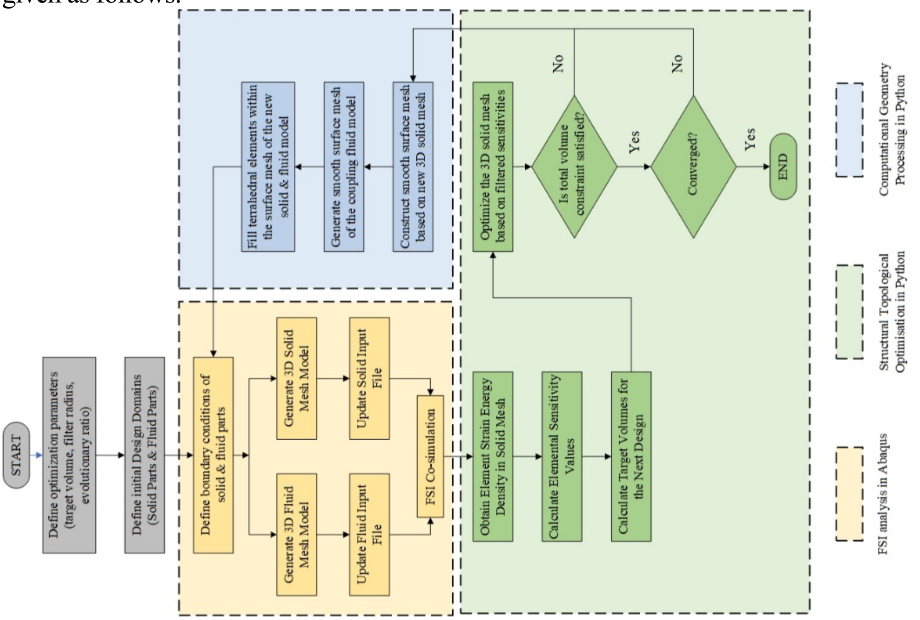


Fig.1 Working flow of FSI-ATO.

3.2. FLUID-STRUCTURE INTERACTION ANALYSIS

Fluid-structure interaction (FSI) refers to the complex interaction between a fluid (liquid or gas) and a structure (solid) that is immersed or interacts with the fluid. It is essential to consider when analysing the behaviour of structures subjected to fluid flow or vice versa. In FSI, the deformation of the structure affects the flow of the surrounding fluid, and, in turn, the fluid exerts forces on the structure. The complexity of these calculations demands a substantial number of simulations and time. For this study, the Abaqus 2016 has been introduced to conduct Fluid-Structure Interaction (FSI) computations effectively.

In Abaqus 2016, the FSI co-simulation is based on two separate modules: the CFD solver and the standard/explicit solver. The computational core will generate two distinct input files for the fluid and solid models respectively, which are also the fundamental data exchange carriers between our python codes and the Abaqus software. The input files contain all the necessary solid/fluid data for FSI, including the mesh information, the boundary conditions, material assumptions, analysis settings, output recording requirements, etc. During the whole evolution process, both the solid and fluid input files update their mesh data and the solid-fluid interface information simultaneously.

3.3. TOPOLOGY OPTIMISATION

Through FSI analysis, the structural mechanical states under the influence of the fluid can be easily analysed. And the structural topology optimisation can be modified according to the multi-volume constraint BESO method (Yan, Xiong, et al., 2023) for optimising multi-building clusters in cities.

According to the BESO theory, the elemental sensitivity numbers α_i, representing how much the element contributes to the structure, can be calculated easily with the FSI output data (e.g., strain energy density, von Mises stress, etc.) as follows,

$$\alpha_i = \sum_{k=1}^{M} w_k^{load} \cdot NS_i^{(k)}$$

$$\sum_{k=1}^{M} w_k^{load} = 1$$

$$NS_i^{(k)} = \frac{S_i^{(k)} - \min(S^{(k)})}{\max(S^{(k)}) - \min(S^{(k)})}$$

where $NS_i^{(k)}$ represent the normalised simulation data of i-th element in k-th loading case, including solid and fluid analysis. And w_k^{load} is the weighting coefficient for k-th load case which is defined according to the project requirements before the evolution and describe how importance the load case is.

Then, the elements with low sensitivity values are removed while the high-sensitivity elements are reserved in the next iteration. Thus, the evolutionary loop stops when the target volume fraction (VF) and following coverage requirements are satisfied,

$$\frac{\left|\sum_{i=1}^{N}\left(C_{k-i+1} - C_{k-N-i+1}\right)\right|}{\sum_{i=1}^{N} C_{k-i+1}} \leq \tau$$

in which k is the current iteration number, τ is an allowable convergence error ($\tau = 0.001$ in this paper), which means stable compliance at least in successive 10 iterations.

One biggest difference between FSI-ATO and classical topology optimsation methods is that the fluid field behaviour can be accurately introduced in optimisation. In conventional topology optimisation like BESO or Ameba software, the only way to simulate the wind load is to simplify the flow pressure with formulas (like uniform distribution or triangular force proportional to the height of the building, etc), which fails to take the eddy current variation and flow interactions between neighbour buildings. However, FSI-ATO method can be used to solve these problems.

3.4. COMPUTATIONAL GEOMETRY PROCESSING

After removing/adding elements, the surface of the solid model tends to exhibit numerous sharp irregularities. These morphological imperfections not only impact the aesthetic appeal and manufacturing complexity but also hinder the flow of surrounding fluids. Therefore, in each optimisation iteration, the optimised architectural mesh needs to undergo a smoothing process.

As shown in Fig. 2(b), topology optimized solid model usually shares sharp angles over the external surface, and thus a Poisson reconstruction smooth method (Kazhdan et al., 2006) is introduced to construct a smooth shell surface for the buildings. After getting the smooth surface mesh of building, its updated surrounding fluid field surface mesh can be easily constructed with the Mesh Boolean Operations. Furthermore, due to both the smooth surface meshes of building and its environment (Fig. 2(c)) only have smooth outer shells without any solid elements inside, a 3D tetrahedral generation method (Hang, 2015) is implemented to fill in the inner space based on the smooth outer surface mesh for the topology optimization evolution in the next step. Subsequently, new tetrahedral computational meshes for both solids and fluids are generated. This ensures the mesh is refined and conducive to fluid flow in the vicinity, contributing to a more visually appealing and manufacturable architectural design.

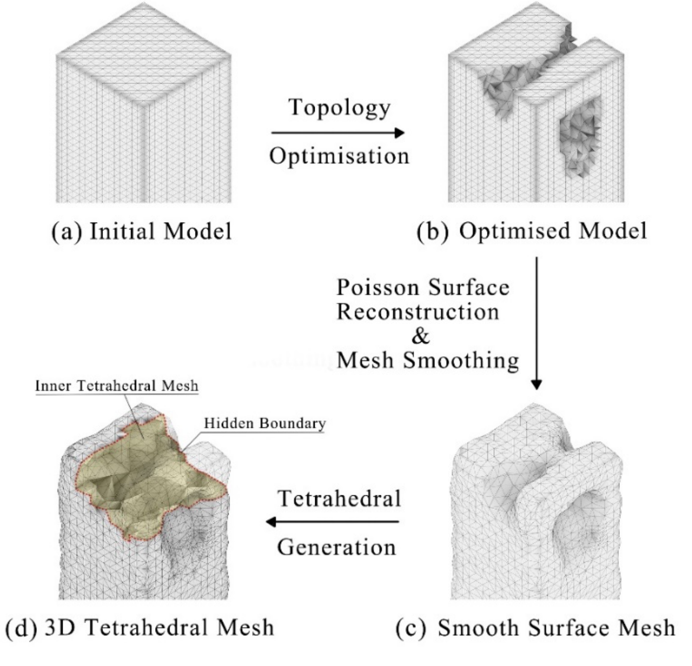


Fig.2 Computational geometric modelling in FSI-ATO.

4. Numerical examples

4.1. FAÇADE OPTIMISATION OF SINGLE BUILDING

In this part, a traditional 2D high-rise building with horizontal wind load is introduced. The initial architectural form domain is assumed as a long rectangle with width of 30m and height of 90m. The solid material is set as steel with density of $7800\ \mathrm{kg/m^3}$, Young's modulus of $210\ \mathrm{GPa}$ and Poisson ratio of 0.3. The values of target volume fraction, filter radius and evolutionary ratio are 50%, 2m and 5%, respectively. The bottom boundary is fixed in three directions.

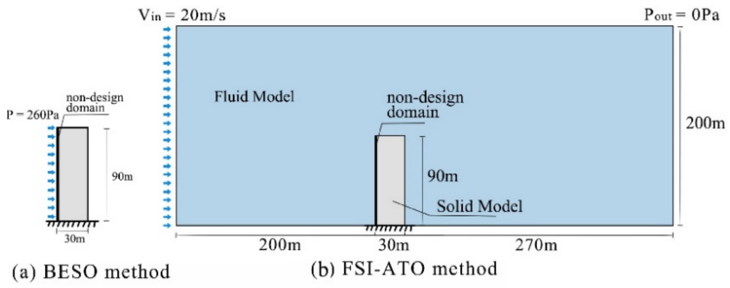
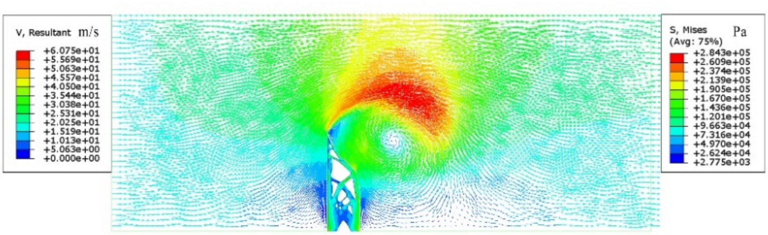
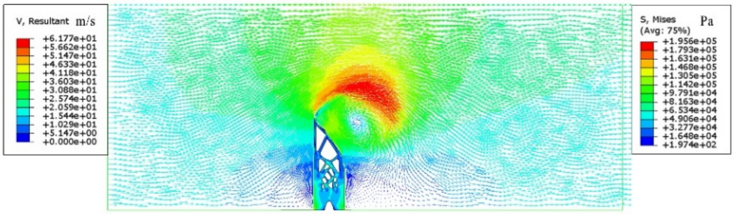


Fig.3 Boundary conditions of BESO and FSI-ATO methods.

As illustrated in Fig. 3, in the traditional BESO method, the Class 8 horizontal wind load (with wind speed of 20 m/s) has to be represented by a uniform wind pressure of $260\ \mathrm{N/m^2}$ due to there is not surrounding fluid model. Meanwhile, the left boundary must be set as non-design domain to resist the pressure, which limits the scopes of architectural form-finding. However, with FSI-ATO method, the additional fluid model can be defined as incompressible navier stokes with the air density of $1.29\ \mathrm{kg/m^3}$. Thus, the initial wind velocity can be directly defined on the left inlet boundary. Fig. 4 lists the two results of BESO and FSI-ATO method respectively. It is obvious that the BESO optimised design holds more structural Mises stress and lower wind speed than FSI-ATO design. This means that FSI-ATO can generate more accurate and reasonable architecture forms in real wind fluid environment to reduce its own inner stress level and make the surrounding air easier to circulate at the same time. It is noted that although Fig. 4(b) shows an example of FSI-ATO with fixed left boundary, FSI-ATO has advantage of defining non-design domains freely, as shown in Fig. 5(b).



(a) BESO optimised design



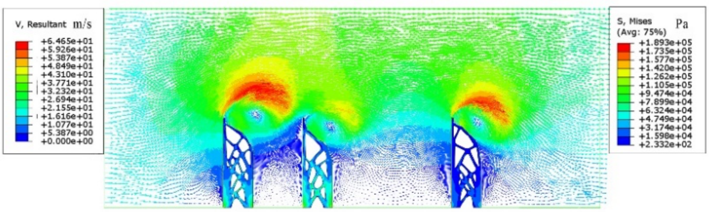
(b) FSI-ATO optimised design

Fig.4 Wind speed and Mises stress analysis of BESO and FSI-ATO structures.

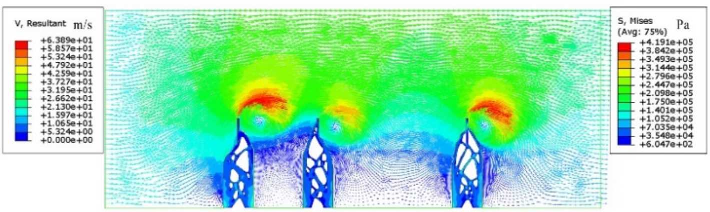
4.2. INTERACTIVE OPTIMISATION AMONG BUILDING CLUSTERS

With modelling the fluid model, FSI-ATO shares the advantage of optimising the buildings clusters due to its ability to deal with the complex fluid interactive behaviours among different buildings. Fig. 5 shows two examples about three buildings under similar settings in Section 4.1. The only difference between these two examples is the non-design domain, which locates at left boundary in Fig. 5 (a) and the top middle parts in Fig. 5 (b).

The optimised designs of the three buildings are totally different in both Fig. 5. It indicates that although with the same macroscopical conditions (wind speed, material assumptions, boundary conditions, etc), the complex air flow among buildings place a huge impact on the architecture forms.



(a) Left boundaries as non-design domains.



(b) Middel top parts as non-design domains.

Fig.5 Wind speed and Mises stress analysis of FSI-ATO structures.

5. Conclusion

This work develops an innovative architectural topological form-finding method integrating CFD with FEA using the FSI technique. This method can effectively balance the influences of the building inner structural performance and the outer fluid field performance. Architects can use it to obtain diverse optimal individual building or building cluster form drafts in 2D/3D. Several numerical examples about single high-rise building, and urban building clusters are also implemented to demonstrate the effectiveness of the method in optimising structure design and building environment. The approach aims to harmonize various physical requirements in the form-finding process within intricate architectural contexts. This innovation holds substantial practical promise for applications in architectural and urban design.

From these examples, it is obvious that FSI-ATO method has many advantages than the previous methods:

(1) FSI-ATO can balance the influence of inner solid mechanics and surrounding fluid environments and generate innovative forms with high performances.

(2) There are less constraints about non-design domains in FSI-ATO method, which means that architects have more freedom in optimising and designing building forms with FSI-ATO.

(3) Benefit from additional CFD model meshes in FSI-ATO, the interactive influences among building clusters can be incorporated into optimization considerations, and as a result, the urban building clusters can be simultaneously optimised.

(4) With more material settings and constraints, FSI-ATO shares the potential of automatically generating other form designs besides the high-rise building facades, such as Taihu stone-like porous structures in water flows.

Due to FSI requires more computing power than the conventional methods, it is essential to researching large scale high-performance computing methods in the future work.

Acknowledgements

Author 1 and author 2 contributed equally to this work.

The authors would like to thank several colleagues whose support helped fulfill the research project described in this paper: Mr. Yulin Xiong, Dr. Zicheng Zhuang, and Prof. Yi Min Xie.

This paper is supported by China Postdoctoral Science Foundation (2023M741955) and Australian Research Council (IH220100016).

References

Bao, D. W., Yan, X., & Xie, Y. M. (2022). Fabricating topologically optimised tree-like pavilions using large-scale robotic 3D printing techniques. Journal of the International Association for Shell and Spatial Structures, 63(2), 122-131.

Bernardini, F., Mittleman, J., Rushmeier, H., Silva, C., & Taubin, G. (1999). The ball-pivoting algorithm for surface reconstruction. IEEE transactions on visualization and computer graphics, 5(4), 349-359.

Duan, C. B., Shen, S. Y., Bao, D. W., & Yan, X. (2023). Innovative design solutions for contemporary Tou-Kung based on topological optimisation. Architectural Intelligence, 2(1), 10.

Edelsbrunner, H., & Mücke, E. P. (1994). Three-dimensional alpha shapes. ACM Transactions On Graphics (TOG), 13(1), 43-72.

Feng, Z., Gu, P., Zheng, M., Yan, X., & Bao, D. (2022). Environmental data-driven performance-based topological optimisation for morphology evolution of artificial Taihu stone. Proceedings of the 2021 DigitalFUTURES: The 3rd International Conference on Computational Design and Robotic Fabrication (CDRF 2021) 3,

Hang, S. (2015). TetGen, a Delaunay-based quality tetrahedral mesh generator. ACM Trans. Math. Softw, 41(2), 11.

Kazhdan, M., Bolitho, M., & Hoppe, H. (2006). Poisson surface reconstruction. Proceedings of the fourth Eurographics symposium on Geometry processing,

Li, Y., & Xie, Y. M. (2021). Evolutionary topology optimisation for structures made of multiple materials with different properties in tension and compression. Composite Structures, 259, 113497.

Lin, P. F. Y. Y. (2019). Research on High-Rise Building Group Morphology Generative Design Method Based on Physical Wind Tunnel and Neural Network Algorithm. Journal of Human Settlements in West China, 34(1), 22-30.

Lorensen, W. E., & Cline, H. E. (1998). Marching cubes: A high resolution 3D surface construction algorithm. In Seminal graphics: pioneering efforts that shaped the field (pp. 347-353).

Ma, J., Lu, H., Lee, T.-U., Liu, Y., Bao, D. W., & Xie, Y. M. (2023). Topology optimisation of shell structures in architectural design. Architectural Intelligence, 2(1), 22.

Ohmori, H. (2011). Computational Morphogenesis: Its Current State and Possibility for the Future. International Journal of Space Structures, 26(3), 269-276.

Sasaki, M., Itō, T., & Isozaki, A. (2007). Morphogenesis of flux structure. Aa Publications.

Song, Y., & Yuan, P. F. (2021). A Research On Building Cluster Morphology Formation Based On Wind Environmental Performance And Deep Reinforcement Learning. eCAADe. Novi Sad, Serbia: Faculty of Technical Sciences, University of Novi Sad, 335-344.

Xie, Y. M. (2022). Generalized topology optimisation for architectural design. Architectural Intelligence, 1(1), 2.

Yan, X., Bao, D., Zhou, Y., Xie, Y., & Cui, T. (2022). Detail control strategies for topology optimisation in architectural design and development. Frontiers of Architectural Research, 11(2), 340-356.

Yan, X., Bao, D. W., Cai, K., Zhou, Y. F., & Xie, Y. M. (2019). A new form-finding method for shell structures based on BESO algorithm. Proceedings of IASS Annual Symposia,

Yan, X., Bao, D. W., Ren, C., & Xie, Y. M. (2023). Constructing topologically optimised spatial structure using innovative mortise-and-tenon joints. Proceedings of IASS Annual Symposia,

Yan, X., Bao, D. W., Xiong, Y., Snooks, R., & Xie, Y. M. (2023). Structural topology optimisation based on a multi-agent model. Engineering Structures, 296, 116978.

Yan, X., Xiong, Y., Bao, D. W., Xie, Y. M., & Peng, X. (2023). A Multi-volume constraint approach to diverse form designs from topology optimisation. Engineering Structures, 279, 115525.

Yuan, P. F., Song, Y., Lin, Y., Beh, H. S., Chao, Y., Xiao, T., Huang, S., Zheng, J., & Wu, Z. (2021). An architectural building cluster morphology generation method to perceive, derive, and form based on cyborg-physical wind tunnel (CPWT). Building and Environment, 203, 108045.

CLICKING IS ALL YOU NEED: IMPLEMENTING WAVE FUNCTION COLLAPSE IN EARLY-STAGE DESIGN FOR MANUFACTURING AND ASSEMBLY PROJECTS

OWEN ZHIYUAN LU[1], LEO LIN MENG[2], CRISTINA RAMOS JAIME[3] and M. HANK HAEUSLER[4]
[1234]Computational Design, School of Built Environment, University of New South Wales
[2]Data-Driven Design, HDR Inc.
[1]Zhiyuan.lu@student.unsw.edu.au, 0009-0006-1683-4714
[2]Leo.meng@hdrinc.com, 0000-0002-6279-9052
[3]c.ramos@unsw.edu.au, 0000-0002-6868-8855
[4]m.haeusler@unsw.edu.au, 0000-0002-8405-0819

Abstract. Wave Function Collapse (WFC) is a constraint-solving algorithm inspired by the quantum mechanics process. However, few attempts have been made in the Architectural, Engineering, and Construction (AEC) industry. WFC literature indicates that it is constrained by its low-fidelity, stochastic process, making it hard to apply in real-world designs, hence its potential lack of application in the AEC sector. Yet this research sees an opportunity in Design for Manufacturing and Assembly (DfMA). Unlike typical architectural projects, DfMA is often more constrained due to modularity. How the DfMA modularity benefits and constricts the spatial planning process, and if such a priori modular definition better informs the design process, is yet to be explored. Thus, how can the highly constrained spatial rules in DfMA architectural design be used in implementing WFC for higher-fidelity fast design concept prototyping? During the research, a prototype was experimented with and implemented while demonstrating several advantages jointly inherited from both the DfMA and WFC, namely (a) high-resolution rapid prototyping with little user intervention for early-stage DfMA and (b) further building material and topological analytics, were enabled for decision support. Hence, this paper addressed the rarely discussed early-stage design problems in the DfMA lifecycle and contributed to a real-world architectural project-based implementation of WFC integrated into an automated computer-aided architectural design workflow inspired by DfMA's modularity that aligns with Sustainable Development Goals (SDGs) of 11 Sustainable Cities and Communities and 12 Responsible Consumption and Production.

Keywords. Wave Function Collapse (WFC), Decision Support Tool, Computational Design, Design for Manufacturing and Assembly (DfMA), Modular Building and Construction.

1. Introduction

The wave function collapse (WFC) is a novel heuristic algorithm based on given spatial constraint rules. Compared to its rarity in AEC applications, it has been widely adopted in the video gaming field for generating mass assets procedurally with little human input. While it offers a simple approach for generating multiple outcomes according to neighbouring rules, the process is noted to be stochastic (Chasioti, 2020; Villaggi et al., 2023). Design for Manufacturing and Assembly (DfMA) is a field with increased interest in academia and the AEC industry for its promise of shortening project duration from the design to construction stages and maximising off-site prefabrication and functions here as a case study for WFC (Laovisutthichai & Lu, 2021). While this research applies DfMA as a framework of exploration, it does not aim to contribute directly to DfMA – thus the discussion on DfMA will be limited however we refer to the following listed references as literature review in the field (Abd Razak et al., 2022; Abrishami & Martín-Durán, 2021; Mesa et al., 2020; Tan et al., 2020; Yuan et al., 2018). In the architectural industry, the typical approach towards spatial planning in early design stages is usually illustrated via massing blocks due to the lack of detailed design information on day one of the projects. Such a technique normally trades off its fidelity in return for speed of prototyping, and it limits the applicability of computational analysis that requires higher model fidelity. We ask the following research question: How can we use WFC and DfMA to (1) rapidly generate higher fidelity spatial models with limited user input, (2) allow the generated model to adhere to given spatial planning constraints, and (3) use the generated model for further geometrical-intense computational analysis? Consequently, the research provides a workflow in early-stage architectural spatial planning that is quicker than massing modelling and generates higher-resolution models, which simultaneously benefits the decision-making process.

2. Literature Review

Wave Function Collapse (WFC) is derived from quantum mechanics. It describes the process where multiple possibilities at each superposition with unique different outcomes are constrained to one single state due to external interactions after an observation (Von Neumann, 1955). It has been widely adapted and applied in the video game industry as a constraint-solving algorithm, due to its simplicity and efficiency in procedurally generating mass assets with flexibility using little human input (Gumin, 2016). Its principle could be akin to an automatic sudoku solver, with multiple possibilities in each grid that are affected by results in adjacent cells and will influence subsequent moves. In this case, numerous building modules could be fitted onto the grid system, similar to placing integers until the sudoku is solved or the wave function is collapsed (Sevkli & Hamza, 2019). It has been rarely discussed in AEC challenges, since only a few attempts were made on considerably low-resolution meshes based on 2D WFC neighbouring rules. For instance, Lin et al. (2020) combined WFC and Convolutional Neural Networks (CNNs) to train computers in the design of urban spaces using the existing urban database and then produced 3D Block models. Similarly, Lioret et al. (2022) assigned 2D WFC rules for creating 3D meshes using Generative Adversarial Networks (GAN), another type of machine learning. On the other hand, Chasioti (2020) used pre-modelled tilesets of various rooms as input to investigate WFC constraint rules in a higher dimension while providing arguably detailed building model results. In the meantime, Villaggi et al. (2023) research based on massing blocks could offer site solutions with multiple buildings and analytical data during its iterative generation process. Nevertheless, as Chasioti (2020) and Villaggi et al. (2023) noted, WFC has limitations that its constraints must be purely spatial, and the generation process is simple but stochastic.

In contrast with WFC, other heuristic algorithms, such as Genetic Algorithms (GA) and Cellular Automata (CA), are often implemented in research on computer-aided AEC fields. Cellular Automata (CA) shares similarities to WFC, it can also generate complex patterns on its grid system following given neighbouring rules, which has been well explored in AEC (Herr & Ford, 2016; Watanabe, 2002). Even though both CA and WFC are both grid-based algorithms, but CA is deterministic while WFC is probabilistic. Meanwhile, GA is also a widely researched heuristic algorithm inspired by the natural process of evolution, in which through iterations of the generation process, the best-fit candidates' parameters (gene) are selected, crossed, and mutated for the generation of the subsequent iterations. As a heuristic algorithm, GA does not guarantee the optimal outcome, but produces near-optimal results within the desired computational power. GA heuristics can be applied to a wide range of problems and optimise towards single and multiple objectives in numerous AEC applications (Latifi et al., 2016; Nisztuk & Myszkowski, 2019; Wang & Wei, 2021). Compared with algorithms for sole generative purposes, GA relies on iterations of generation and evaluation, which is inefficient for the design task assessed in this research. Hence, WFC may exhibit greater flexibility and variation in procedural generation compared with CA, while it can be computed efficiently compared with GA, allowing higher resolution models as cell inputs, as well as multiple possible states in each cell, with the potential of generating models with greater resolution, complexity, and flexibility.

As Tan et al. (2020) argued, Design for Manufacturing and Assembly (DfMA) has become a field of enquiry in academia and the AEC sector. It can be divided into levels such as prefabricated components, volumetric assembly, and modular buildings, according to the Royal Institute of British Architects (RIBA). DfMA highlights a unique construction approach that emphasises off-site prefabrication to the maximum possible extent, while on-site construction is kept to a minimum (Abd Razak et al., 2022). The ability to upgrade, repair, or modify each module or component allows for flexibility throughout the lifecycle and extended usage (Mesa et al., 2020). DfMA is closely tied with BIM integration processes, which Abrishami & Martín-Durán (2021) proposed to model prefabricated building parts as BIM family components, and Yuan et al. (2018) a BIM-based parametric design for DfMA prefabricated components, which can save time in the project design stage. These advantages led to additional gains in terms of sustainability, such as minimisation of material waste and embodied carbon, which aligned with Sustainable Development Goals (SDG) 11 Sustainable Cities and Communities and 12 Responsible Consumption and Production (Laovisutthichai & Lu, 2021).

WFC is an efficient heuristic algorithm for various generation purposes, but there are limited instances of successful implementation in AEC research. The research in DfMA suggests that high fidelity modular DfMA BIM models can be created at family component level, but it still requires manual design and modelling input to assemble these components into an integrated building model, thus may be unsuitable for early design stages where time and design input are limited. Therefore, in what ways can we use WFC to generate high-fidelity building prototypes using DfMA as a guideline that can be beneficial for early-stage spatial planning decision-making?

3. Research Methodology

Action Design Research (ADR) methodology is applied to this research, where four essential steps, including problem formulation, action planning, action taking, and evaluation and reflection are undertaken. This methodology further enables launching of additional research cycles to revise and adjust the initial plan, then act, reflect, and evaluate again to benefit industry-collaborated research, allowing opportunities to check and gain additional feedback from industry experts, ensuring the mentioned research outcome is produced successfully while demonstrating its advantages for the early-stage decision-making process. (Sein et al., 2011).

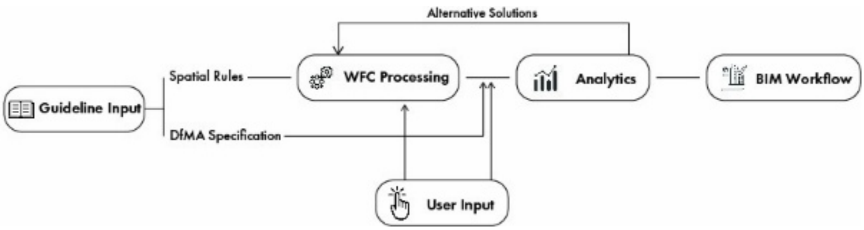


Figure 1 Project Workflow Diagram

During the first ADR stage of problem formulation, this research is partnered with AEC industry practice HDR Inc., to investigate a real-world issue regarding how WFC can be integrated as a new way of generating fast design prototyping. After identifying an issue in AEC practices with HDR, the actions for establishing a backend workflow were planned, and existing literature has been reviewed in terms of understanding the context background, state-of-the-art technology, software feasibility, etc. Then, the workflow outline was discussed and mapped out with the industry partner. Where the proposed tool is developed based on Rhino Grasshopper, including significant features such as DfMA guideline translation from spatial regulation pattern to adjunct rules, WFC processing and generation upon given rules, providing decision support and design analysis based on quick prototyping, along with an iterative process for continuous improvement. The plug-in, Monoceros, developed by SubDigitals is used for WFC processing in Rhino, since the tool is also planned to merge with HDR's BIM workflow via Rhino Inside Revit. During the action-taking stage of ADR, all mentioned features are developed in Grasshopper, under industrial partner HDR's support, regarding both technical and real-world project content throughout weekly progress meetings. This research uses HDR's real-world primary school projects as a case study, developed under New South Wales school infrastructure's DfMA guidelines and principles (DfMA System Guideline, 2020).

4. Case Study

This research uses a primary school design project, a case study in New South Wales

(NSW), Australia to develop and test the WFC DfMA tool following the methodology explained above. The School Infrastructure department of NSW released guidelines to provide a consistent approach across NSW school projects and promote the DfMA technique for its advantages of sustainability, productivity and efficiencies.

Under the guidelines, Primary Schools have specific rules as listed in PS Example Standard Hub Layouts (2021). Projects involve a standard structural grid system of 7.5*9 (x*y) metres. Consequently, rooms with different functionalities can be volumetrically placed and assembled into the system. Within the document, precise spatial arrangement regulations have also been clearly stated. For instance, each Learning Common (LC) must be connected to multiple General Learning Spaces (GLS). LC can be indoor or outdoor, which needs to be placed on the sides of the generated scheme and linked with 2 GLSs as an outdoor LC, or 4 GLSs as an indoors LC. Hence, such restricted spatial constraints allow the potential of implementing WFC to generate rapid building prototyping. Thus, how can the translation from guideline room arrangement regulations to WFC adjacency rules be beneficial to the decision-making process for architectural spatial planning?

4.1. ITERATION 1 - INITIAL EXPLORATION

In the first iteration, the essential aim is to extract spatial arrangement regulations and other parameters for instance the designated number of students, which is decisive regarding school sizing and number of modules required. This would be beneficial towards mapping out decision support logic from the guideline, followed by creating module models for the generation of WFC according to the guideline and testing the algorithm permissibility in the selected software and plugin, Grasshopper Monoceros. These extracted neighbouring rules needs to be scripted and validated, while ensuring the robustness during further modifications in upcoming iterations. For the moment, three different module types have been included, with corresponding 2D WFC adjacent

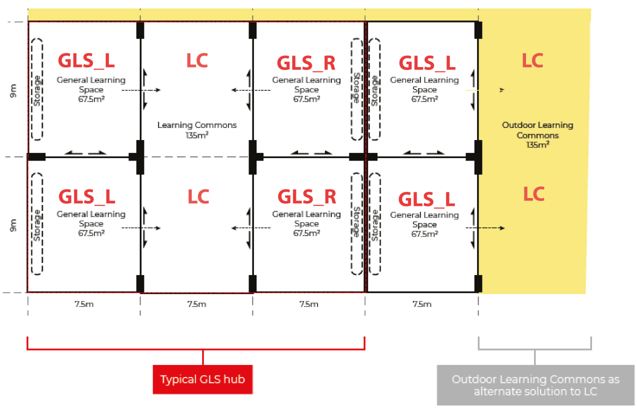


Figure 2 Module Definition to WFC Input

rules, inherited from the typical hub layout sample (fig.2). General Learning Spaces are further categorised into GLS_L and GLS_R. This indicates whether they are located on Learning Commons's left- or right-hand side associated with the typical hub layout. Even though functionalities are identical, they are still spatially different from the modelling perspective. The first iteration's results showed that the tool can quickly generate low-resolution spatial planning options based on the grids controlled by sliders. It also highlighted some issues that merited further investigation: The assigned rules are not precise enough to guarantee that all the generated solutions are deterministic and meet the highly constrained DfMA guidelines. Additionally, the WFC-solving process is currently stochastic and uncontrollable by users, and the idea of decision support hasn't yet been addressed.

4.2. ITERATION 2 - SIGNIFICANT IMPROVEMENT

This stage further implements interaction, data analysis, and geometrical analysis components to provide the necessary functionality required in the conceptual design process. In this stage, module definitions were further divided to allow for more precise adjacency rules, because the established rules in iteration 1 could not guarantee that the produced solutions comply with the mentioned regulation. Besides that, Grasshopper Plugin Human UI and Mouse Rat have been used to create an interface with an interactive grid system to minimise stochastic process. As a result, users can click to generate design options at specific locations on the selected grids. Each click represents the addition of a new module, which the WFC solver will periodically activate to update and adapt to the latest user input. Furthermore, the interface offers information regarding school sizing categories, required GLS and LC numbers, and placed module numbers, subject to the student number input and current generated solutions relating to the guideline regulation. Overall, compared to the first iteration, the second iteration offers more user control over the generation process, rather than being completely random and reliant on spatial rules. Limitations arise when only a minimal level of decision support is addressed, and the current iteration can only produce solutions for a single level because 2D adjacency rules have been used.

Figure 3 Iteration2 Results

4.3. FINAL OUTCOME

WFC adjacency rules are expanded to 3D to generate multi-level spatial planning solutions. The decision-support functionality of the tool has been improved with the capacity to provide beneficial feedback that helps assess each design option during rapid prototyping. For example, benefiting from the modularity of NSW's DfMA primary schools, the number of prefabricated components by each category and their volume, can be calculated given that existing DfMA system information is provided by NSW School Infrastructure. More importantly, building embodied carbon footprint breakdown can also be further estimated, since volume and materiality based on component category, are also assumed subject to the guideline. After that, the Grasshopper plugin conduit is used to provide live analytical graphs, updated every time a new module is placed, to assist users in exploring numerous design options in a short time. In addition, solar access analysis is also available via ladybug, to establish an understanding between the building and its surrounding environment. The tool has been integrated with the industrial partner's developed BIM model generation workflow. Once the design option is explored in the WFC-powered prototyping tool, the user can export the solution to Revit and complete a BIM model within seconds, as part of the automated workflow benefits of computer-aided design (fig. 5). Thus, the research demonstrates the potential of implementing WFC to generate detailed models and addressing highly restricted spatial planning constraints inherited from a DfMA process.

Figure 4 Final Iteration Results

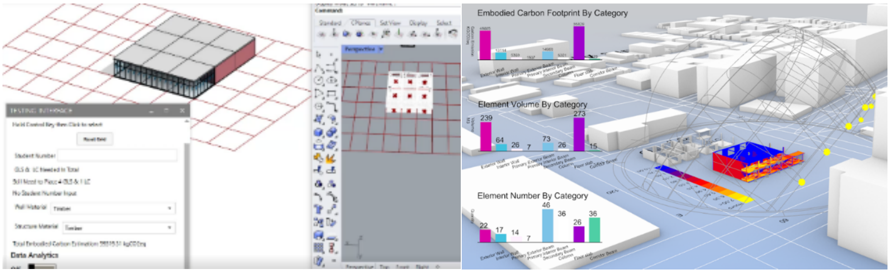


Figure 5 Merge with BIM quick generation workflow & Solar Analysis

5. Discussion and Future Steps

To highlight the aim of this study, A tool that is applicable to real-world AEC practices and their DFMA project workflows has been explored, inspired by both DFMA's spatial modularity, and the novel generative algorithm of Wave Function Collapse's characteristics of simple asset generation process based on constrain solving. This successful attempt indicated by utilising WFC in real-world DfMA projects, the generated model has higher geometrical resolution inherited from the DfMA guideline, contrasting the typical industry practice approach of low-resolution massing. Those high-fidelity models further enable a variety of computational analyses, including prefabricated component counts, building embodied carbon breakdown, and solar access insights, allowing to compare each design option in real time. These details would not usually be accessible in the early design stage due to many factors have not yet been decided.

Yet, the tool is still under development and has some limitations. Firstly, the WFC implementation through the Monoceros plugin is limited to a rectilinear grid system, which prevents more complex building typologies, as curved building blocks, or multiple buildings, from being placed at an angle. Secondly, the control of the WFC stochastic process is yet to be fully addressed. In the case study, we have used DfMA principles and spatial adjacency rules to provide controlled outcomes for the project. This applies to most DfMA projects with associated strictly defined spatial rules and modules, rather than specific to this case study. DfMA can be seen as a special case of building design, and the adaptability for non-DfMA building typology is subject to future testing. Based on the limitation discussed above, a series of further works are proposed: (1) implement WFC on an irregular grid basis for AEC applications and further test irregular modules and angled building placement (2) introduce a global and local weight system to introduce more finetuned user control over probability during tile generation, similar to the generation process proposed Villaggi et al. (2023), (3) include more analytical functions that are relevant to early stage spatial planning, such as traversal distance, wind/ventilation analysis, or regulatory compliance.

6. Conclusion

The study investigates the integration of the novel WFC algorithm into the initial phase of DfMA in architectural spatial planning. Through literature review, the research highlights previous WFC implementations that could be relevant to AEC applications, while also identifying limitations in tileset control. Further review of DfMA indicates that combining WFC with DfMA guidelines may enhance the early-stage architectural design. The study follows an ADR methodology to iteratively develop a tool and tested it in a real-world project as a case study and presents a real-world problem-focused implementation of the WFC technique in the AEC field. By incorporating DfMA principles into the generation process, the developed tool partly overcomes the previous constraint of lack of control. Moreover, the integration between WFC and DfMA enables the production of a high-fidelity model, generating analytics that were previously only possible at later design stages, rather than early conceptual designing via common massing blocks illustration. Compared to existing approaches using CA and GA, this tool significantly reduces generation waiting time with higher model resolution, and simplicity and intuitiveness of input. In addition, the analytical insights provide designers with better decision-making tools, and deliver the potential of data-driven design. In conclusion, the study shows how early-stage manual spatial planning can be streamlined to produce higher fidelity models in rapid speed and provides a novel means of client interaction that is more informative and engaging. It also emphasises the significance of DfMA's continued adaptation as a modern construction approach, aligning with SDG11 of sustainable city and community and SDG12 sustainable consumption and production patterns. Thus, such an automated process demonstrates the advancement in AEC computer-aided architectural design workflow integration and optimisation, indicating new opportunities for real-time design decision support for early-stage design projects.

Acknowledgements
The authors would like to thank Leo Meng for supporting the research project, and thanks to Luka-Luke Jovanovic, Rena Wang from HDR Inc. who have contributed architectural and modelling knowledge.

References

Abd Razak, M. I., Khoiry, M. A., Wan Badaruzzaman, W. H., & Hussain, A. H. (2022). DfMA for a Better Industrialised Building System. *Buildings*, 12(6), Article 6. https://doi.org/10.3390/buildings12060794

Abrishami, S., & Martín-Durán, R. (2021). BIM and DfMA: A Paradigm of New Opportunities. *Sustainability*, 13(17), Article 17. https://doi.org/10.3390/su13179591

Chasioti, E. (2020). Gameplay with encoded architectural tilesets: A computational framework for building massing design using the Wave Function Collapse algorithm. https://www.academia.edu/44870033/Gameplay_with_encoded_architectural_tilesets_A_computational_framework_for_building_massing_design_using_the_Wave_Function_Collapse_algorithm

DfMA System Guideline. (2020). School Infrastructure NSW.

Gumin, M. (2016). Wave Function Collapse Algorithm (1.0) [C#]. https://github.com/mxgmn/WaveFunctionCollapse (Original work published 2016)

Herr, C. M., & Ford, R. C. (2016). Cellular automata in architectural design: From generic systems to specific design tools. *Automation in Construction*, 72, 39–45. https://doi.org/10.1016/j.autcon.2016.07.005

Laovisutthichai, V., & Lu, W. (2021). Architectural Design for Manufacturing and Assembly for Sustainability. In S. S. Y. Lau, J. Li, S. Hao, & S. Lu (Eds.), *Design and*

Technological Applications in Sustainable Architecture (pp. 219–233). Springer International Publishing. https://doi.org/10.1007/978-3-030-80034-5_15
Latifi, M., Mahdavinezhad, M. J., & Diba, D. (2016). UNDERSTANDING GENETIC ALGORITHMS IN ARCHITECTURE. THE TURKISH ONLINE *JOURNAL OF DESIGN, ART AND COMMUNICATION*, 6(AGSE), 1385–1400. https://doi.org/10.7456/1060AGSE/023
Lin, B., Jabi, W., & Diao, R. (2020). Urban space simulation based on wave function collapse and convolutional neural network. *Proceedings of the 11th Annual Symposium on Simulation for Architecture and Urban Design*, 1–8.
Lioret, A., Ruche, N., Gibiat, E., & Chopin, C. (2022). GAN applied to Wave Function Collapse for procedural map generation. *ACM SIGGRAPH 2022 Posters*, 1–2. https://doi.org/10.1145/3532719.3543198
Mesa, J. A., Esparragoza, I., & Maury, H. (2020). Modular architecture principles – MAPs: A key factor in the development of sustainable open architecture products. International *Journal of Sustainable Engineering*, 13(2), 108–122. https://doi.org/10.1080/19397038.2019.1634157
Nisztuk, M., & Myszkowski, P. B. (2019). Hybrid Evolutionary Algorithm applied to Automated Floor Plan Generation. *International Journal of Architectural Computing*, 17(3), 260–283. https://doi.org/10.1177/1478077119832982
PS Example Standard Hub Layouts. (2021). School Infrastructure NSW.
Sein, M. K., Henfridsson, O., Purao, S., Rossi, M., & Lindgren, R. (2011). Action Design Research. *MIS Quarterly*, 35(1), 37–56. https://doi.org/10.2307/23043488
Sevkli, A. Z., & Hamza, K. A. (2019). General variable neighborhood search for solving Sudoku puzzles: Unfiltered and filtered models. *Soft Computing*, 23(15), 6585–6601. https://doi.org/10.1007/s00500-018-3307-6
Tan, T., Lu, W., Tan, G., Xue, F., Chen, K., Xu, J., Wang, J., & Shang, G. (2020). Construction-Oriented Design for Manufacture and Assembly (DfMA) Guidelines. *Journal of Construction Engineering and Management*, 146, 04020085. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001877
Villaggi, L., Stoddart, J., & Gaier, A. (2023). Harnessing Game-Inspired Content Creation for Intuitive Generative Design and Optimization. In C. Gengnagel, O. Baverel, G. Betti, M. Popescu, M. R. Thomsen, & J. Wurm (Eds.), *Towards Radical Regeneration* (pp. 149–160). Springer International Publishing. https://doi.org/10.1007/978-3-031-13249-0_13
Von Neumann, J. (1955). *Mathematical foundations of quantum mechanics.* Princeton, N.J. : Princeton University Press. http://archive.org/details/mathematicalfoun0613vonn
Wang, Y., & Wei, C. (2021). Design optimization of office building envelope based on quantum genetic algorithm for energy conservation. *Journal of Building Engineering*, 35, 102048. https://doi.org/10.1016/j.jobe.2020.102048
Watanabe, M. S. (2002). *Induction Design: A Method for Evolutionary Design* | SpringerLink (1st ed.). Birkhäuser Basel. https://link.springer.com/book/9783764366414
Yuan, Z., Sun, C., & Wang, Y. (2018). Design for Manufacture and Assembly-oriented parametric design of prefabricated buildings. *Automation in Construction*, 88, 13–22. https://doi.org/10.1016/j.autcon.2017.12.021

EXPLORING RULE-BASED DISCRETE TIMBER DESIGN WITH MORTISE-TENON JOINTS

QIANBIN XU[1], FREDERICO FIALHO TEIXEIRA[2] and MARYAM SHAFIEI[3]
[1,2]The University of Queensland.
[3]American University in Dubai.
[1]qianbin.xu@uq.edu.au, 0000-0002-2971-7107
[2]f.frederico@uq.edu.au, 0000-0003-2853-2343
[3]mshafiei@aud.edu, 0000-0002-7505-8245

Abstract. Digitisation in timber-frame architecture offers potentials to revitalise traditional building methods within a contemporary setting. This paper explores an automated design workflow for creating digital timber frame systems, incorporating mortise-tenon joints (MTJs) and principles from traditional Chinese architecture, including Dougong brackets and rural housing. The novelty resides in developing a comprehensive digital library of MTJs in the form of discrete timber blocks, which in return, facilitates the design of adaptable and reconfigurable timber housing. Addressing the classification and algorithmic reproduction of MTJs, the study introduces a methodology for rule-based generative design, which includes the identification of principles governing joint interlocking, automated joint generation, and block combination and aggregation. The approach includes analysing traditional timber jointing types, parameters, and constraints, followed by coding a MTJ library and combinatorial patterns for assembling discrete timber blocks. The resultant discrete timber forms, informed by Dougong's compositional rules and patterns observed in rural timber housing, present the possibility of extendibility and reconfiguration. This research culminates in a digital toolset that enables the fusion of time-honoured MTJs with contemporary computational design, aiming to extend the capabilities of traditional timber frames.

Keywords. Mortise-Tenon Joints, Discrete Timber Blocks, Digital Timber Frame, Combinatorial Design, Rule-Based Generative Design

1. Introduction

As the industry endeavours to diminish the impact of embodied carbon and ultimately pursue a more sustainable built environment future, timber building has gained growing attention as a green alternative to conventional concrete and steel (Reyes et al., 2021). Computational design tools and mass timber construction have been employed to explore innovative formal possibilities of intricate, flexible, and adaptable timber architectures (Retsin, 2019b). This approach thus holds the potential to

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 313-322. © 2024 and published by the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA), Hong Kong.

revolutionise timber structure systems by employing mortise-tenon joints (MTJs) as the primary connection method in traditional housing designs (Qiao et al., 2021).

Historically, the study of the mortise-tenon principle in Chinese timber frame architecture dates back to the ancient Song dynasty. Li Jie authored the Yingzao Fashi (1103), establishing a Cai-fen modular system for timber frames with MTJs (Liang, 1803). Based on Yingzao Fashi, MTJ principles have been investigated in generative design in a limited number of previous studies. Traditional Chinese timber frames and their Dougong brackets were studied by applying the concept of shape grammars and creating a set of rule-based grammars for generating their forms (A. Li, 2001; Wu, 2003). These shape grammars suggested an extension of Yingzao Fashi to the computational field; however, they have not been used to generate innovative designs.

Recent studies highlight the discrete design strategy's potential for innovative timber architectures, using modular blocks numerically understood (Retsin, 2019b). While Retsin's Tallinn Pavilion 2017 and Royal Academy Installation 2019 show this strategy's creative applications, they face reconfigurability limitations due to threaded rod connections. Conversely, MTJs offer interlocking, flexible, and reversible connections, allowing for reducing these restrictions. In the latest studies, Dougong components have been simplified and applied to parametric design (Zhao et al., 2021), but these studies have primarily focused on digital fabrication rather than emphasizing a patterned and replicable design process. Moreover, with the computational analysis of the topological system of the Dougong, the logic of its composition has been developed (Lin & Hou, 2022). While it has generated new types of Dougong design, its capability to generate architectural scale projects remains limited, and the comprehensive investigation of the intricacies of MTJs has yet to be thoroughly undertaken. In order to fill part of this knowledge gap, this paper seeks to propose a novel digital design method that combines computational logic with MTJs to facilitate their effective application in contemporary houses.

Based on the above state-of-art, this research aims to address the complexities and limitations of traditional MTJ systems inherent in the context of digital design transformation. Therefore, to simplify traditional MTJ while maintaining its advantages and to break through the limitations of traditional timber applications in contemporary houses, this research proposes a spatial combination design method based on the geometry and connection logic of traditional Chinese MTJs. This approach transforms MTJs into a shape library in the form of discrete blocks, digitally redefining the shapes and joining patterns of MTJs. Additionally, it develops geometric and spatial connection rules for combinatorial design to enable aggregation, which together with the shape library results in a workflow for flexible timber forms design.

2. Framework and Methods

The research framework is grounded in the principles outlined in our previous work (Xu et al., 2023). Leveraging a rule-based generative approach for a flexible timber design system, this framework consists of three parts: first, developing modular systems for MTJs based on Yingzao Fashi; second, exploring discrete strategies for timber block combinatorics; and third, merging these strategies with shape grammar for MTJ-based discrete timber design. This framework attempts to adapt traditional Chinese timber frame techniques, utilising MTJs in contemporary contexts, thereby developing a digital approach for flexible timber design.

Following the foundational framework, the methods employed in this study provide an in-depth exploration of the principles and modularity of MTJs and rule-based discrete timber design. The paper integrates the algorithmic logic for more complex and varied design iterations based on the establishment of MTJs libraries, along with the Grasshopper Python component developed as a design tool. The focus lies on exploring the combinatorics of discrete blocks and integrating shape rules for more customisable and reconfigurable aggregations of timber forms. These methods systematically reinterpret traditional principles through digital tools, intending to combine historical craftsmanship with contemporary architectural needs.

3. Discrete Mortise-Tenon Timber Design Workflow

This paper proposes a design workflow for discrete timber blocks incorporating MTJs, delineating a comprehensive methodology from the development of an MTJ library to the aggregation of timber forms, as shown in Figure 1. This systematic workflow begins with the identification of traditional MTJs in order to establish a digital foundation. The focus then shifted to exploring the computational attributes of discrete timber blocks suitable for MTJs, emphasising that the combination of blocks adheres to MTJ principles. The third phase integrates the data on various MTJ types and block combinations to automatically generate available MTJ geometries. The final phase extends the block combinations base and incorporates rules extracted from traditional Chinese timber frames to enable the aggregation of timber forms. These forms are then applied in generating MTJ geometries and visualising assembly sequences.

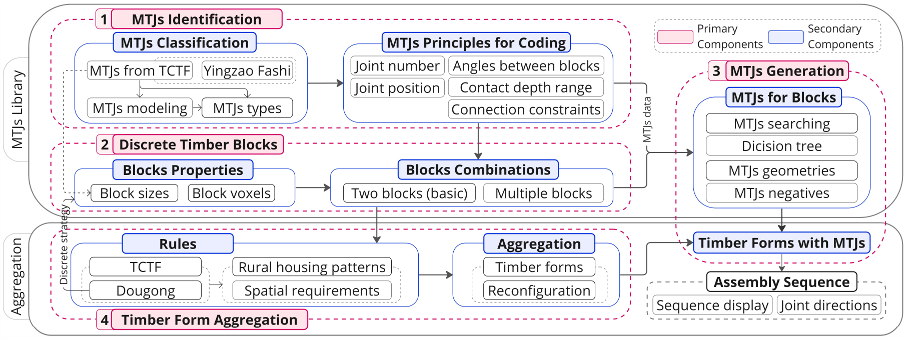


Figure 1. Overall workflow from MTJs library to aggregation of timber forms (figure by author)

3.1. MORTISE-TENON JOINTS IDENTIFICATION

The identification of MTJs forms the foundation of this study, which involves the classification of traditional joints and their subsequent encoding for digital manipulation. This process aims to create a shape library and establish coding principles that are essential for the innovative design and assembly of timber structures. Subsequent analyses will expand on these classifications and principles, bridging historical timber craftsmanship with computational design.

MTJs achieve interlocking through imposition of geometric constraints, enabling reversible connections. In traditional Chinese timber frames (TCTF), various MTJ types can be identified by analysing different connection positions of timber elements. Figure 2a illustrates five basic MTJ types with 0- or 90-degree angles, which support the assembly of straight-shape, L-shape, T-shape, and X-shape for specific joining needs. Additionally, there are three other MTJ types with non-0 or non-90-degree angles allow for connections between horizontal and oblique members, such as beams and cantilevers in TCTF. Different connection positions impact the geometric contact volume that must be subtracted from timber blocks to create concavity or convexity.

Based on these classifications, this paper creates a shape library of MTJs containing the MTJs' geometry for connection logic in the x, y, and z axes. Figure 2b shows the matrix of these shapes records the three-dimensional connection directions for each specific MTJ, corresponding to the axis of the spatial relations between discrete timber blocks for combinatorial design in a later stage. Moreover, these spatial relations determine the suitable MTJs that fulfil the requisites of the jointing sequence and spatial geometric constraints to ensure stable mechanics in the connection.

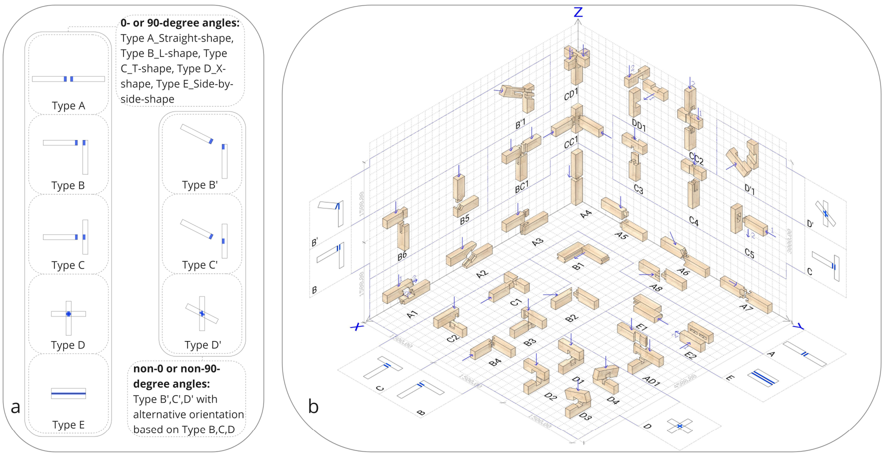
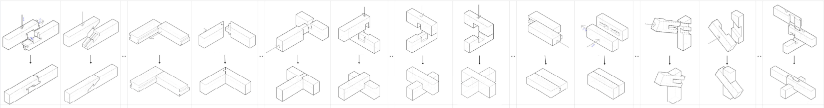


Figure 2. a: MTJs types based on connection positions and angles, b: MTJ shape library with the geometries and connection directions of are sorted on the x, y, and z axes (figure by author)

To facilitate the application of classified MTJs in subsequent discrete design processes and automated generation of joint geometries, this paper converts MTJ shapes into a database equipped with parametric variables. Figure 3 illustrates the data transformation for a selection of MTJ shapes from the shape library shown in Figure 2b, with the approach maintained for other MTJs to systematically establish parametric variables. Different MTJs are distinguished by MTJ types and numbers. For instance, Type A is assigned '0' as the path number, and the first MTJ of Type A is given '0' as the item number; hence, MTJ A1 is encoded as (0,0) for the identification and retrieval of corresponding MTJ parameters during computational processes. The second critical parameter is joint position, identified by marking the contact position on the top, end, or side surfaces of two timber blocks, which will be encoded for computational correspondence in the next step of discrete block combinations. In this paper, angles between blocks are primarily 0 or 90 degrees, but additional angles are also developed and explored for more complex designs. The contact volume range usually includes the whole or half in traditional MTJs, but in this design system, this parameter varies within a range depending on the density or stability required by the stacking blocks. The final parameter documents the geometric or gravity constraints of MTJs on the X, Y, and Z axes, potentially utilised to assess structural performance and the sequence of connection directions. This MTJ data will expand with the refinement of the MTJ shape library to meet specific design requirements.



MTJs Schematic Diagram			..			..			..			..			..			..		..
MTJs Number	A1 (0,0)	A2 (0,1)	..	B1 (1,0)	B2 (1,1)	..	C1 (2,0)	C2 (2,1)	..	D1 (3,0)	D2 (3,1)	..	E1 (4,0)	E2 (4,1)	..	B'1 (5,0)	D'1 (7,0)	..	AD1 (8,0)	..
MTJs Types	Type A_Straight-shape	Type A_Straight-shape	..	Type B_L-shape	Type B_L-shape	..	Type C_T-shape	Type C_T-shape	..	Type D_X-shape	Type D_X-shape	..	Type E_Side-by-side-shape	Type E_Side-by-side-shape	..	Type B'_L'-shape	Type D'_X'-shape	..	Type A_Straight-shape, Type C_T-shape	..
Joint Position	End-End	End-End	..	End-End_side	End-End	..	End-Side	Top-Top	..	Top-Top	Top-Top	..	Side-Side	Side-Side	..	End-Top	Side-Side	..	End-End, End-Top	..
Angles between Blocks	0°	0°	..	90°	90°	..	90°	90°	..	90°	90°	..	0°	0°	..	others	others	..	0°, 90°	..
Contact Volume Range	Whole	Whole	..	Edge	Whole	..	Whole	Half	..	Half	Whole	..	Edge	Edge	..	Whole	Whole	..	Half, Whole	..
Constraints in Axes	X, Y, Z (Geometry & Gravity Constrains)	X, Y, Z (Geometry & Gravity Constrains)	..	Y, Z (Geometry Constrains)	Z (Geometry Constrains)	..	Y, Z (Geometry Constrains)	X, Y, Z (Gravity Constrains)	..	X, Y, Z (Gravity Constrains)	X, Y, Z (Gravity Constrains)	..	X, Z (Geometry Constrains)	X, Z (Geometry Constrains)	..	X, Y, Z (Geometry Constrains)	Y, Z (Gravity Constrains)	..	X, Y, Z (Geometry & Gravity Constrains)	..

Figure 3. Parameter variables for the transformation of partial MTJs from shapes to data (figure by author)

3.2. DISCRETE TIMBER BLOCKS

The concept of part-to-whole in discrete architecture comes from mereology, which involves utilising a limited set of parts to consist of a whole design (Koehler, 2018; Retsin, 2019a). This strategy goes beyond conventional modular design and facilitate the digital manipulation of MTJs in the form of discrete timber blocks. This section explores how encoded timber blocks, related to Dougong and MTJs and adhering to various combinatorial rules, offer the potential to reconfigure spatial patterns.

The Dougong in TCTF exemplifies a highly modular design, achieving outstanding structural stability and aesthetics through the connection of MTJs stacked up layer by layer. Incorporating the discrete approach, the Dougong form is adaptable to flexible timber architecture through discrete blocks. As illustrated in Figure 4a, a section of the Dougong bracket on the left has three tiers with three different lengths of timber, and each horizontal bracket bay is 300 mm according to the 8th-grade timber in Yingzao Fashi. The discrete form of Dougong utilises all timber blocks with the basic size, including cross-section dimensions of 150mm x 100mm and length of 700 mm (the minimum Gong block length from Dougong). The block length is determined by the number of bays and accounts for the subtractive volume required for the MTJs of timber blocks. Additionally, an extra 50mm is allocated to both ends of the block to accommodate the overlapping part. Hence, the block length is calculated by the number of bays x 300mm + 2 x 50mm as shown in Figure 4b. Block lengths of 1300mm and 400mm are also considered for the efficiency and flexibility of combinatorial design. To enable the recognition of relative connection positions of discrete blocks combined with MTJs data, encoded voxels of identical sizes are generated, reflecting the properties of the blocks. As shown in Figure 4c, the sequencing of voxel generation within each block begins at the centre and incrementally extends towards the edges. The numbers assigned to voxels facilitate the matching of with connection relations of various MTJ types during block combinations and MTJ geometries generation.

Combinatorial designs are derived from permutations and combinations of discrete parts (Sanchez, 2016). In the context of discrete timber with MTJs, this involves the application of algorithms based on shape grammars that consider spatial relations such as joint position, angles, and contact volume range. These algorithms correspond to a comprehensive MTJ database, setting the foundation for the specific style of design. Based on this, the paper introduces a custom component for block combinations developed in Grasshopper Python, which operates by employing the types of MTJs as conditions to calculate all combinatorial possibilities and to invoke other MTJs data under that type for spatial operations. As shown in Figure 4d, the primary combinatorial logic translates the spatial relations between blocks for each type of MTJ into relative positions and angles between voxels within the blocks. This is followed by the duplication, rotation, and movement of the initial block to generate and iterate blocks for combination. Taking MTJs Type D as an example, its characteristic connection position at the non-terminal intersection of blocks is interpreted as excluding both ends n and n-1 of the voxels in a block. Thus, the reference voxels for combinatorial arrangements are from voxel 0 to n-2 of the initial block and from voxel 0 to n-2 of the iterative block. Once the list of voxel combinations is obtained, the voxel centre of each iterative block serves as the starting point for movement, with the corresponding voxel centre of the original block as the target point. Additionally, adjustments to various angles and contact volume ranges can be made as required.

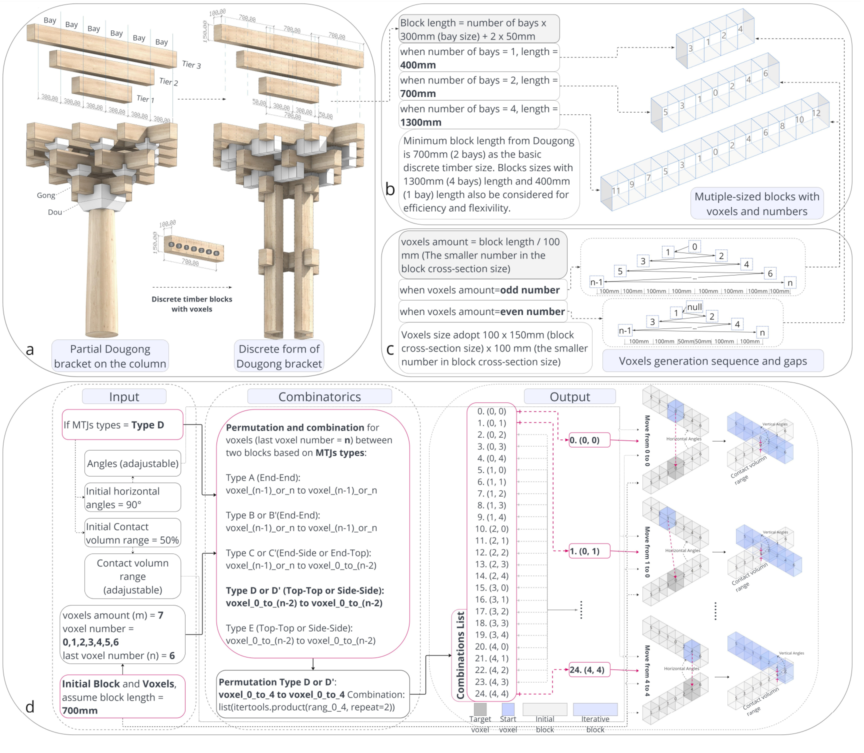


Figure 4. a: Discretization of Dougong by discrete timber blocks with voxels, b: Timber blocks size calculations, c: Voxels generation and sizes, d: Combinatorial Logic of Discrete Timber Blocks linking MTJs types and combination graphs in MTJs Type D as an example (figure by author)

3.3. MORTISE-TENON JOINTS GENERATION

The automated generation of MTJs geometry is based on integrating MTJs data and discrete blocks combinatorics, both derived from the previous two phases. It begins with searching for available MTJ types of intersecting blocks and then proceeds to generate geometries based on the matched MTJs. This process allows to generation of suitable MTJs for discrete timber architecture forms with various block combinations.

A critical prerequisite for generating MTJ geometries is the systematic search for potential MTJ types, crucial for precisely aligning the discrete block combinations. This phase is pivotal as it sets the foundation for the subsequent generation of joint geometries. It requires a methodical classification of intersecting block types, their spatial relations, and possible interlocking patterns. Following this classification, the method for identifying suitable MTJ types for discrete blocks, demonstrated in Figure 5, employs computational representation. This includes categorising contact voxel relations and applying vector mathematics to define block interactions. In terms of voxel relations, the ranges used to filter the voxel contacts between two blocks are synchronized with the list of block combinations to initially determine MTJ types. On this basis, the parallel, orthogonal, or other properties of the block interactions are determined by evaluating the vector cross-products and dot-products, ultimately leading to a match with the available MTJ types. This establishes a clear protocol for identifying various MTJ configurations ranging from simple linear connections to complex intersecting forms. This structured approach ensures that the geometry generation process is both efficient and precise, allowing for the seamless integration of MTJs into the diverse array of timber forms.

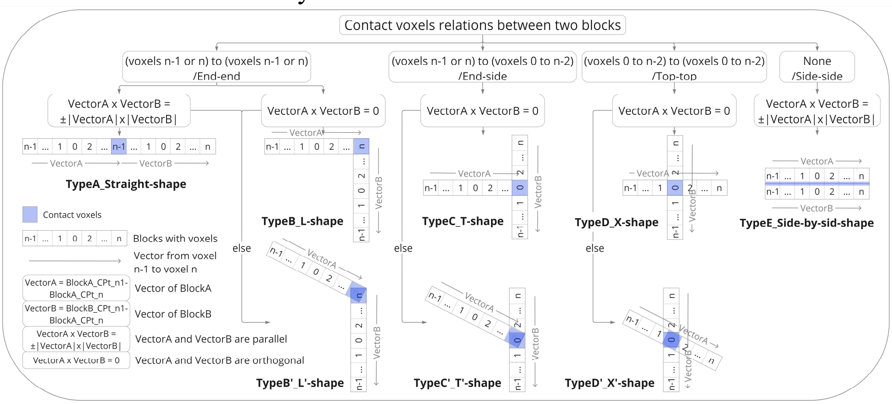


Figure 5. Searching method of identifying available MTJs types for discrete blocks (figure by author)

The proposed automated generation of MTJs relies on the underlying logic based on the principles of MTJs. The whole process commences with the input of discrete timber forms to extract the combinational scenarios for each block. It then proceeds to search for available MTJs within the defined spatial relations, culminating in the algorithmic and spatial manipulation to generate the MTJ geometries. As highlighted in Figure 6, the method initiates by calling the geometric properties of the block and voxel from the combination scenario to obtain the intersecting volume of the blocks and the voxel centroid, which is then used to calculate a plane to split the intersection to get two MTJ negatives. Subtractive operations are subsequently applied to the blocks to generate blocks incorporating MTJs. Owing to its reliance on block volumes and cutting planes calculated from voxels, this method remains applicable for generating MTJs in combinatorial scenarios with varying angles or contact volume parameters.

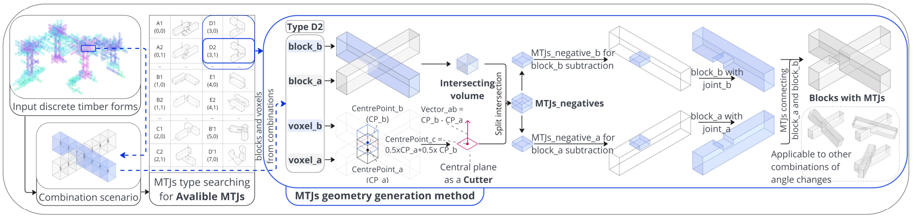


Figure 6. MTJs generation process highlighting generating Type D2 geometries (figure by author)

3.4. TIMBER FORM AGGREGATION

The aggregation of discrete timber forms represents a further exploration of the fundamental combinatorial principles associated with discrete blocks, alongside the Dougong and TCTF rules, to generate coherent patterns by composing architectural elements to the whole timber framing form. Before aggregating to the timber form, it is essential to determine the architectural elements that aligned along horizontal and vertical orientations. The generation of these elements is guided by rules that stem from the topological relations inherent in the highly modular Dougong system. Based on the previously acquired discrete form of Dougong, which exhibits the potential to stack upwards with increasing tiers and extend horizontally, Figure 7a demonstrates the rules of the Dougong pattern. These rules are formulated by encoding the sequential aggregation of discrete blocks in ascending order and connecting the intersections of block pairs using the spatial relations defined in the fundamental combinatorial principles. Additionally, to enhance the efficiency of vertical aggregation, a set of rules for the column pattern has been established, based on the spatial relations between the lowermost horizontal blocks of the Dougong, integrated with vertical blocks as shown in Figure 7b. This allows linking Dougong and column patterns to obtain beam-column-like timber architectural elements, as Figure 7c shows, which also provides the potential for the extension or replacement of discrete blocks for reconfiguration.

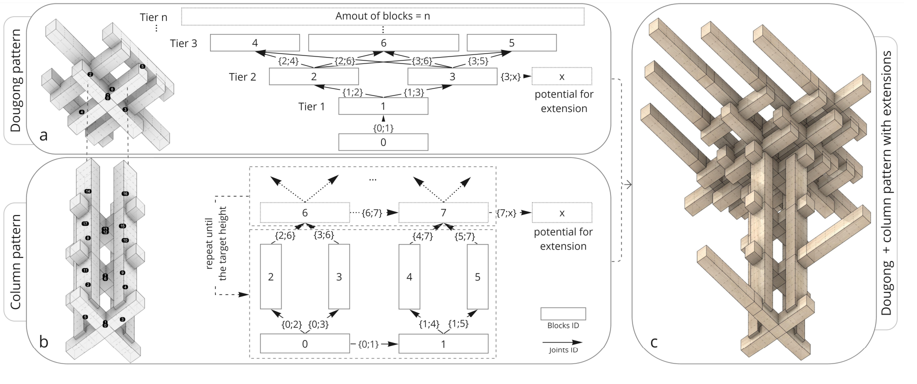


Figure 7. Rules for Dougong-based discrete architectural elements generation (figure by author)

Beyond the scale of the architectural elements, a more comprehensive hierarchy of principles, grounded in Dougong rules, is required to govern the starting, connecting and ending points for timber form aggregations. As an approach for developing these principles and as a platform for testing the proposed design strategy, the aggregation rules are derived from both TCFT and rural housing patterns in China. Typically, rural timber houses feature rooms enclosing various types of courtyards, such as open, three-sided, and four-sided courtyards, as illustrated in Figure 8a. The variation in courtyard patterns can be parametrically interpreted from the relative positioning of rooms in a plan, providing layout outlines for planar generation of timber forms. Moreover, the inflection points of these outlines serve as starting points for upward aggregation in the column pattern. Additionally, a key observation of TCTF is the similarity of Tailiang frame's generation logic to the tiered upward progression dictated by Dougong rules, progressively narrowing towards the apex, as shown in Figure 8b. By extracting this stepwise upward rule, the height of each Dougong pattern aggregation and the extent of its horizontal extension are constrained. This approach shapes the discrete timber frame form, as demonstrated in Figure 8c, integrating the structural logic of traditional rural architecture with contemporary computational design techniques.

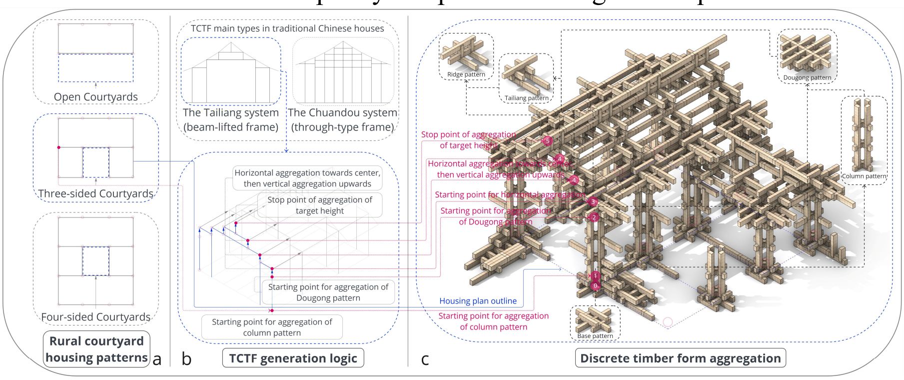


Figure 8. Timber framing logic for discrete timber form aggregation (figure by author)

4. Results

This research innovatively integrates traditional MTJs with digital design, developing a comprehensive method for the design of timber frame architecture. Moving beyond previous studies primarily focused on digital fabrication or aesthetics, this study introduces a detailed and systematic approach that emphasises the utilisation of MTJs to augment digital timber design. The methodology encompasses the identification, digitisation, and application of MTJs in a combinatorial design process, specifically aimed at investigating their functional and structural potential within digital timber architecture. The creation of a digital MTJ library, featuring diverse joint types classified by their geometric and interlocking properties, establishes a versatile foundation for designing adaptable timber forms. This method creates a solution that combines traditional craftsmanship with contemporary design principles to preserve traditional aesthetics while facilitating the automated generation of timber structures. However, the wider application of this approach is potentially constrained by the computational complexity involved in generating and assembling discrete timber blocks, in addition to the requisite further alignment and refinement between computational design methodologies and traditional joinery practices.

5. Conclusion

This research contributes to the development of digital timber architecture through merging traditional mortise-tenon techniques with contemporary computational design strategies. The innovative design approach, developing from the establishment of a digital MTJ library coupled with a rule-based discrete timber design workflow, is inspired by traditional Chinese architectural techniques yet extended through digital technology and design thinking. This approach facilitates novel explorations into adaptability and reconfigurability of contemporary architectural forms. Nevertheless, the transition of traditional joinery into a digital design context and the complexity inherent in the computational processes for multiple combinations of discrete timber blocks present challenges that require further investigation. Overall, this research bridges the gap between traditional MTJs and digital innovations, opening up possibilities for flexible timber architectural design. Additionally, while the system exhibits considerable customisability, further research should focus on refining computational strategies and assessing the scalability and practical applicability of this approach, particularly in the context of rural timber housing.

References

Koehler, D. (2018). Mereologies-Combinatorial Design and the Description of Urban Form. *Computing for a Better Tomorrow-Proceedings of the 36th eCAADe Conference*, Volume 2, 85-94.

Li, A. (2001). *A shape grammar for teaching the architectural style of the Yingzao Fashi* [Doctoral thesis]. Massachusetts Institute of Technology, Department of Architecture.

Liang, S. (1983). *Yingzao Fashi Zhushi*. China Architecture & Building Press.

Lin, H.-T., & Hou, J.-H. (2022). Exploring the Topological System of Dougong. *POST-CARBON-Proceedings of the 27th CAADRIA Conference*, 667-676.

Qiao, W., Wang, Z., Wang, D., & Zhang, L. (2021). A new mortise and tenon timber structure and its automatic construction system. *Journal of Building Engineering*, 44, 103369.

Retsin, G. (2019a). Discrete Architecture in the Age of Automation. *Architectural Design*, 89(2), 6-13. https://doi.org/10.1002/ad.2406

Retsin, G. (2019b). Discrete: Reappraising the Digital in Architecture (Vol. 89). *Architectural Design*. https://doi.org/10.1002/ad.2019.89.issue-2

Reyes, N., Rodríguez, B., Wiegand, E., Zilic, F., Ramage, M., Bukauskas, A., … Ahumada, M. (2021). Achieving zero carbon emissions in the construction sector: The role of timber in decarbonising building structures. *Cambridge Open Engage*.

Sanchez, J. (2016). Combinatorial design: Non-parametric computational design strategies. *ACADIA 2016: Posthuman Frontiers: Data, Designers, and Cognitive Machines*, 44-53.

Wu, Q. (2003). *Bracket study: Textual, computational, and digital* [Doctoral thesis]. Massachusetts Institute of Technology, Department of Architecture.

Xu, Q., Fialho Teixeira, F., & Shafiei, M. (2023). Frameworks in Flexible Timber Design Systems Combining Mortise-Tenon Joints and Discrete Timber Blocks. *Digital Design Reconsidered - Proceedings of the 41st eCAADe Conference*, Volume 1, 283-292.

Zhao, J., Lombardi, D., Chen, H., & Agkathidis, A. (2021). Reinterpretation of the Dougong Joint by the use of Parametric Tools and Robotic Fabrication Techniques. *eCAADe 2021: Towards a New*, *Configurable Architecture*, 233-242.

EXTRACTING ACTIONABLE INFORMATION FROM THE SITE CONTEXT USING A PHENOTYPE-BASED STRATEGY

BANLIANG LING[1] and BIGE TUNÇER[2]
[1]*Singapore University of Technology and Design*
[2]*Eindhoven University of Technology*
[1]*banliang_ling@mymail.sutd.edu.sg, 0000-0002-9855-2860*
[2]*b.tuncer@tue.nl, 0000-0002-1344-9160*

Abstract. As more computational design tools are developed, solution generation has been accelerated to provide real time feedback. However, a human designer is still required to translate generated data into actionable information. This is especially so for diverse design scenarios, where the data structure differs, and the computer is unable to draw conclusions across both scenarios. The site context is one key parameter that contributes towards the difference in scenarios. In short, how can an algorithm extract design-related information from diverse scenarios? To address this issue, a phenotype-based strategy is proposed as a representation method, and it re-parameterises diverse site conditions by focusing on their geometrical properties. Instead of parameterising the site context, street-view images are captured, and Gabor filters extract relevant geometrical properties, such that site conditions with different compositions, forms, and density can be organised. This method quantifies compositional and density-based properties of the surrounding building blocks, thereby enabling the computer to digest generated information and provide design suggestions. A new sample site is then used to demonstrate a query of the phenotype space, where suggestions about solar radiation performance is feedback to a human designer.

Keywords. Performance-based design, phenotype-based strategy, computer vision, site context representation, geometry-to-performance information

1. Introduction

Computational strategies accelerate repetitive actions, increasing the speed of performance calculations. With predictive methods, computational methods can even provide real time feedback (Duering et al, 2020). However, despite the large amounts of data generated, a human is still required to translate data into actionable information. This is because different data spaces are constructed using different structures. As a result, related design conclusions could exist in differently defined data spaces, but the computer is unable to discover them. For example, a previous data space might reveal that a building orientation leads to poor solar performance, yet a computer must

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 323-332.

rediscover this information using a new data structure. This contrasts human designers who can easily bridge information across data spaces. In other words, computers cannot repurpose information from past experiences and are only able to perform actions based on the current data structure. For example, Ampanavos and Malkawi (2022) trained a Variational Auto-Encoder (VAE) to generate high performing solutions using a context prompt. However, they mentioned that when prompts are outside of the training conditions, generated results become unpredictable. Compared to a computer, designers form flexible constraints that allow them to quickly ideate across various design spaces. A designer can therefore identify trends outside of the initial constraints, digest them into design actions and apply these to a new scenario. This also represents the gap between design and computational thinking.

This paper bridges this gap in knowledge using a deterministic method to extract design directions from the site geometry. It also addresses the larger research question: How can an algorithm extract design-related information from diverse scenarios? Looking at the site context, diversity refers to the various combinations of context buildings that affect the design performance. Focusing on the geometry, diversity looks at the composition, form, and the density of context blocks. By extracting geometry-related information from the site, computational design search can 1) avoid low performing solutions, and 2) provide early-stage performance-related suggestions based on past design scenarios. These are demonstrated using 19,683 different site contexts that describe a diverse range of possible scenarios.

With regards to the extraction of actionable information, Brown (2020) compared various computational strategies and concluded that an interactive method provides designers with the best tool for design exploration. On the other hand, a free exploration led to less quality outcomes. These results were based on the premise that design-related information involves an interplay between geometry and performance. Current methods differ upon this interplay with some prioritising geometry diversity, and others prioritising the design performance. Newton (2018) presents an open-ended search example that proceeds while designers interact and select intermediate solutions based on its looks. Through this human-computer interaction (HCI), a designer discovers the impact of façade folds on visual, thermal and condensation performance. While the above example is qualitative in nature, Brown (2020) presents an alternative that works off solution re-parameterisation. Using the re-calculated parameters, a computational analysis is conducted to extract design directions towards higher performance areas. The above represent different methods of extraction for design-related information.

While a computer can construct design spaces using the same parametric representation, a difference in data structure makes it difficult to computationally identify a design logic. This means that when different solutions are represented using different structures, it becomes impossible to compare. The site context has the same limitation as each site context is modelled to different levels of detail and to a different extent. This suggests that a purely parameter-based representation does not facilitate comparison across diverse cases. For algorithms to maximise its potential and tap on the increasing amounts of generated data, building representations should instead be geometry-based so that geometry-to-performance links can be discovered. This will also provide a different approach to the interplay between geometry and performance.

From the perspective of the site context, the mentioned research question poses two main challenges. Firstly, there is the extraction of design-related information. The basis of any extraction relies on the available data. This affects the type of design suggestions that a computer can provide. Secondly, there is the consideration of diverse scenarios. For a designer, it is easy to re-structure constraints and look at a design from a fresh perspective. However, the computer does not have this initiative. Both challenges will be looked at in detail and a phenotype-based strategy is proposed.

2. Phenotype-Based Strategy

In computational design, there are three ways to represent solutions (fig 1): 1) genotype — describes the actions that form a solution, 2) phenotype — describes the geometry, and 3) performance — describes the value of a solution. For the site context, genotypes describe the density and heights of neighbouring blocks as a list of numbers (Ling, 2019), while phenotypes describe the "look" of surrounding blocks using image pixel values (Ampanavos and Malkawi, 2022). Compared to genotypes, phenotypes are not described using their generating actions. This means that new site contexts, generated using different parametric models but are captured as images can be related to an existing pool of site context images. On the other hand, a genotype description necessitates one to keep to certain data structures (eg. number of surrounding buildings for density and heights). It also means that a phenotype-based strategy better allows the computer to appreciate geometrical properties.

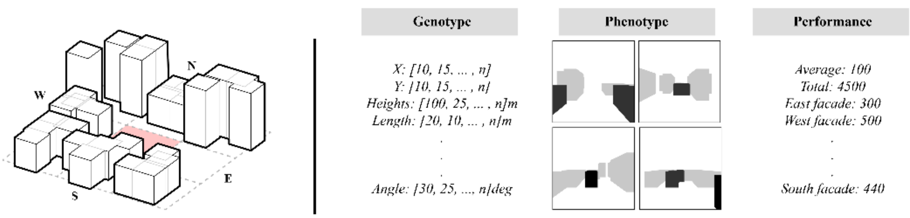


Figure 1. Genotype, phenotype, and performance representations for the site context

Shape composition is one geometrical property that is hard to quantify using a parametric approach. Yet it has an obvious impact on design performance. For example, a close composition of two squarish blocks might have a similar sun shading effect to one long slab block. If a genotype approach is used, this comparison is lost due to them having different data structures. Computer vision techniques offer an alternative to capture and extract desired geometrical properties from images. One technique is the use of Zernike polynomials to measure pixels radially, quantifying compositions in a circular direction and therefore acting as a measurement of shape similarity (Ling and Tunçer, 2022). The authors went on to demonstrate that a sketch could recall similar shapes, all constructed with different numbers of composite blocks. Gabor filters is another technique for capturing pixel compositions. Filters focus on different angles or colour intensity, and they generate a unique identifier for each unique pattern. Stuart-smith and Danahy (2022) used this method to quantify roof design patterns, contributing towards an overall aesthetic measure. Ling and Tunçer (2023) instead used Gabor filters to categorise daylight performance patterns. Given their simple floorplans, the filters covered the eight cardinal directions which was enough to differentiate amongst available patterns. Furthermore, they applied Principal Component Analysis (PCA) to identify dominant data patterns within these daylight images. Geometrical trends that relate to the look of daylight patterns are then extracted and used to influence computational design decisions.

2.1. GENERATING THE SITE CONTEXT DATABASE

The site context consists of context buildings that are built to different levels of detail and extent. Diversity in this sense refers to the block configurations, whereby distance and density result in different performance possibilities for the actual design. This paper will also adopt Gabor filters as the number of filters can be updated according to the level of detail and the defined filters focus on desired properties of the blocks. There are two parts to preparing the context images (fig 2). Firstly, a parametric model populates the land plots. Four square shapes per plot can move in the x- and y-directions, creating shape amalgamations that replicate real life contexts. They are each assigned a height (0, 50, 100m) to imitate an empty plot, medium height block, and a tall block respectively. Across all eight surrounding plots, a unique site condition is then created. Secondly, a camera in the middle of the site takes 1000x1000px pictures of the North, South, East, and West directions. Nearby buildings are coloured black (RGB 0,0,0) and faraway buildings are coloured grey (RGB 200,200,200). This creates a scale for nearby blocks (0) and faraway blocks (200) while differentiating from the background (256). Four images form the phenotype for one site condition. Compared to an overhead shot, a street-view capture was preferred as it does not limit the extent of modelling. For example, wind simulations require a large context and solar simulations work with a small context. The street-view allows a comparison of both types of site conditions, focusing on the impact of immediate contextual buildings.

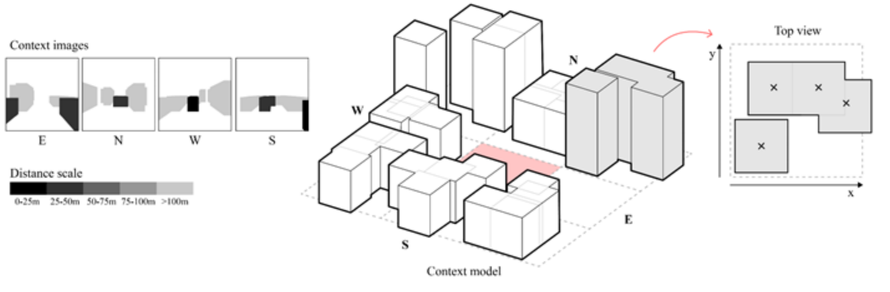


Figure 2. Example of context images (left) and context generation (right)

Two sets of parameters determined the difference in site context. The first relates to placement (x & y shifts) and the second relates to the building heights. Three sets of placements were created and the plots in these three sets were permutated with different heights (0, 50, 100m). Each set would therefore have 3^8= 6561 contexts and 3 sets provide 19,683 contexts. Placement set 1 and 2 have similar placements while set 3 is designed to be different. This relationship is also reflected in the PCA plot in figure 3.

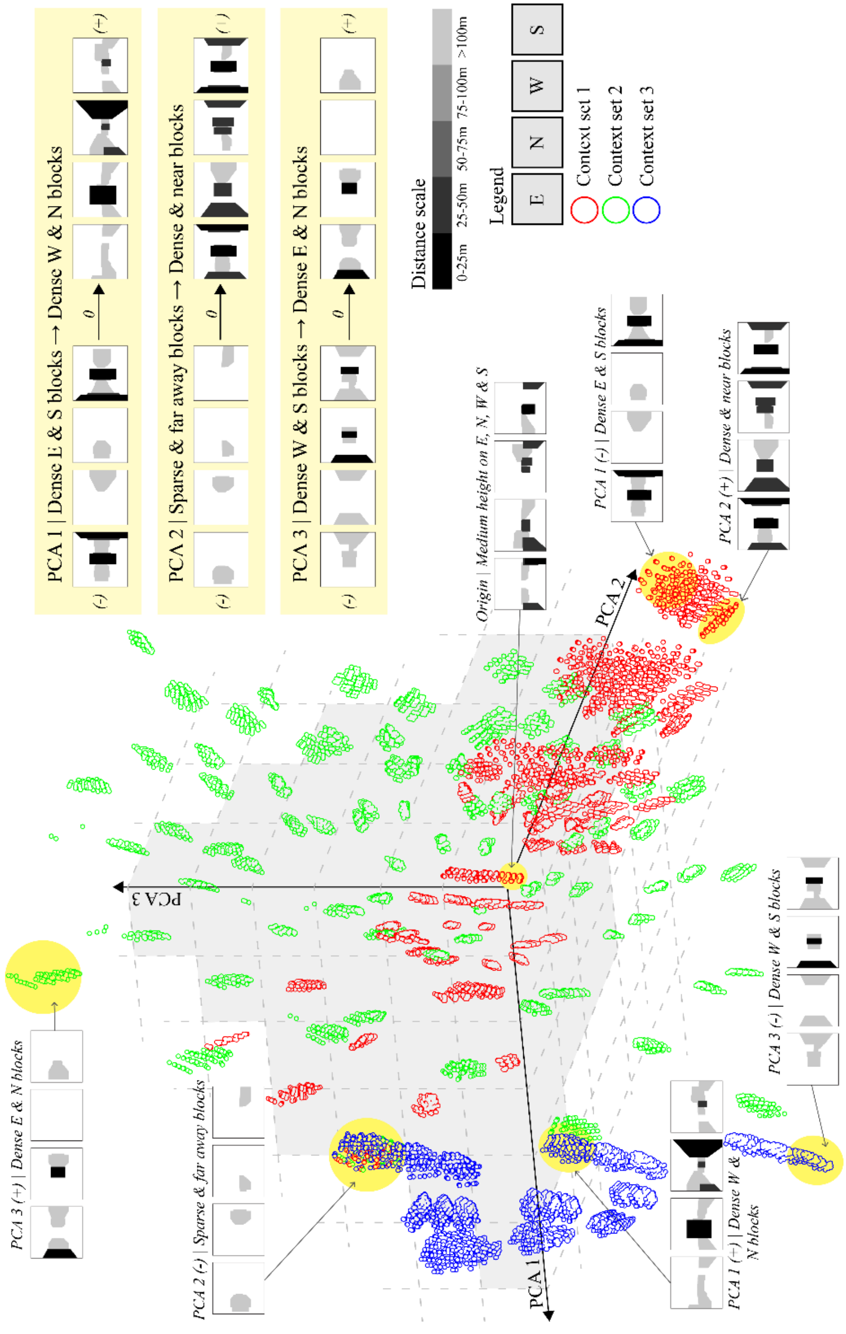


Figure 3. PCA plot of site context with each PCA axis demonstrating different geometrical trends. Similar looking site contexts nearby and different looking ones are further away.

2.2. PRESENTING GEOMETRICAL TRENDS

Geometrical trends should reflect the main types of site conditions within a database. This first layer of organisation must offer an understanding of geometrical information. This includes broad trends like density or compositional changes. A second layer would be the relative contributions of each geometrical trend. It is important to understand the ratio that each trend provides in the current database. Lastly, these trends should ideally be transferable and able to be re-applied into future analysis. PCA offers a solution for all three layers and is therefore presented in this paper. Figure 3 shows a plot of all the site context across PCA 1, 2, and 3. The phenotypes are derived using Gabor filters set at 0°, 90°, 180° and 270° with two variations of sigma (1 & 3) and frequency (0.05 & 0.25). These values account for macro- and micro-level geometrical properties, where a two-block composition has similar macro-level numbers to a slab block, yet they remain differentiated at the micro-level. PCA was conducted on the Gabor filters, and the first three axis accounted for 97.5% variance. Across PCA 1, the site context goes from having denser East and South conditions towards having denser West and North conditions. PCA 2 describes a shift from sparse and faraway blocks towards having densely packed site conditions. Lastly, PCA 3 goes from having densely packed conditions on the West and South, to having densely packed conditions on the East and North. Accordingly, this means that the origin consists of medium-sized blocks on all four directions. These results demonstrate that Gabor filters provide good descriptions of the geometrical properties of the context buildings, and PCA organises the phenotypes to reflect geometrical trends.

More importantly, PCA is deterministic, and results change according to available information. When there is a large diversity of site context, PCA might only extract generic geometrical trends. This is also described as overfitting, and it means that the PCA has not arrived at a conclusive trend which a designer can also arrive at by briefly interacting with the data. Conversely, if the diversity is minimal, the PCA might extract minute details of the site context that does not relate to new scenarios. Also described as an underfitting, the available data is not enough for the algorithm to make an appropriate geometrical conclusion. For the computer to overcome such challenges, the Gabor identities can be re-structured. One option is to cluster the identities such that an initial spread of data can be filtered into focused parts. Another option is to identify outliers that contribute towards overfitting. These are explored further in later sections.

3. Extracting Actionable Information

Building on fig 3, when a new context is under consideration, it can be placed onto the PCA plot. Since PCA is deterministic, the new site context holds a unique PCA position. If the PCA axis provides a good enough geometrical logic, then the new plot point will demonstrate to the designer, geometrical qualities that are similar to existing scenarios. This is in turn helpful for understanding possible constraints for the new site context. For example, if a similar site context had been used for running an optimisation for solar radiation, that scenario will have a substantial set of performance data. The PCA space would then enable the computer to cross reference based on site similarity and have a better understanding of the performance bounds. Using the old data as a basis, designers can now understand what the maximum and minimum solar radiation scores could possibly be. This offers a suggestion into the type of performance diversity that a designer can expect. To demonstrate, clusters are extracted using DBSCAN — density-based spatial clustering of applications with noise. This clustering algorithm was selected as it does not require one to preset the number of clusters and it detects outliers using a distance tree. By altering this distance parameter, hierarchical clusters can be identified. This means that similar looking site contexts that are located nearby in the PCA plot can be hierarchically categorised into representative groups. At the top of the tree would be site contexts that have broad similarities and at the bottom of the tree branch, site contexts still in the same cluster would be highly similar.

3.1. QUERYING A NEW SITE CONTEXT

A new site context (fig 4) based on a town in Tampines, Singapore is added to the PCA plot. The two closest Euclidean clusters were then determined, and the closest site context within each cluster were presented (fig 5). The queried Tampines model is deliberately modelled with a higher level of detail, and the extent has a long proportion compared to the squarish ones within the current database. However, the street view images still demonstrate a good ability to relate the new condition to the database. Furthermore, the images capture shape composition and describe building relationships better than a purely parametric approach. While the new site context has many surrounding buildings, the image focuses on a few significant nearby blocks (black). This is translated into the closest match in cluster 20, which has no blocks on the Southern and Eastern sides. The proportion and position of such nearby blocks (black) are also similar across both cases. Facing North, both sites have a looming presence to its left, and when facing West, both sides have a skewed imposing block to the right.

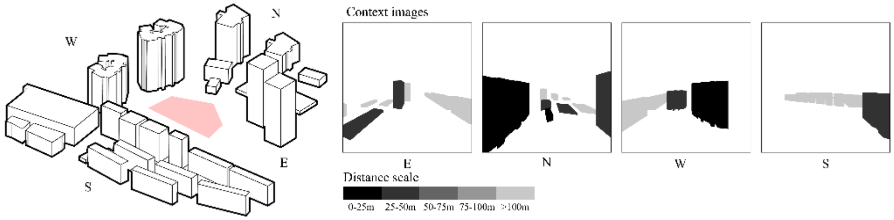


Figure 4. Model and phenotype of the new site context (Tampines, Singapore) added to the database

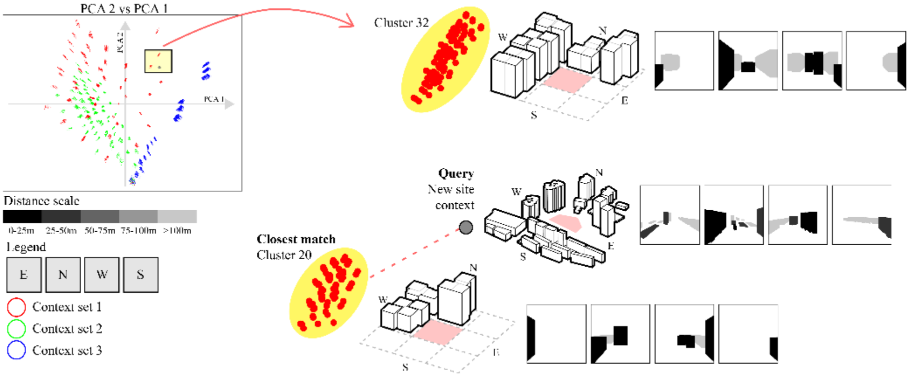


Figure 5. Most similar site contexts to the query option (fig 4) are denser on the N and W sides. This query allows the designer to understand performance ranges based on cluster 20 and 32.

This example demonstrates the computational actions that are enabled with the PCA space. While the street view images create a unified site data space, it also quantifies compositions of buildings and allow the computer to categorise models into clusters. Secondly, these clusters are arranged by geometrical properties that are discovered from the available data. In other words, this visualisation offers information about potential constraints, related to similar site contexts, that might have been unavailable had the design process begun from a clean slate.

Annual solar radiation based on the Singapore weather file is now added as performance data. The building block placed on site is a square block and the façades are captured as images. Using the query in fig 4 as a base, the three closest context clusters are extracted. Fig 6 shows the PCA 1 against PCA 2 plot (90% variance) of solar radiation patterns across all four façades. Since these were extracted from nearby context clusters, patterns are similar. However, PCA 1 still describes an increase in solar radiation on the East and South facing façades. PCA 2 focuses on the East façade, describing a decrease in solar radiation across its axis. Figure 7 is a comparison between the actual solar radiation performance between the query site and the closest matching site. Patterns are most similar on the West and South facing facades, while solar intensity differs on the East and North facades.

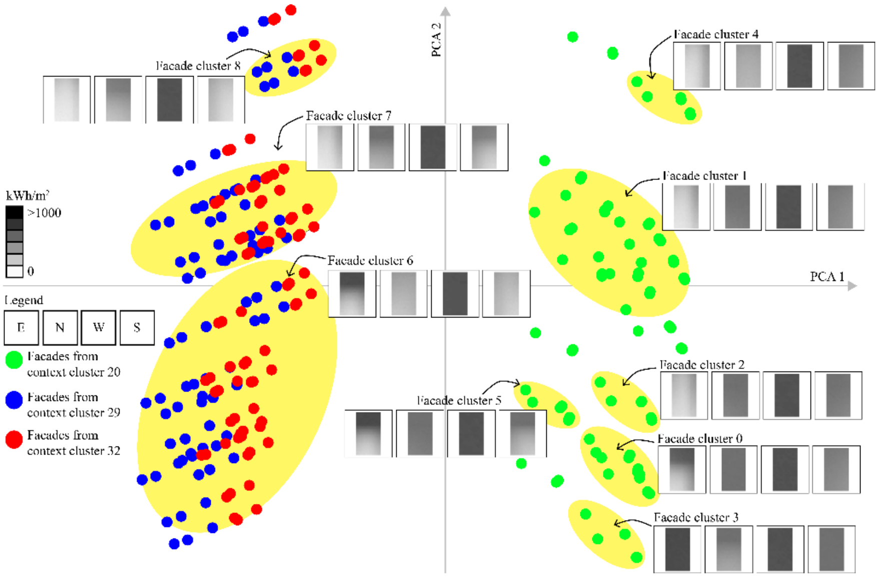


Figure 6. PCA plot of façade solar radiation patterns, demonstrate differences on the E and S facades (PCA 1) and a general decrease on the E façade (PCA 2)

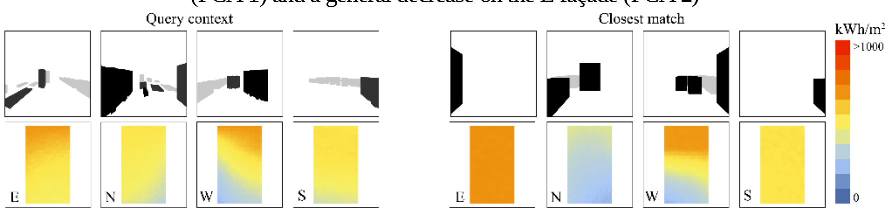


Figure 7. Comparing façade patterns from query site context and closest match

4. Discussion

Paper results aimed to address the research question: How can an algorithm extract design-related information from diverse scenarios? This was addressed from the perspective of site context representation, leading towards a phenotype-based strategy. This unified and consistent phenotype data space was visualised using the first three PCA axis. Being unified, it combines site contexts generated from different parametric processes and allows the addition of new contexts if a set of North, South, East, and West images can be furnished. This data space is also consistent because the Gabor filters always provide unique identities to unique site patterns. Extraction of geometrical trends is also consistent since PCA is deterministic. This is therefore a useful method for classifying site contexts and for relating performance bounds based on past search scenarios. In short, the proposed phenotype-based strategy bridges diverse design scenarios by focusing on the representation of geometrical properties.

4.1. ALTERNATIVE VISUALISATION

Figure 8 offers an alternative visualisation that describes the relationship between context clusters, façade clusters, and total solar radiation performance. The data is based on figure 6 and it uses a parallel categories plot to show links between each cluster type. For example, cluster 20 offers a different façade pattern compared to clusters 29 and 32. This is also evident in fig 6, where cluster 20 facades are all on the right side of the PCA plot. Branching from the façade clusters, there is no clear trend for total solar radiation. Across all categories of solar radiation, there is a good spread of contributors from all previous categories. This is attributed to the façade representation method, where solutions are coloured from white (0) to black (>1000). As a result, white-black patterns take precedence over performance intensity. This contrasts the perception of performance in figure 7, where the blue-yellow-red colour scale prioritises a different aspect of the façade pattern. Nevertheless, figure 8 is still helpful as it demonstrates to a designer, the various façade performance patterns that exist and the range of performance values that they contribute to.

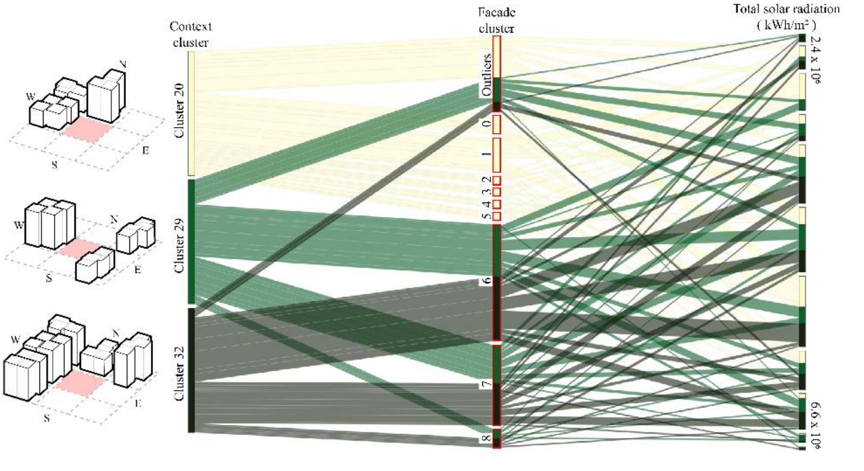


Figure 8. Parallel categories plot between context clusters, façade clusters, and total solar radiation provide a different understanding of the data space as each category is connected using bands

4.2. LIMITATIONS AND FUTURE WORK

While the presented method can relate diverse site conditions, it still depends on image related factors like the camera lens length and the image size. Beyond the preparation of images, the type of colour scale and the selection of computer vision techniques also impact any extraction of design-related information.

This paper has presented the main parts of a phenotype-based strategy, where an image-based representation method is first proposed before computational methods are used to extract design information. The presented camera capture method can be extended to represent design solutions, using images to capture key geometrical properties. On the other hand, phenotypes represent the "look" of any design artifact and a data structure that relates the site conditions, design solution, and performance-related patterns will contribute towards a more refined phenotype-based approach.

References

Ampanavos, S., & Malkawi, A. (2022). Early-phase performance-driven design using generative models. In: *Computer-Aided Architectural Design. Design Imperatives: The Future is Now. Communications in Computer and Information Science, vol 1465.*

Brown, N. (2020). Design performance and designer preference in an interactive, data-driven conceptual building design scenario. *Design Studies,* 68, 1–33. https://doi.org/10.1016/j.destud.2020.01.001.

Brown, N. (2020). Suggesting Design Directions: Early Examples of Simulation-Based Guidance for Common Model Types. In: *Symposium on Simulation in Architecture + Urban Design, SimAUD 2020.*

Duering, S., Chronis, A., & Koenig, R. (2020). Optimising Urban Systems: Integrated Optimisation of Spatial Configurations. In: *Symposium on Simulation in Architecture + Urban Design, SimAUD 2020.*

Ling, B.L. (2019). Early-stage design decision-making informed by a wind performance design space. In: *Urban Tropicality: 7th International Network of Tropical Architecture*

Ling, B.L., & Tunçer, B. (2022). A phenotype-based representation that quantifies aesthetic variables. In: *Computer-Aided Architectural Design. Design Imperatives: The Future is Now. Communications in Computer and Information Science, vol 1465.*

Ling, B.L., & Tunçer, B. (2023). Enabling flexible architectural design re-representations using a phenotype-based strategy. In: *Computer-Aided Architectural Design. Interconnections: Co-computing Beyond Boundaries. Communications in Computer and Information Science, vol 1819.*

Newton, D. (2018). Accommodating change and open-ended search in design optimisation. In: *Learning, Adapting and Prototyping, Proceedings of the 23rd International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2018, Volume 2, 175-184.*

Stuart-Smith, R., & Danahy, P. (2022). Visual character analysis within algorithmic design: Quantifying aesthetics relative to structural and geometric design criteria. In: *POST-CARBON, Proceedings of the 27th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2022, 131-140.*

FORMAL VARIATION EXPLORATION FOR PERFORMANCE-BASED BUILDING MASSING DESIGN OPTIMIZATION

TONG SHAO[1], LIKAI WANG[2] and GUOHUA JI[3]
[1,3]Nanjing University.
[2]Xi'an Jiaotong-Liverpool University.
[1]mg20360016@smail.nju.edu.cn, 0000-0002-5646-3376
[2]wang.likai@outlook.com, 0000-0003-4054-649X
[3]jgh@nju.edu.com, 0000-0003-3237-1263

Abstract. Due to the requirement of design abstraction and simplification for early-stage building design optimization tasks, most applications of performance-based building design optimization are based on the designs generated using orthogonal and cubical geometries, which allows for simpler geometrical operations and sufficient design variability and differentiation in terms of geometrical configuration. However, these applications are only able to produce coarse solutions with room for improvement. In order to address this issue, this study proposes a method focusing on the formal variation in performance-based building massing design optimization to produce more detailed and precise solutions. In this study, two formal variation algorithms are developed using a volume-based and a boundary-based approach, which can modify the input orthogonal geometries produced by EvoMass, a design tool for building massing optimization and exploration. To demonstrate the efficacy of the proposed approach, a case study is presented, which shows that the use of the two formal variation algorithms can further improve the design performance and also allow designers to extract more accurate information related to the design strategies and performative design implications.

Keywords. Design Optimization, Formal Variation, Performance-based Design, Design Exploration, Parametric Design, EvoMass

1. Introduction

With the increasing attention to energy consumption, performance-based building design has become an important topic for researchers and designers. It is widely accepted that early-stage design optimization and exploration play an important role in enhancing overall building performance (Li et al., 2020a, 2020b). Among various factors, building massing forms are one of the most critical factors determining the design performance of buildings. In order to improve the efficiency of performance-based building massing form finding, computational design optimization has been widely applied to relevant research and applications (Wang, 2022b; Wang, Janssen, et al., 2023). A computational design optimization framework typically consists of three

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 333-342.

components: a parametric model for design generation (design generative models), performance simulators for design evaluation, and an optimization algorithm for design evolution. Among these three components, the design generative model defines the design space for the optimization search and has a decisive impact on the information revealed by the optimization result (Wang, 2022a). Two typical approaches have been often adopted in existing studies and applications to build up design generative models.

The first approach focuses on the topological variation, which often creates building massing forms using voxels or cubical geometry. By varying the topological configuration of the building form, this approach can create design variants with significant differentiations, which can further reflect different design strategies or typologies, such as courtyards, self-shading, and solar envelopes (De Luca, 2017; Huang et al., 2015). The optimization incorporating this design generative approach is typically aimed at producing design variants that can convey abstract information related to overarching design strategies or typologies. Moreover, using such voxel and cubical volumes also facilitates geometrical operations for complex topological configurations, such as alignment and gridding, which can ensure the feasibility and rationality of the generated design to be maintained. Nevertheless, this approach is only able to produce abstract but coarse design solutions as the design variation is mostly at the building configuration and topological levels, and as a result, the optimized design often requires further refinement to exploit its performance potential.

The second approach focuses on the design manipulation of relatively subtle and detailed geometrical variations, and it often adopts operations such as twisted, slant, and taper (Chen et al., 2019). This approach can overcome the limitation inherited in the first approach and can produce more precise and detailed design solutions that are capable of harnessing greater performance potential. However, as the overall building configuration and typology remain fixed, using this approach often creates a confined design space for the optimization process to explore. As a result, solely using this approach makes it difficult to help architects achieve a systematic design exploration of different building typologies.

1.1. PAPER OVERVIEW

The above discussion highlights the advantages and disadvantages of the two design generative approaches frequently adopted in performance-based building design optimization. It is evident that these two approaches can complement each other and address the limitations inherent in each approach. In light of this, this study aims to incorporate these two approaches for performance-based building design, aiming to explore how their integration can further enhance the efficacy of computational design optimization during early-stage building design and streamline the process of design exploration and optimization. Herein, this study primarily focuses on a second-phase building massing design optimization workflow based on the design generated by a Rhino-Grasshopper plugin, called EvoMass (Wang, Luo, et al., 2023) with the use of two formal variation algorithms to further diversify the optimized design produced by EvoMass.

EvoMass is a design tool that generates and optimizes building massing based on cubical volumes. Nevertheless, it also holds the limitation mentioned above and is only able to produce coarse solutions for designers. As a result, two formal variation algorithms are developed respectively using a volume-based and a boundary-based approach to introduce additional formal variability to the optimized design. The details of the implementation of the two developed algorithms are elaborated in the next section, with a case study demonstrating the utility of the proposed workflow afterward. Finally, the paper concludes by discussing the relevance of the study as well as the limitations and future research directions.

2. Method

This section describes the workflow that incorporates EvoMass with the two developed formal variation algorithms for building massing design generation and optimization as well as the details of the algorithm implementation.

2.1. WORKFLOW

The proposed second-phase design optimization process is based on the optimization result produced by EvoMass. As mentioned above, EvoMass is a design tool aimed at generating and optimizing building massing designs for various environmental performance objectives. It contains two generative models that create building volumes based on the subtractive and additive principles (Wang, 2022b). Due to operations such as gridding and surface alignment, all the elements used in EvoMass are cubical, and as a result, all generated building geometries are orthogonal. While the optimized design produced by EvoMass can help designers identify promising design typologies and strategies related to energy-saving or passive design (Wang, Luo, et al., 2023), the cubical geometry still leaves great room for further performance improvement (Wang et al., 2019; WANG et al., 2019). To this end, designers often need to manually refine and develop the optimized design to exploit the performance potential, which could be time-consuming and laborious.

Considering the generative approach and the orthogonal building geometry produced by EvoMass, the proposed second-phase optimization connects the optimized design produced by EvoMass with two formal variation algorithms, each corresponding to one of the two generative models. The two algorithms can provide additional formal variations to the optimized design and enhance design performance by refining the building massing form. When using these two algorithms, an optimized design produced by EvoMass is first employed as the input for the corresponding formal variation algorithm. Subsequently, a second-phase optimization process is executed to further enhance the performance of the design through subordinate formal variations. This sequential design workflow enables designers to conduct secondary optimization and provides them with more precise and accurate solutions.

2.2. VARIATION ALGORITHMS

The two variation algorithms: *Volume-Based Algorithm* (VBA) and *Boundary-Based Algorithm* (BBA), correspond to the design produced by the two generative models in EvoMass. In addition, the two algorithms provide a set of user-defined parameters for designers to tailor the generated design, including mass selection, range of control points, displacement distances, etc. (Fig. 1).

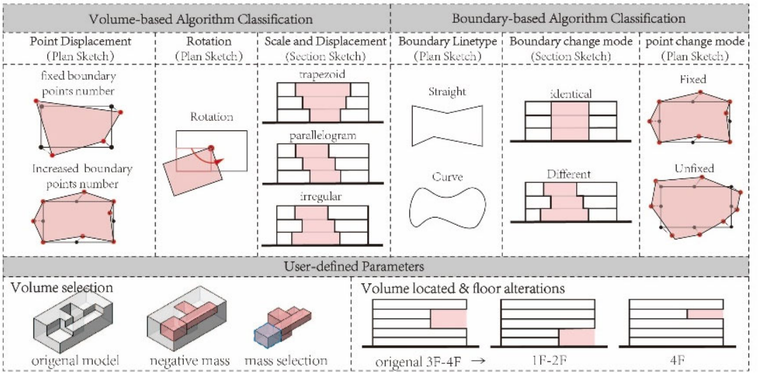


Figure 1 Overview of the Two Formal Variation Algorithms

2.2.1. Volume-Based Algorithm

Figure 2 illustrates the operational workflow of using VBA, which is designed to be connected with the design generated by the subtractive model in EvoMass. First, the negative volume of the subtracted void needs to be extracted, and the selected negative volume is then connected to the algorithm. There are three variation strategies provided for VBA, including 1) Control Point Displacement, 2) Scale and Rotation, and 3) Scale and Displacement. Finally, the varied design can be integrated with an evolutionary optimization for performance-based design optimization. The three strategies are elaborated as follows.

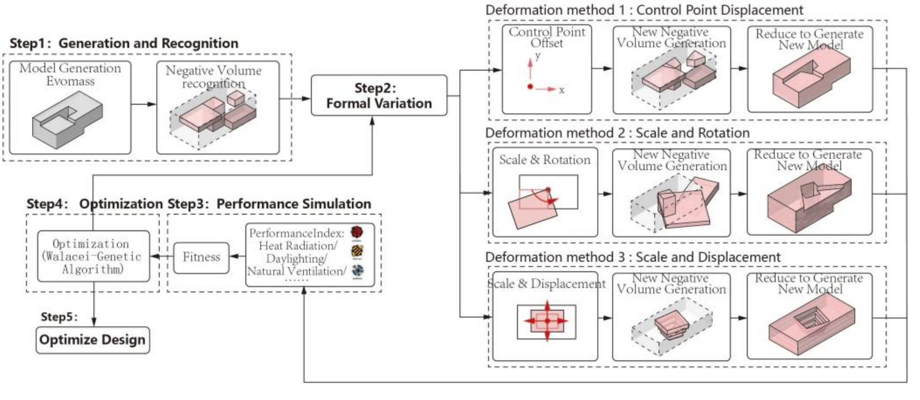


Figure 2 Formal Variation Operations of VBA

Regarding the control point displacement strategy, this strategy alters the geometry by varying the control point position of the selected negative volume (Figure 3). The displacement direction is determined by the position of each control point. The control point at the corner of the exterior building massing boundary is fixed. Displacements alongside the boundary surface are only applied to the control point on the exterior building massing boundary. Free displacements are applied to control points within the exterior building massing boundary. The example of the original geometries produced by EvoMass and the varied geometries are illustrated in Figure 4.

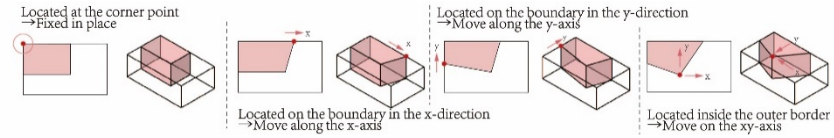


Figure3 Control point displacement strategies

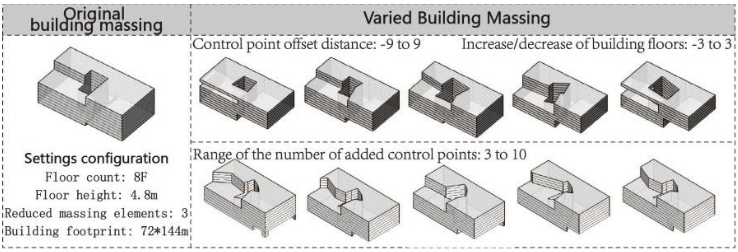


Figure 4 Random Sampling Designs Generated Using the Control Point Displacement Mode

Regarding the scale and rotation strategies, the selected negative volume is scaled and rotated, and the scaling and rotation center is determined according to the relationship between the negative volume and the input building massing exterior boundary (Fig. 5). This strategy can increase the design variability, while the change of the floor area is also relatively controllable. Designers can specify the scaling ratio, the range of rotation angles, and the range of affected floors. Figure 6 demonstrates a group of reshaped designs with the degree of variations controlled by different scaling ratios, ranges of rotation angles, and ranges of affected floors.

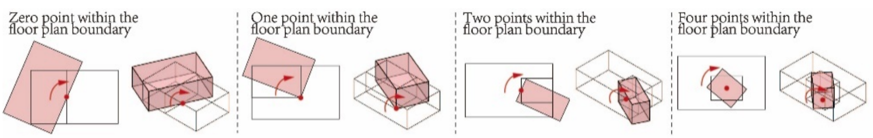


Figure 5. Rotation Center Selection Approach

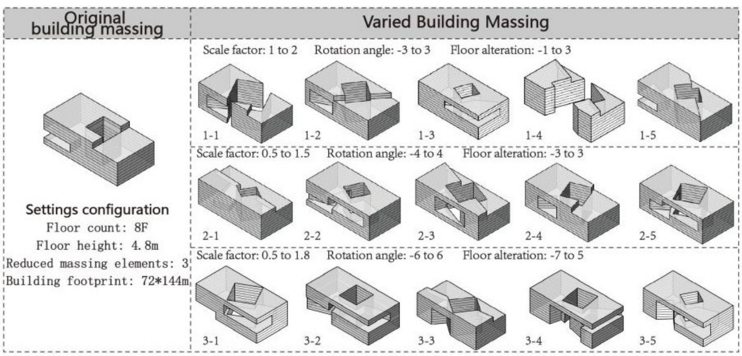


Figure 6. Random Sampling Designs Generated Using the Scale & Rotation Mode

Regarding the scale and displacement strategy, this strategy modifies the building form by scaling and moving the negative volume, and the variation is also applied to the vertical direction based on approaches including rectangular, trapezoidal, irregular, and parallelogram (Fig. 7). Likewise, designers can configure this strategy through parameters such as displacement distance, scaling ratio, and vertical alternation modes. Figure 8 demonstrates the effects of this strategy. As shown, significant change in the building massing form can be observed, and architectural features such as terraces can be clearly identified.

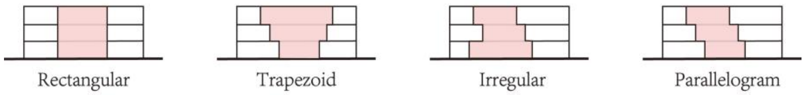


Figure 7. Vertical Variation Modes

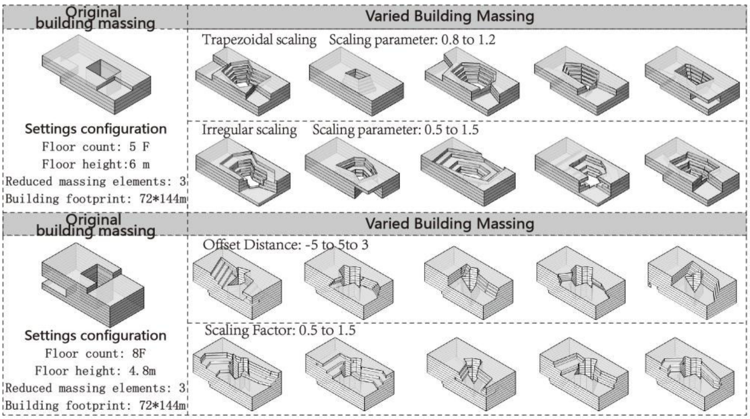


Figure 8. Random Sampling Designs Generated Using the Scale and Displacement Mode

2.2.2. Boundary-Based Algorithm

BBA (Boundary-Based Algorithm) contains a series of steps primarily to handle the building massing generated by the additive model in EvoMass, which, unlikely the design generated by the subtractive model, the sub-volume constituting the final aggregated volume is difficult to separate and extract. As a result, to address this issue, BBA alters the building form using the floor plan boundary of each floor level. BBA first creates a boundary polyline for each floor level, based on which a set of control points are created according to a user-defined interval. The position of these control points is then varied by the algorithm to change the shape of the boundary of each floor level. The use of boundary polylines also enhances the algorithm's generalizability, rendering it applicable to designs generated by the subtractive model as well.

The workflow of BBA is illustrated in Figure 9. In BBA, a set of user-defined parameters is provided to tailor the generated design, which includes the interval distances of the control point, boundary point adjustability mode, and the vertical consistency of the boundary variation (Fig. 10). Figure 11 demonstrates a set of example designs generated using BBA based on the additive models. As shown, BBA can generate buildings with curvilinear shapes. At the same time, using inconsistent vertical variation mode can produce architectural forms with fluidity.

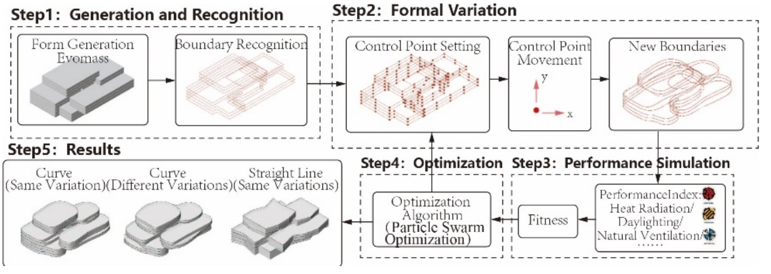


Figure 9 Formal Variation Operations of BBA

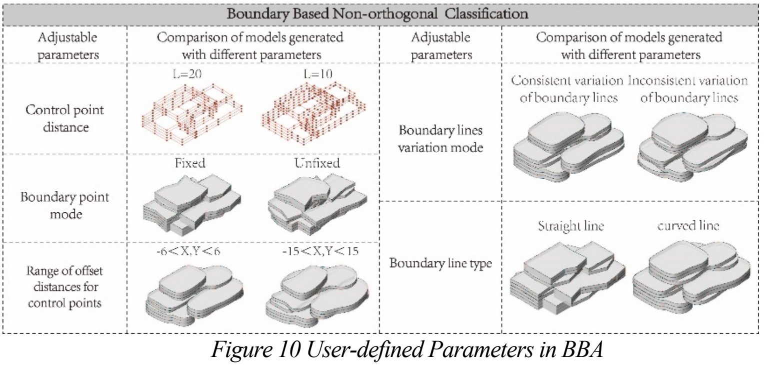


Figure 10 User-defined Parameters in BBA

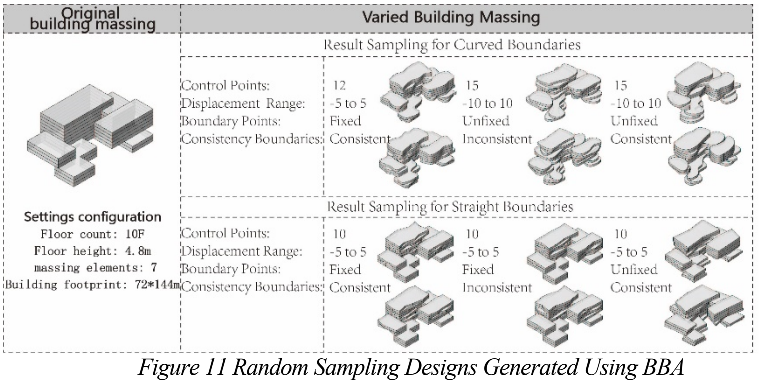


Figure 11 Random Sampling Designs Generated Using BBA

3. CASE STUDY

To demonstrate the effectiveness of proposed formal variation algorithms in performance-based building design tasks, a case-study design is conducted, which describes a teaching complex building in Nanjing, Jiangsu. The existing building is first deformed and then optimized. The total area of the building is 11,000 square meters, 7 floors with 3.5 meters height for each floor. The site, surrounded by multi-story buildings, has complex shading conditions, posing significant challenges for architects to balance various performance indicators using conventional methods.

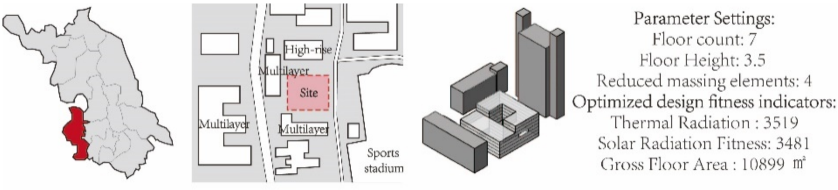


Figure 12 Site Condition and an Optimized Form Generated by EvoMass

Nanjing is characterized by a hot-summer-and-cold-winter climate, which stresses the importance of balancing solar radiation in both seasons, e.g. minimizing summer solar radiation and maximizing winter solar radiation. Thus, the optimization objective is defined as minimizing the difference between the received solar radiance between summers and winters, which ensures that the optimized design either has lower solar heat gain in summer or higher solar heat gain in winter.

The design presented in Figure 12 shows an optimized design produced by EvoMass, which is used for further formal variation. Featuring a central atrium, the selected optimized design contains a stepping terrace on the south side and a cantilevered block on the west side for self-shading. This design can obtain a summer-winter radiation difference of 3481 kWh. On this basis, the formal variation optimization was performed, aiming at refining the facade orientation, atrium size, and self-shading positions of the original design.

In test one, the genetic algorithm in Wallacei was employed. With solar radiation differences (SRD) used as the optimization objective, factors, including the gross floor area (GFA) and the number of floors, are also integrated into the objective function. The optimization produces 800 iterations of design generations and evaluations (8 generations, 100 individuals per generation, 0.9 crossover rate). The scale and rotation mode of the VBA was selected. The top-ranking optimized design, as well as those results on the Pareto curve, are summarized in the upper part of Figure 13.

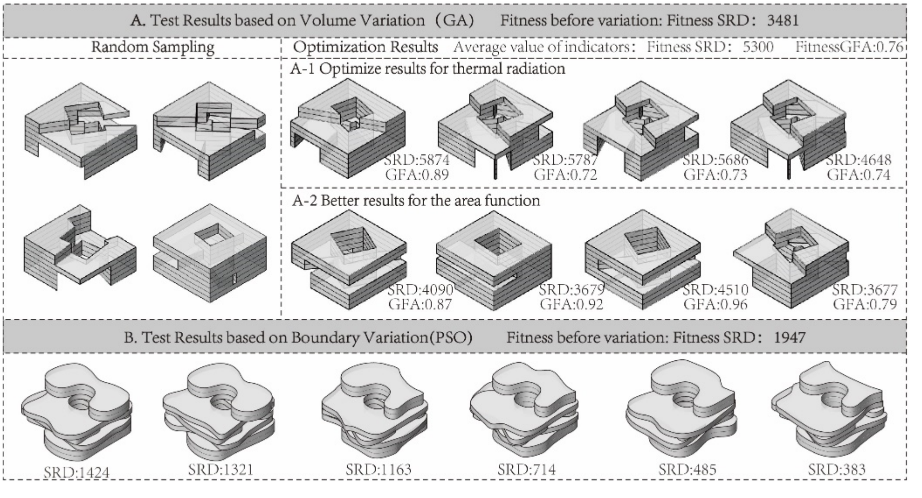


Figure 13 Test Results (SRD: solar radiance)

Overall, the optimized building forms in this test perform better than those originally produced by EvoMass. The optimized designs are characterized by a setback volume with internal corridors (A-1). Large self-shading blocks can be also identified among those designs that can better satisfy the GFA and floor number objectives (A-2). Further comparisons between optimization results and random samplings reveal that the optimized courtyard typically has a 30° clockwise rotation towards the north-south axis, and has a smaller self-shading area on the west side.

In test two, a Particle Swarm Optimization (PSO) algorithm is used, which is specifically designed for minimizing a single objective function. Currently, there is no available plugin having a PSO in Grasshopper, and therefore, the PSO algorithm was imported using Python through Jupyter Notebook to Rhino-Grasshopper. In the second test, with the use of PSO, BBA was used to modify the optimized design, and the optimized results in the lower part of Figure 13 present a significant performance improvement compared to the original design.

In comparison, VBA is able to achieve more substantial design variation by varying the orientation or cross-sectional shape of the negative void, which has a greater impact on performance improvement. In contrast, BBA is more focused on detailed formal variation and is more suitable for design refinement after the design solution has been primarily determined.

4. DISCUSSION AND CONLUSION

This paper investigates formal variation approaches for early-stage building massing design exploration, which is aimed to provide more detailed and precise solutions based on the input abstract design created using orthogonal geometric operations. Two formal variation algorithms are developed on the basis of Evomass, which can be widely adopted in various application scenarios such as courtyards, facades, and self-shading designs. The case study shows that with the use of the second-phase optimization incorporated with the selected optimized design and the two algorithms, the design performance can be further enhanced. In addition, designers can also extract more pertinent information from the refined design solution. Moreover, the two formal variation algorithms also include different geometrical variation modes, which provides greater flexibility for designers to conduct the design exploration using different variation algorithms and variation modes.

While it is possible to integrate the formal variation within the original optimization process based on the orthogonal geometry, one critical issue of using this approach is that there will be a large number of parameters required to control the building massing and the formal variation. This will further cause the failure of the optimization due to the huge design space for the optimization search. In contrast, the proposed second-phase optimization approach can avoid this issue, and at the same time, it also allows the designer to be more involved in the design process, as they can select different solutions for the secondary optimization.

To conclude, the paper presents a study focusing on the optimization for early-stage building design exploration, which is aimed to overcome the limitation inherited in the simplified geometric operation and generation that are often adopted in the initial design stage. The two developed algorithms for formal variations allow designers to further exploit the performance potential of the selected design solution. At the same time, it also helps designers identify more precise and detailed design implications related to the building performance. The case study shows the efficacy of the proposed algorithms, in which the design performance can be further improved by using the second-phase optimization. Regarding future research, more user-friendly and integrated tools are needed to enhance its applicability for a wider range of users. Furthermore, incorporating general formal variation algorithms that extend beyond the geometric aspects generated by EvoMass would contribute to enhancing the proposed approach's generalizability in diverse design scenarios.

References

Chen, K. W., Choo, T. S., & Norford, L. (2019). Enabling Algorithm-Assisted Architectural Design Exploration for Computational Design Novices. Computer-Aided Design & Applications, 16(2), 269–288. https://doi.org/https://doi.org/10.14733/cadaps.2019.269-288

De Luca, F. (2017). From Envelope to Layout - Buildings Massing and Layout Generation for Solar Access in Urban Environments. Proceedings of the International Conference on Education and Research in Computer Aided Architectural Design in Europe, 2, 431–440. https://doi.org/10.52842/conf.ecaade.2017.2.431

Huang, Y. S., Chang, W. S., & Shih, S. G. (2015). Building Massing Optimization in the Conceptual Design Phase. Computer-Aided Design and Applications, 12(3), 344–354. https://doi.org/10.1080/16864360.2014.981465

Li, S., Liu, L., & Peng, C. (2020a). A Review of Performance-Oriented Architectural Design and Optimization in the Context of Sustainability: Dividends and Challenges. Sustainability, 12(4), 1427. https://doi.org/10.3390/su12041427

Li, S., Liu, L., & Peng, C. (2020b). Energy strategy pattern for climate responsive architecture: workflow in the early stages of design. Architectural Science Review, 63(6), 494–506. https://doi.org/10.1080/00038628.2020.1724071

Wang, L. (2022a). Understanding the Span of Design Spaces. In D. Gerber, E. Pantazis, B. Bogosian, A. Nahmad, & C. Miltiadis (Eds.), CAAD Futures 2021: Computer-Aided Architectural Design. Design Imperatives: The Future is Now (Vol. 1465, pp. 288–297). Springer Singapore. https://doi.org/10.1007/978-981-19-1280-1_18

Wang, L. (2022b). 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., Bui, T. D. P., & Chen, K. W. (2023). COMPARING DESIGN STRATEGIES - A system for optimization-based design exploration. HUMAN-CENTRIC, Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA), 1, 221–230.

Wang, L., Janssen, P., & Ji, G. (2019). Reshaping Design Search Spaces for Efficient Computational Design Optimization in Architecture. 10th International Conference on Computational Creativity, 100–107.

WANG, L., Janssen, P., & Ji, G. (2019). Progressive Modelling for Parametric Design Optimization - An Example of How Parametric Design Optimization Can Support Reflection. In T. F. M. Haeusler, M. A. Schnabel (Ed.), Intelligent and Informed - Proceedings of the 24th International Conference on Computer-Aided Architectural Design Research in Asia, CAADRIA 2019 (Vol. 1, pp. 383–392). https://doi.org/10.52842/conf.caadria.2019.1.383

Wang, L., Luo, T., Shao, T., & Ji, G. (2023). Reverse passive strategy exploration for building massing design-An optimization-aided approach. International Journal of Architectural Computing, 21(3), 445–461. https://doi.org/10.1177/14780771231177514

FROM SKETCH TO DESIGN

A Cross-scale Workflow for Procedural Generative Urban Design

JIN GAO[1] and SAYJEL VIJAY PATEL[2]
[1] *Massachusetts Institute of Technology.*
[2] *Digital Blue Foam*
[1]*gaojin@mit.edu, 0000-0002-1595-6895*
[2]*sayjel@digitalbluefoam.com*

Abstract. This study aims to transform urban generative design by aligning it with the symbolic thinking of urban designers, creating an interactive web-based tool. Urban planners and designers use graphic and symbolic thinking. Yet, computer-aided design software accepts inputs by numeric parameters, while recent generative design tools focus on quantity rather than controllability, where in most cases there are "black boxes" hiding the algorithm and procedural logic behind the hood. As a result, there is an established need to tailor tools to the mindset of urban designers to make them more user-friendly and practical. To address this need, we develop a tool that takes the users' intuitive sketch as symbolic inputs to guide 3d urban form generation. It seamlessly combines symbolic thinking and urban form generation using ML-based gesture recognition middleware, and Tensor Field-based procedural urban network generation algorithm. The contribution of this work is: a review of urban generation algorithms and methods of human computation interaction, conceptual overview and methodology, and case study evaluation. We analyzed traditional and state-of-the-art urban design generative algorithms, such as cellular automata-based and machine learning image generation-based methods. It also compared different methods of human-computer interaction and selected appropriate urban design symbol systems. Combining both, the study develops a web-based interactive program backed by a Tensor Field-based procedural urban network generation algorithm, with JavaScript and React framework based on ML-based gesture recognition middleware. The research proposes a workflow for generative urban design. It further evaluates the possibility of integrating block or building scale, with generative architectural design platforms, to automatically refine and analyze rough urban volumes. Ultimately, this work points to a new future of tool development with tools that conform to the logic of urban growth while complying with the designer's mind.

Keywords. generative urban design, procedural modeling, urban semiotics, gesture recognition, tensor fields, creativity support tools

1. Introduction

Urban design emphasizes rational thinking with graphics. Among them, symbols, the abstract expressions of urban characteristics, are the fundamental language forming complex urban design proposals. The process of design is essentially a process of defining symbols. As a shared graphical language, they help urban designers understand and think simultaneously about multiple facets of urban attributes, including urban axis, density, form, function, etc. Urban theories also introduce semiotics as a way to understand the city. In the Image of City (Lynch, 1964), the conceptual "elements" are used to map the perception of cities, while Space, symbol, city (Zhu, 1993) categorizes cities as a collective entity of symbolic elements.

Generative urban modeling has been a long-term interest for planners, designers, and scholars. Due to the varied historical and planning backgrounds of each city, along with their complex and ever-changing nature, there is no universal methodology and tools for procedural urban modeling. The parameters controlling generation vary widely and may involve randomized parameters, user-input text boxes or sliders, or derived from GIS data, which is far from the actual graphical design thinking process. In addition, one set of input parameters often corresponds to multiple results, and it is often difficult for users to control the generation process. However, the designer's way of thinking is based on symbols and sketches, and a process of continuous adjustment based on visual feedback.

Currently, interactive urban design tools that can understand the symbolic language of designers are in the early stages of research. Designers use common architectural modeling workflows, including Rhino and Sketchup, to operate a large volume of masses. The workflow is unable to perform rapid modification and includes unnecessary details that are irrelevant to the core principles of urban design. As Humans are the primary decision-makers in the design process, there is also a lack of intuitive input and output methods that align with designers' symbolic language. These challenges highlight the need for tools that are more intuitive for designers and adaptable across different use cases.

The objective of this study is to propose a cross-platform generative urban design tool that simplifies urban design workflow. This tool aims to bridge the gap between core urban design methodologies and the intuitive, creative process of designers. The significance of this study lies in its potential to make urban design more accessible, efficient, and integrated with existing design workflows.

2. Related Works

2.1. VOCABULARY OF URBAN DESIGN

In both design education and practices, architectural and urban elements are usually interpreted as a "language" consisting of a vocabulary of patterns, as categorized in A Pattern Language (Alexander, 1977) and Concept Sourcebook (White, 1975). In discussions and collaborations, hand sketches and symbol languages are commonly used in all kinds of design processes.

2.2. GENERATIVE URBAN DESIGN

As computational design becomes widely used, scholars have tried to model the generation of urban environments from various perspectives. Learned from the self-organizing properties of urban systems, the Cellular Automata (CA) (Batty, 1997), and its derived SLEUTH model (slope, land use, exclusion, urban extent, transportation, and hillshade) (Triantakonstantis & Mountrakis, 2012) is widely used in regional scale modeling. Many generative models are based on transforming designers' urban planning and design logic into procedural algorithms. The literature (Beirao, Mendes, Duarte, & Stouffs, 2010, Mei, Pan, Cheng, & Garcia Del Castillo Lopez, 2021) encodes the urban planner's decision into parametric moves with which designers can generate organic or grid-like city models with assigned building height and programs based on common urban patterns. Scholars have also introduced architecture-scale computational design methods into urban scale. This includes the tools (Wilson, Danforth, Davila, & Harvey, 2019) to facilitate decision-making by generating correlated economic indicators, and the data-driven method (Olascoaga & Emilio, 2016, 2021) to model the city utilizing shape grammar. Recently, generative modeling methods, such as Machine Learning (ML), have been used in tasks including finding the figure-ground relationship to predict missing urban networks (Boim, Dortheimer, & Sprecher, 2022). However, compared with traditional methods that are based on logic, ML methods face the difficulties of extracting meaningful data from complex urban environments, the computation resources to train and implement models, and unpredictable outcomes.

The tensor-based method, applied in our research, was first proposed by (Chen, Esch, Wonka, Muller, & Zhang, 2008) and further improved and implemented with TypeScript by ProbableTrain (Keir, 2023), who produced an open-source, browser-based city generation program. This method abstracts the city network into a collection of tensors which represents the flow of common urban structures: grid and radius. Environmental factors, such as terrain and coastlines, can affect the tensor field thereby producing continuous results no matter the complexity of environmental constraints. On top of the city's basic grid, we can perform actions including classifying multiple levels of roads and merging city parcels. The advantage of this approach lies in its balance between the randomness of generation and the adherence to urban planning principles, facilitating the generation of cities that are both adaptive to context and well-structured as real-life cities.

This paper focuses on exam:

- How do symbols function as a fundamental language in complex urban design proposals, and how can they be defined and utilized in the design process?
- What are the challenges and variations in parameters controlling procedural urban modeling, and how can a generative urban design tool address them to align with the visual thinking process while providing efficient and adaptable solutions?
- What are the current limitations and challenges in existing interactive urban design tools, particularly in relation to their understanding of the symbolic language used by designers, and how can a cross-platform generative urban design tool bridge the gap between traditional urban design workflow and the creative designing process.

2.3. DESIGN TOOLS AND HUMAN-COMPUTER INTERACTION

Traditional urban design aid tools are mostly developed in the form of plugins set in commonly used professional software (Rhino, Sketchup, and ArcGIS). Exemplified works include the DeCodingSpaces Toolbox (Koenig, Miao, Bus, Knecht, & Chang, 2017) and the Urban Network Analysis Toolbox (Sevtsuk, 2021) as Rhino plugins, and the CityEngine (Kelly, 2021) as standalone desktop software based on the ArcGIS ecosystem. These tools have advanced analysis and generation capabilities, but their use requires dedicated training and professional knowledge. Their inability to cross-platform requires users to manually configure the running environment and rely on the host software for data exchange.

The pursuit to enhance the interactivity and ease of use of professional urban design tools has been a long-standing endeavor. Research groups represented by MIT Media Lab have proposed multiple forms of urban interactive modeling tools, such as the Augmented Urban Planning Workbench (Ishii et al., 2002), and CityScope (Alonso et al., 2018). These projects often feature a backend that simulates the built environment, including wind analysis, crowd dynamics, and traffic flows. Then, using an interactive frontend, such as projection combined with a touchable three-dimensional interface, to provide feedback on people's operations. With an emphasis on user interaction, these explorations enabled more dynamic and participatory approaches to urban design that foster a more inclusive and democratic urban planning process. The challenges they face are the sophisticated setup of hardware-software systems and limited use cases can make the technology less accessible.

Recently, more urban design tools have been developed using web technology. Thanks to the advances of WebGL and front-end frameworks, building cross-platform programs based on web pages is becoming more agile and efficient. Sophisticated functions, like advanced AI engines, can be implemented through high-performance servers running on the backend servers while maintaining lightweight, responsive user interfaces on the client side. These tools are often conveniently offered as Software as a service (SaaS), providing opportunities for business success. Representatives include Digital Blue Foam's Generative Design tool (DBF, n.d.), Delve (Delve by Sidewalk Labs - Real Estate Generative Design, n.d.), and KPF Urban Interface's Scout (Scout, n.d.). The emergence of web-based urban design tools represents an evolution in making them cross-platform, out-of-the-box accessibility, and SaaS-friendly.

3. Methodology

3.1. DEFINE GESTURE LANGUAGE SYSTEM FROM USER STUDY

To understand how designers utilize symbolic language in their sketches, we collected and analysed samples of urban design diagrams, and conducted experiments under a controlled setting by asking graduate-level urban design major students to draw interpretations of specific sites using their own symbolic language. We then categorized commonly used symbolic languages into basic shapes like stars, squares, and circles, to define a system of symbols. As samples listed in Figure 1, each of these symbols is meant to represent different types of influential parameters in urban design, including height, density, and program, which eventually shape the physical and functional aspects of the generated urban spaces.

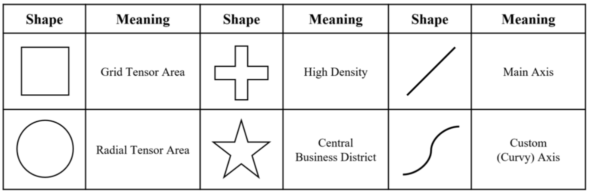


Figure 1. Example of urban design symbolic language

3.2. CROSS-SCALE WORKFLOW

The research proposes a cross-scale workflow for generative urban design. As shown in Figure 2, the workflow takes the user's intuitive ideas and context information as input, generates urban design scale massing, and then zooms into a plot to perform architectural scale massing design and room division. Then, both the urban scale data: proximity, accessibility, porosity, FAR, and architectural scale data: floor area, structure, material usage, and energy efficiency, are evaluated through a comprehensive analysis process. Eventually, the generated project can be output as standard 3D models, allowing for refined rendering or detailed analysis with dedicated professional software, such as structural calculation and performance reports.

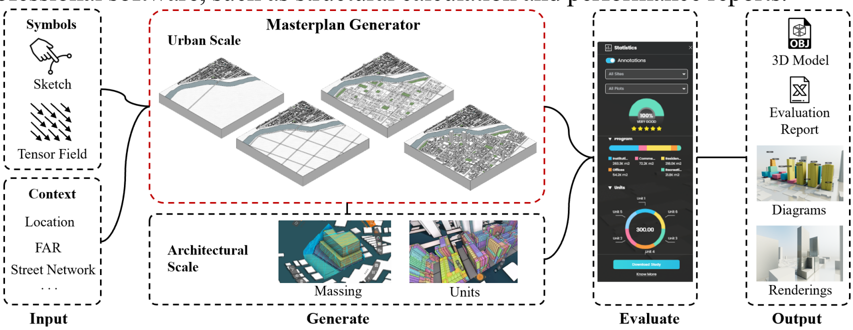


Figure 2. A Generative urban design workflow

3.3. PROTOTYPING OF THE TOOL

3.3.1. Software Architecture

The prototype is a web application that mainly features urban masterplan generation. Shown in Figure 3, it mainly consists of three components, a generation engine running at the backend, a render engine to visualize the generated models, and a frontend framework taking users' inputs. Between each component, the information is interchanged through inner APIs in the format of JSON objects.

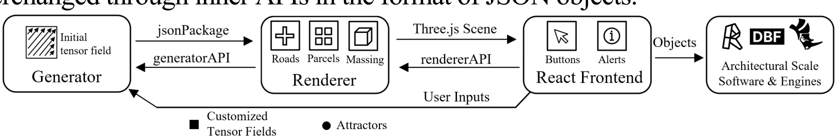


Figure 3. Software Architecture

3.3.2. Symbolic inputs and gesture recognition

The prototype featured symbol recognition of user sketches using the $1 Unistroke Recognizer Library (Wobbrock, Wilson, & Li, 2007). $1 represents the machine learning algorithm that uses an instance-based nearest-neighbor classifier. The hand sketches are translated into attractors that can be read by Tensor Field. Just like designing on paper, users can switch between tools such as pencils, erasers, and lines through the graphical buttons. Sketches will be automatically recognized and translated as symbols with their position and size. Furthermore, these symbolic languages serve as inputs to the multi-dimensional grid that stores calculated parameters to influence the generative process. Details will be explained in the next paragraph.

3.3.3. Urban Environment Generation

In our study, we explored state-of-the-art methods for generating urban networks and selected tensor-based methods due to their controllability, expandability, predictability, and high performance. Our prototype incorporates the urban generation module from the open-source program MapGenerator (Keir, 2023), which is based on the algorithm proposed by (Chen et al., 2008) and implemented in TypeScript. We further developed a control layer to take user inputs to influence the generation process. A multi-dimensional grid with nodes is used to record urban parameters. After receiving user input for tensor locations, and symbolic information, the program generates the 3D models in real time based on the tensor field and corresponding parameter grid.

Step 1 Environment Definition: Within a set rectangle boundary with a range of 1000m by 1000m (can be modified), the algorithm first randomly generates a coastline that defines the boundary between land and water. The boundary serves as the base input for the context of the urban layout and influences the overall structure of the tensor field, therefore influencing the generated urban network. The tensor field can be represented mathematically by (Chen et al., 2008):

$$R = \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix}$$

$where\ R \geq 0\ and\ \theta\ \in [0,2\pi).$

Step 2 Urban Network Generation: The street network is modeled as a graph G = (V, E), where V represents a set of nodes (intersections and endpoints) and E embodies a set of edges (streets). Initially, the site has a default grid tensor field. Users can introduce attractors by adding symbols. The attractors represent different characteristics of the base tensor field. Currently, the base tensor field has grid and radial types. The algorithm computes gradients between each tensor field defined by the attractor and updates dynamically. The grid field is defined as (Chen et al., 2008):

$$\ell = \sqrt{v_x^2 + v_y^2} \quad \text{and} \quad \theta = \arctan\left(\frac{v_y}{v_x}\right), \quad T(p) = e^{-d\|p-p_0\|^2} \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix}$$

For the radial pattern, the basis field is given by:

$$T(p) = e^{-d\|p-p_0\|^2} \begin{bmatrix} y^2 - x^2 & -2xy \\ -2xy & -(y^2 - x^2) \end{bmatrix},$$

$where\ x = x_p - x_0\ and\ y = y_p - y_0.$

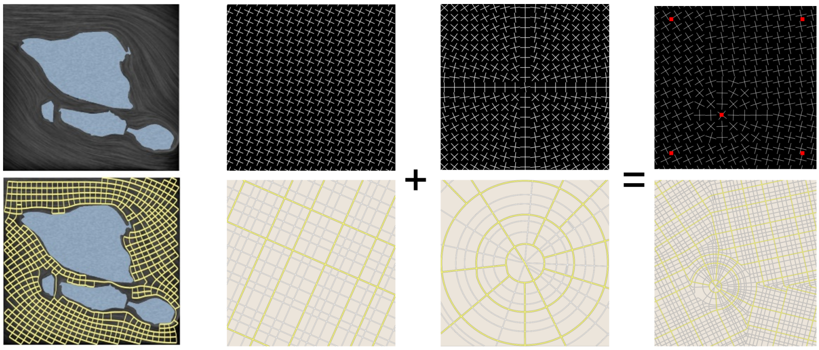


Figure 4. Transform the tensor field into street graphs (Chen et al., 2008)

Figure 4 describes how the algorithm finds street networks as streamlines within tensor fields. The street network is then divided into two hierarchies: major and minor roads. The hierarchy is defined by the distance and numbers between streamlines.

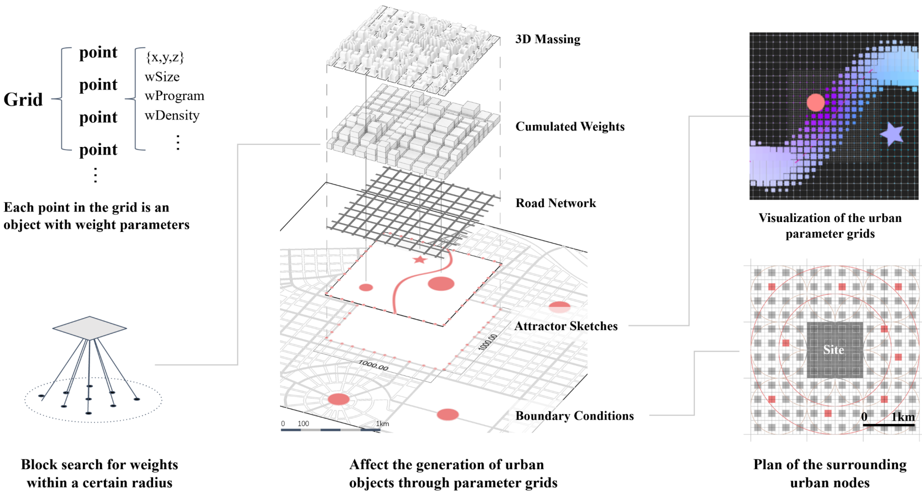


Figure 5. 3D Massing Generation

Step 3 3D Massing Generation: Shown in Figure 5, the 3D massing is further defined by a grid system that incorporates various weighted parameters based on user inputs and surrounding contexts. For instance, a "star" shape is interpreted as representing a "central area". It will have a high influence on the density and program score of the grid parameters around its central point. Then, based on the combined influences of scores, the building base plane will be assigned a weighted height and program parameter that will be represented by the final massing outcomes. Currently, the massing is extruded from the base planes. For further development, the massing and its corresponding parameters can be exported to architectural scale generative design tools, allowing further massing division and detailed architectural design.

Step 4 Visualization and Representation: The prototype's visualization component is powered by Three.js, a widely used WebGL library, to achieve a high-performance real-time rendering of the generated urban environment. The React framework is used to create a control layer that receives user sketches, with the dat.GUI as general parameter input panel. The program featured augmented representation by adjusting material properties based on inherent urban parameters. Furthermore, the user can switch between display modes (perspective and isometric) and perform real-time rendering by additional libraries like three-gpu-pathtracer (Gkjohnson, n.d.). Figure 6 shows the live interactions a user creates with the prototype.

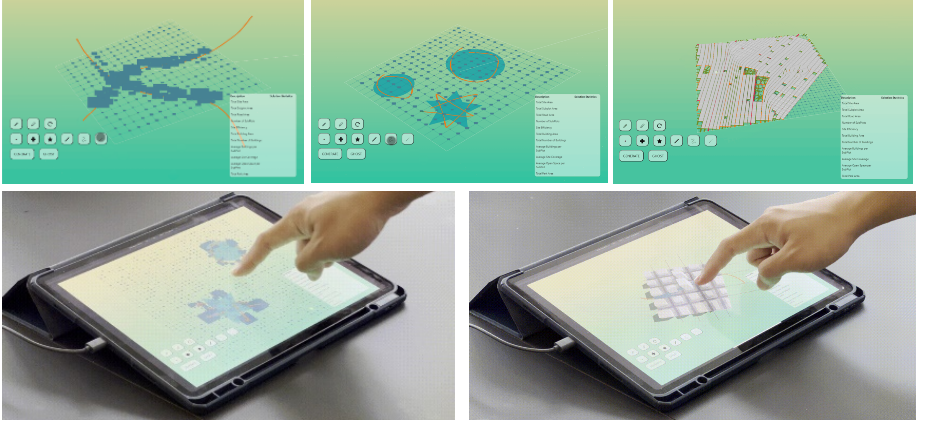


Figure 6. Real-world Use Cases

Step 5 Post-process of Generation Results: The generation results are designed to be exported as compatible formats for architectural scale generative software, ensuring further manipulation and refined design by the next step of workflows. Additionally, the model can be exported as a standalone entity and imported into common modeling software utilized as the foundational model for proposal graphics and further rendering.

4. Discussion and Future Works

4.1. DISCUSSION

Unlike traditional urban design tools or modeling software that often rely on intricate inputs and non-dedicated workflow for urban designers, this tool proposes an intuitive approach by interpreting designers' symbolic languages and sketches into the computer-aided urban design workflow. The tool stresses interactivity, which is fundamental in fitting the actual needs of designers. Currently, with more advanced architecture-level generative design tools coming out, the research aims at filling the gap in urban-scale design tools and allows a seamless transition from urban-scale design to detailed architectural design. The tool can be used in many use cases: First, it can be used by urban planners and designers to quickly conceptualize and manipulate urban layouts, and preview the outcome in a cross-platform, interactive user interface. In academic settings, it can be used as a teaching tool to demonstrate key principles of urban design. The students can experiment with multiple design approaches with an easy learning curve. The tool can also contribute to the participatory design process by allowing non-professionals to contribute to design discussions in a more interactive and intuitive approach.

4.2. FUTURE WORKS

As the prototype is still in its development phase, there are several key areas of enhancement planned for the next phase to make the proposed workflow more robust and adaptable. First, import real-world map data as context, to provide an adaptative urban network based on surrounding environments. Defining more types of tensor fields will make the tool more adaptable in representing multiple typologies of urban networks. Furthermore, by expanding and customizing the symbol language system, we can make the tool adapt to a broad range of users. We foresee opportunities below:

- Seamless cross-scale design tool: Currently, the prototype is a middleware between the user's input and architectural scale design software. This barrier can be broken by establishing dynamic links and automatic sync between tools.
- Collaborative symbol language-aided design: Like real-world design practice, there are usually several people sketching on the same piece of masterplan together. With the collaborative editing feature, multiple designers or community members can discuss and contribute to the design outcome under a shared platform.
- Multi-model symbol-driven modeling: The emergence of large language models provides powerful tools for understanding human language. Following the principle of this paper, we can categorize the designer's lingual symbol vocabularies and integrate them with the workflow, to achieve the connection between language and interactive modeling. Moreover, emerging visualization technologies like augmented and virtual reality enable spatial symbol input and an immersive preview of the results. We believe its corresponding workflow will open opportunities for future urban planning and design with more creativity and possibilities.

Acknowledgments

The authors would like to thank the Digital Blue Foam team for their generous help and guidance during the process of development of the prototype.

References

Alexander, C. (1977). *A pattern language: towns, buildings, construction.* Oxford University Press.

Alonso, L., Zhang, Y. R., Grignard, A., Noyman, A., Sakai, Y., ElKatsha, M., ... Larson, K. (2018). Cityscope: a data-driven interactive simulation tool for urban design. use case volpe. In *Unifying themes in complex systems ix: Proceedings of the ninth international conference on complex systems 9* (pp. 253–261).

Batty, M. (1997). Cellular automata and urban form: A primer. *Journal of the American Planning Association, 63(2),* 266-274. doi: 10.1080/01944369708975918

Beirao, J., Mendes, G., Duarte, J., & Stouffs, R. (2010, 09). Implementing a generative urban design model: Grammar-based design patterns for urban design. doi: 10.52842/conf.ecaade.2010.265

Boim, A., Dortheimer, J., & Sprecher, A. (2022). *A Machine-Learning Approach to Urban Design Interventions in Non-Planned Settlements.* doi: 10.52842/conf.caadria.2022.1.223
Chen, G., Esch, G., Wonka, P., Mu¨ller, P., & Zhang, E. (2008, August). Interactive Procedural Street Modeling. *ACM Trans. Graph., 27.* doi: 10.1145/1399504.1360702
Digital Blue Foam. (2023, December 18). *DBF: AI Generative Design and Spatial Analytics.* Retrieved 2023, December 18, from https://www.digitalbluefoam.com/.
Sidewalk Labs. (2023, December 18). *Delve by Sidewalk Labs - Real Estate Generative Design.* Retrieved 2023, December 18, from https://www.sidewalklabs.com
Gkjohnson (2023, December) *Gkjohnson/three-gpu-pathtracer: Path tracing renderer and utilities for three.js built on top of threemesh-bvh.* Retrieved 2023, December 18, from https://github.com/gkjohnson/three-gpu-pathtracer.
Ishii, H., Underkoffler, J., Chak, D., Piper, B., Ben-Joseph, E., Yeung, L., & Kanji, Z. (2002). Augmented urban planning workbench: overlaying drawings, physical models and digital simulation. In *Proceedings. international symposium on mixed and augmented reality* (pp. 203–211).
Keir. (2023, December). *ProbableTrain/MapGenerator.* Retrieved 2023-12-03, from https://github.com/ProbableTrain/MapGenerator (original-date: 2020-04-01T17:01:11Z)
Kelly, T. (2021). CityEngine: An Introduction to Rule-Based Modeling. In W. Shi, M. F. Goodchild, M. Batty, M.-P. Kwan, & A. Zhang (Eds.), *Urban Informatics* (pp. 637–662). Singapore: Springer. doi: 10.1007/978-981-15-8983-6 35
Koenig, R., Miao, Y., Buˇs, P., Knecht, K., & Chang, M.-C. (2017). *Interactive Urban Synthesis Computational Methods for fast Prototyping of Urban Design Proposals.*
Lynch, K. (1964). *The image of the city*. MIT press.
Mei, Z., Pan, Y., Cheng, J., & Garcia Del Castillo Lopez, J. L. (2021). Cross-Scale and Density-Driven City Generator—Parametric assistance to designers in prototyping stage. 563–570. https://doi.org/10.52842/conf.ecaade.2021.1.563
Olascoaga, S., & Emilio, C. (2016). *Painting with data: from a computational history of urban models to an alternative urban computing* (Thesis, Massachusetts Institute of Technology). Retrieved 2023-12-03, from https://dspace.mit.edu/handle/1721.1/106680
Olascoaga, S., & Emilio, C. (2021). *Drawing Participation: Histories of Geospatial Computing, Professional Silos, and Computational Potentials for Collaboration in Planning and Design* (Thesis, Massachusetts Institute of Technology). Retrieved 2023-12-03, from https://dspace.mit.edu/ handle/1721.1/140089
KPF Urban Interface(2020)., *Scout.* (n.d.). Retrieved 2023-12-17, from https://scout.build/
Sevtsuk, A. (2021, October). Estimating Pedestrian Flows on Street Networks. *Journal of the American Planning Association*, 87(4), 512–526. doi: 10.1080/01944363.2020.1864758
mrdoob(2023)., *Three.js – JavaScript 3D Library.* (n.d.). https://threejs.org/.
Triantakonstantis, D., & Mountrakis, G. (2012, December). Urban Growth Prediction: A Review of Computational Models and Human Perceptions. *Journal of Geographic Information System*, 4(6), 555–587. doi: 10.4236/jgis.2012.46060
White, E. T. (1975). *Concept sourcebook: a vocabulary of architectural forms*.
Wilson, L., Danforth, J., Davila, C. C., & Harvey, D. (2019). How to generate a thousand masterplans: A framework for computational urban design. In *Proceedings of the symposium on simulation for architecture and urban design.* San Diego, CA, USA: Society for Computer Simulation International
Wilson, A. D., & Li, Y. (2007, October). Gestures without libraries, toolkits or training: a $1 recognizer for user interface prototypes. In *Proceedings of the 20th annual ACM symposium on User interface software and technology* (pp. 159–168). Newport Rhode Island USA: ACM. Retrieved 2023-11-26, doi: 10.1145/1294211.1294238
Zhu, W. (1993). *Space, symbol, city: A design theory of city.* China Architecture & Building Press, Beijing.

RE-IMAGINING THE URBAN DEVELOPMENT OF WESTERN SYDNEY: THE CASE STUDY OF ORAN PARK

MONIKA KRNCEVIC[1], NIOUSHA ZENDEHDEL ARJAGHI[2,] MOHAMMED MAKKI[3] and JORDAN MATHERS[4]
[1,2,3]University of Technology Sydney.
[1]monika.krncevic@student.uts.edu.au, 0009-0005-8665-9555
[2]niousha.zendehdelArjaghi@student.uts.edu.au,0009-0009-8925-224X
[3]mohammed.makki@uts.edu.au, 0000-0002-8338-6134
[4]jmathers@sjb.com.au

Abstract. This paper addresses the challenges of rapid urban expansion in Western Sydney, Australia, using the suburb of Oran Park as a case study. With the region's population projected to more than double by 2041, and an expected influx of an additional 400,000 people by 2030, there is a pressing need for sustainable and environmentally responsive urban development. Current approaches have prioritised space utilisation over environmental and social considerations, leading to homogeneity and poor urban quality. In response to these challenges, this study proposes a four-stage generative model for Oran Park, emphasising environmental restoration, agricultural integration, and housing diversification. This model aims to balance economic growth with environmental sustainability, contrasting with the density-focused development prevalent in the area. By implementing multi-objective optimisation, this research presents an algorithm-driven approach to urban planning, catering to the diverse needs of the expanding population.

Keywords. Western Sydney, Oran Park, Sequential Simulations, Evolutionary Algorithm, Computational Design, Urban Growth, Housing Development

1. Introduction

Sydney, Australia is currently experiencing rapid growth in urban settlement, resulting in significant demand for housing in the most populous and urbanised state within Australia, New South Wales (Farid Uddin & Piracha, 2023). Coupled with a significantly increasing housing market, and limited space within existing infrastructure in the city, Sydney's urban sprawl has expanded rapidly towards the western part of the state, known as the Greater Western Sydney region (Lawton & Morrison, 2022). This expansion presents both opportunities and challenges. On the one hand, it offers a chance for economic growth and the development of new communities; on the other hand, it underscores the pressing need for strategic urban planning to ensure that growth is sustainable, equitable, and integrates well

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 353-362. © 2024 and published by the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA), Hong Kong.

with the existing urban fabric through environmental, historic and future climate considerations.

The rapid urban expansion in Western Sydney, a region projected to see its population double from 2.4 million in 2016 to 4.1 million by 2041, presents critical challenges in urban planning and environmental sustainability (DPIE, 2019). This phenomenon is not confined to temperature changes. Western Sydney, Australia's third-largest economy and a rapidly growing centre for population and employment, faces challenges beyond just urban heat (Forster 2006). The Australian Government's Centre for Population predicts an influx of 400,000 people into Western Parkland City's eight local government areas (LGAs) by 2030 (DPIE, 2019). By 2031, over half of Sydney's population will reside in Western Sydney (Morrison et al. 2022). Historic precedents of rapid urban growth triggered by population demands showcase repetitive housing developments that prioritise the use of space over all else and do not enable the development of supportive infrastructure such as schools, open spaces and community facilities (Minns, 2023). This results in a complete disregard for the quality of the urban setting, its relationship to the environment surrounding it, and the community that it supports (Minns, 2023). Such practices underscore the urgency of sustainable urban planning despite the temporal stresses on urban growth.

2. Background and Context

2.1 ENVIRONMENTAL AND GEOGRAPHIC RESPONSES OF NEW URBAN DEVELOPMENTS

According to the Department of Environment and Climate Change (2008), the Sydney region, including Western Sydney, is experiencing significant climate change impacts. The Greater Sydney Region Plan, anticipates a population surge to 5.3 million by 2031, necessitating an additional 640,000 homes and 500,000 jobs (Department of Planning NSW, 2005). Western Sydney's exposure is heightened due to its lack of cooling sea breezes, unlike Sydney's eastern coastal suburbs. A University of Technology Sydney research analysis revealed that temperatures in Western Sydney are rising at over twice the rate of coastal suburbs or the global warming average (Lewins, 2023). This exacerbates the urban heat island effect, where urban development and dark-coloured surfaces increase local temperatures, impacting health and infrastructure (Lewins, 2023).

Addressing the challenges of cooling urban environments demands a comprehensive approach that extends beyond the introduction of tree canopy. It involves rethinking suburb planning, valuing open vegetated spaces, reducing car dependency, and reconsidering dwelling sizes and construction materials (Morrison et al. 2022). Water harvesting and reuse also play a crucial role in supporting plant growth for cooling. However, current master planning practices have often led to a lack of green spaces, contributing to urban overheating (Morrison et al. 2022).

2.2 HOUSING DEVELOPMENTS

The prevailing urban development strategy in Western Sydney has primarily revolved around the expansion of broadacre housing estates, often neglecting the essential provision of community facilities and services (Johnson, 2016). This emphasis on expansion raises concerns about the preservation of an adequate 'quality of place,' a concept frequently underscored in global city discourse (Mee, 2010). However, there is a increasing shift in the political landscape, with various authorities pushing for a redirection of focus towards higher-density developments, particularly in areas with high amenities already established.

The New South Wales Premier Chris Minns is pushing for the construction of 75,000 homes annually, doubling the Department of Planning's 2022 forecast, to address the state's housing crisis (Wikramanayake, 2023). The proposed plan includes the establishment of high-density towers near train stations, complemented by detached housing initiatives. Minns envisions the annual creation of 10,000 to 20,000 homes through six-story apartment complexes. However, this ambitious plan encounters challenges, such as the approval of only 6,000 homes in such complexes by September, resident resistance to medium-density housing, and the need for the construction industry to adapt (Wikramanayake, 2023). Additionally, the government plans to meet housing demands by expanding, demolishing, or redeveloping 50,000 existing apartment complexes built before 2000 (Wikramanayake, 2023).

2.3 CASE STUDY: ORAN PARK

As Western Sydney evolves, it becomes clear that the region is no longer a homogeneous landscape representing the ‘Australian Dream’. Instead, it is a microcosm of global processes, marked by social polarisation and tensions (Morgan 2005). These changes in the region's demography and landscape highlight the ongoing role of Western Sydney as a suburban home but also signal the need for a more nuanced understanding and approach to urban planning in this dynamically evolving area. The Western Sydney region, already witnessing significant growth as evidenced by the plethora of residential developments designed and built in the last decade, is grappling with the consequences of neglecting climate and land considerations in its urban design. Oran Park, aiming to accommodate up to 45,000 occupants, exemplifies this issue with its repetitive housing patterns and lack of basic urban infrastructure, creating a car-dependent suburb with little regard for pedestrian needs or climate adaptability (Morrison et al. 2022).

An alternative urban strategy for Oran Park, and similar estates in Western Sydney, is vital for the sustainable growth of urban development in the region. Through the utilisation of a multi-objective optimisation model, the research presents a proposal that reimagines the urban growth of Oran Park the considers the site's environmental and sociodemographic factors. The first stage of this model focuses on restoring the natural topography and leveraging existing water bodies, addressing the issue of Oran Park being built on flood-zoned farmland, which has caused unintended flooding and toxic runoff in adjacent areas. The second stage involves creating agricultural zones for local workers, promoting medium-scale farming and self-sufficiency. The third stage proposes an irrigation system linked with aquaponic systems for agriculture and floodwater management. Finally, the fourth stage suggests increasing population density with diverse building typologies, supporting new commercial infrastructures like hospitals and schools. These stages aim to harmonise Oran Park's present and future needs, offering a stark contrast to current urban practices in Western Sydney and showcasing the potential of computational models and algorithmic processes in urban planning.

3. Experiment Setup

The application of multi-objective evolutionary algorithms (MOEAs) in design, specifically urban design, has gained momentum over the last decade. The adoption of MOEAs to solve complex design problems has enabled designers to address one of the primary challenges for the design of urban spaces: the ability to simultaneously address multiple conflicting design objectives without necessitating trade-off decisions between the design goals. In this context, the presented research employs the NSGA-2 algorithm (Deb et al., 2000). situated within the Grasshopper plugin Wallacei (Makki et al., 2018).

3.1 DESIGN PROBLEM

The proposed model utilises Oran Park as a case study, proposing a hypothetical scenario in which the presented experiment is situated on Oran Park’s site before its current development. This flood-prone site in its pre-existing condition sees low-lying terrain and two water catchments connected to neighbouring agricultural properties. The first step of the process remodels this topographical condition and maps the water run-off paths on site through particle simulation (using Kangaroo 3D). The evolutionary algorithm selects one run-off line and extends it to the site boundaries, optimising to connect both water catchments at a maximum length (Figure 1A). This initial selection by the algorithm of the primary irrigation line forms the foundation of the proposed model, stipulating the axis on which the urban grid is situated.

Two sets of boundary lines ran parallel to the selected run-off path and were allocated as the shared agricultural corridor. The interior corridor was divided into lot sizes of either 2500m^2, 5000m^2 or 7500m^2. Exterior lots of 5000m^2 + were allocated as communal agricultural space and lots of 2500m^2 were for single-detached residential (SDR) properties (Figure 1B). Building footprints of 400m^2 were placed on SDR lots and extruded to either single or double storey heights based on their proximity to the communal agricultural spaces (Figure 1C). Exterior lots that fell into the water catchment boundary were elevated on structural supports above the water. Houses closer to these spaces were smaller in height, and those further were taller. This height management continued across the entire urban patch in order to ensure minimal solar obstruction on agricultural spaces and neighbouring buildings.

The following stage in the urban proposal introduced an irrigation system that follows the central run-off line, connecting every exterior lot to a system linking the water catchments at either end of the site (Figure 1D). A 40mx40m grid was generated on the same axis as the selected line and divided into four quadrants of varying size (Figure 1E). Cells with an area larger than 350m^2 were allocated as low-rise Residential Flat Buildings (RFB), offset 6m deep and split to create internal courtyards (Figure 1F). Every RFB building was extruded to a single-storey height, creating a mix of ground or level one communal gardens. A second set of extrusions were made above the RFB blocks with level counts informed by their proximity to the nearest communal agricultural space in the central corridor (Figure 1G). The remaining cells that were between 350m^2 and 300m^2 were allocated as high-rise RFB with a storey height ranging from 6-14 according to the same height management proximities as SDR and low-rise RFB. The remaining cells smaller than 300m^2 were offset by 2m to form a joined region allocated as communal agricultural grounds in the higher density zone of the urban plan (Figure 1H).

The remaining space was divided into a 20mx20m grid along the consistent axis. A cull pattern was used to remove 50% of the cells which were then allocated as multi-use buildings. Any connected buildings were joined into a single cell and extruded between a 2-5 storey height range according to height management proximities (Figure 1I). The 6m gap between buildings on the ground plane allowed space for a road network that would run through the high-density zone and along the boundaries of the external agricultural corridor (Figure 1K). To address the higher plain of the site, a set of pedestrian platforms and paths were combined and elevated at 1st storey height (Figure 1J). These paths all stem from a central path following the selected run-off line to form the tradeline that connects urban zones across the agriculture corridor and water bodies. Where the tradeline bridges the water bodies, a minimum of two aquaponic towers were placed, generating an enhanced irrigation loop that feeds the agricultural network across the site (Figure 1L).

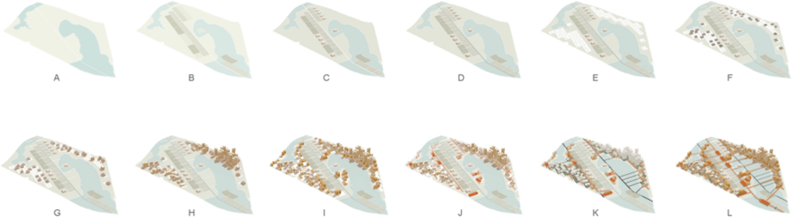


Fig. 1 (Construction of Phenotype – 1A-L)

3.2 FITNESS OBJECTIVES

The design model presented in the previous sections identifies the design variables that the algorithm modified to improve the fitness of each design solution. As such, the fitness functions (or design goals) that the algorithm aims to optimise for are the following five functions: The first is to maximise the length of the run off line to ensure it encompasses the entire site; the second sets a target ratio between single residential dwellings to agricultural land ; the third maximises the area of agricultural spaces; the fourth targets a ratio between low-rise and high-rise typologies; and the final function maximises the number of aquaponic tower to allow for maximum water catchment. The design matrix (Figure 2) outlines the relationship between the five fitness objectives, the design variables (explained in the previous section) and urban form. Visualising the design problem through this matrix allows for a more efficient 'debugging' of the design problem as it directly identifies the impact of each gene on each fitness objective and highlights imbalances that may be built into the developed model. This helps weigh phenotypic indicators in the selection process.

Fitness Criteria	Genes					WLC_ Plugin Point Step no.
	Size of irrigation network	Size of agricultural corridor	Level of solar exposure on ground	Increase Demographic	Connectivity	
Maximise length of run-off line	✓	✓	✓	x	✓	1
Target single-res property to public property (agricultural strip)	x	x	x	x	✓	3
Maximise average area of urban ground agricultural spaces	✓	x	✓	✓	✓	11
Target low-rise footprints to high-rise footprints	x	x	✓	✓	x	13
Maximise number of aquaponic towers	✓	x	x	x	✓	17

Fig. 2 (Design Matrix)

4. Experiment Results and Selection Process

4.1 SELECTION PROCESS

The algorithm evolved a total of 4000 design solutions with a simulation runtime of 39 hrs and 19 minutes. Identifying the best evolved solution in the algorithm requires a collateral analysis of the Pareto Front using additional urban metrics for the purpose of filtration (Showkatbakhsh & Makki, 2022). The algorithm generated 40 Pareto Front solutions, which were then clustered using a K-means clustering algorithm with a K-value of 16. Each cluster centre was selected to represent its respective cluster for further analysis.

A total of three filtration steps were implemented for selecting the top design solution. As part of the first and second filtration steps, the indicators used for analysis were prioritised through the allocation of weighting, thus varying the impact of each indicator on the analysis results; while the third step allocated equal weighting to the analysis metrics primarily due to their small number, which required less differentiation. The first filtration step prioritised the two highest-weighing fitness objectives alongside the highest mode values of urban metrics. These fell under two principal categories; Environmental (FO4 Length of run-off line) and Urban (FO1 Ratio of low-rise footprint to high-rise footprint) (Figure 3).

Phenotypic Indicators			Vector Direction	Weighting
Environmental	Aquaponic Tower distribution	**FO4 Length of run-off line**	+	1
		Number of irrigation branches connecting to water body	+	0.25
		FO5 Number of aquaponic towers	+	0.75
	Agricultural growth	**FO3 Ratio of single-res property to public property (agricultural strip)**	50/50	0.5
		Ratio of single-res agricultural space to urban agricultural space (combined)	20/80	0.75
		Ratio of urban area - ground agriculture to rooftop agriculture	70/30	0.25
		FO2 Average area of urban ground agriculture spaces	+	0.75
	Solar Gain	Sunlight hours on agricultural spaces (SR & urban)	+	1
Urban		Solar exposure on low-level buildings (SR, MR & Multi-use)	+	0.5
	Connectivity	Number of connecting paths between the Tradeline & urban zone	+	1
		Distance between single-res & urban	-	0.25
		Number of Tradeline Platforms	+	0.75
	Demographic Increase	**FO1 Ratio of low-rise footprint to high-rise footprint**	60/40	0.75
		Ratio of single-res footprint to urban footprint (combined)	30/70	0.5

Fig. 3 (Phenotypic Indicators)

The second filtration step took scores ranking in the top 25% of the ranking matrix (Figure 4) and assessed them against a series of manual calculations. The matrix used a colour scheme to score solutions based on calculations specified in the urban metrics. The four Pareto Fronts selected were further assessed on accessibility, connectivity, solar gain, and demographic distribution.

PHENOTYPIC INDICATORS	WEIGHTING	GEN.INDV.															
		23.5	27.2	51.17	60.19	67.2	78.9	82.35	83.31	91.17	97.22	98.1	99.39	20.25	97.28	24.15	66.36
Number of irrigation branches	0.25	0.078	0.148	0.058	0.080	0.033	0.073	0.055	0.063	0.063	0.050	0.073	0.120	0.078	0.055	0.073	0.068
Ratio of single-res agricultural space to urban agricultural space (combined)	0.75	0.603	0.533	0.197	0.500	0.176	0.290	0.440	0.097	0.217	0.116	0.286	0.561	0.689	0.605	0.218	0.653
Ratio of urban area - ground agriculture to rooftop agriculture	0.25	0.028	0.014	0.095	0.123	0.065	0.109	0.173	0.171	0.046	0.220	0.061	0.024	0.098	0.111	0.073	0.111
Sunlight hours on agricultural spaces (SR & urban)	1.00	0.401	0.456	0.482	0.501	0.425	0.454	0.495	0.405	0.495	0.485	0.456	0.469	0.499	0.486	0.463	0.456
Solar exposure on low-level buildings (SR, MR & Multi-use)	0.50	0.262	0.243	0.272	0.267	0.247	0.263	0.249	0.241	0.292	0.295	0.213	0.261	0.305	0.282	0.285	0.298
Number of connecting paths between the Tradeline & urban zone	1.00	0.120	0.100	0.600	0.102	0.140	0.120	0.150	0.900	0.500	0.120	0.110	0.700	0.900	0.500	0.110	0.810
Distance between single-res & urban	0.25	0.050	0.053	0.059	0.057	0.076	0.059	0.058	0.045	0.064	0.051	0.057	0.050	0.052	0.050	0.052	0.049
Number of tradeline platforms	0.75	0.300	0.150	0.225	0.300	0.150	0.225	0.375	0.150	0.150	0.150	0.225	0.150	0.300	0.150	0.300	0.150
Ratio of single-res footprint to urban footprint (combined)	0.5	0.056	0.409	0.334	0.149	0.115	0.279	0.078	0.149	0.126	0.084	0.275	0.102	0.245	0.082	0.074	0.279
F04 Length of run-off line	1.00	0.310	0.194	0.512	0.248	0.836	0.662	0.650	0.361	0.500	0.570	0.691	0.110	0.405	0.287	0.189	0.470
FO5. Number of aquaponic towers	0.75	0.233	0.000	0.233	0.491	0.492	0.244	0.491	0.233	0.491	0.233	0.491	0.077	0.079	0.233	0.077	0.077
F03. Ratio of single-res property to public property (agricultrual strip)	0.50	0.452	0.388	0.055	0.117	0.000	0.000	0.057	0.441	0.401	0.072	0.000	0.184	0.171	0.117	0.084	0.205
F02. Average area of urban ground agriculture spaces	0.75	0.197	0.586	0.158	0.535	0.563	0.563	0.491	0.443	0.017	0.220	0.743	0.709	0.266	0.299	0.126	0.578
F01. Ratio of low-rise footprint to high-rise footprint	0.75	0.317	0.315	0.518	0.389	0.367	0.367	0.037	0.020	0.252	0.147	0.044	0.074	0.409	0.410	0.630	0.239
Totals		3.405	3.588	3.795	3.857	3.683	3.706	3.798	3.715	3.612	2.812	3.724	3.589	4.495	3.666	2.752	4.442
Rank (1-4)					3			4						1			2

Fig. 4 (Scoring of Phenotypic Indicators – Ranking Matrix)

The goal for accessibility was to maximise access to the urban agriculture zones from the RFB blocks, calculating the number of paths generated between the two under a 20m radius (Figure 5A). The intention is to activate hubs within the superblock with community functions of engagement, gathering and collaboration (e.g. community gardens). Similarly, connectivity aimed to maximise site coverage while minimising the average length of a path before it turns, calculating average path lengths against the total pathway area (Figure 5B). The current pedestrian network in Oran Park is partly disconnected and unsafe as it is secondary to that of the vehicle. Optimising pathway connectivity prioritises the pedestrian, generating higher use (coverage) and comfortable access (length). With a heavy emphasis on supporting agricultural networks across site, it was important to maximise solar access on both the ground plane and rooftops. This was done by calculating the percentage of solar hits above 4 (out of 6), prioritising the agricultural corridor (Figure 5C). The demographic distribution analysis was assessed based on two target ratios; (a) 50/50 between Multi-use and Residential typologies and (b) Residential balance ratio of 10% SR, 20% MR (2 person) and 20% MR (4 person) (Figure 5D). Engaging a balanced variety of building typologies allows for a wider range in user demographics. This in turn supports the growth of socio-economic systems within Oran Park and its neighbours.

The third and final filtration step focussed on agricultural distribution and potential across the 5 land typologies; agricultural (private), agricultural (communal), urban ground (communal), urban rooftop (private), urban rooftop (communal) and waterbodies. A calculator was made using a list of fruits and vegetables typically grown in the Camden Council region, as well as poultry and aquaponics. These items were assessed on the estimated growth per acre or square metre of land, alongside rooftop options and sunlight needs (hours and shade tolerances). This calculator was then used to quantify the optimal target ratios for area per land typology; (a) 60/40 between Agricultural Zones and Urban Zones and (b) Zone balance ratio of Agricultural.P 15%, Agricultural.C 45%, Urban Ground 20%, Rooftop.P 5%, Rooftop.C 15% (Figure 5).

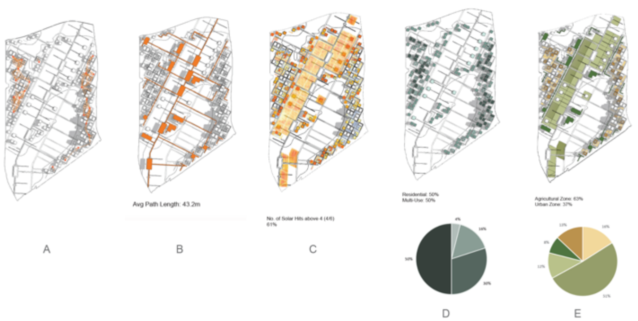


Fig. 5 (Analysis Diagrams – 5A-E)

5. The Selected Solution

Following the three-step filtration process, Generation 20, Individual 25 was selected as the most successful solution from this simulation. To further investigate and understand the relationships between the varying typologies in the selected solution, two focal points were selected and designed at a closer scale. These focal points highlight the two significant thresholds of the site; urban zone to agriculture corridor and tradeline to waterbody. The urban to agriculture threshold includes multiples of each of the four building typologies (Figure 6 (left)), highlighting their unique relationships with each other and the surrounding site. Implementation of communal platforms and garden spaces aims to consolidate these distinctive building typologies, allowing them to not only exist within close proximity but establish a high level of connectivity and user engagement. In the second focal point, the tradeline to waterbody threshold includes multiples of two building typologies connected to the trade market and aquaponic towers (Figure 6 (right)). This focal point captures the central line of the agriculture corridor and irrigation network, highlighting how the communal platforms and garden spaces land in this zone differently to the urban zone. Identifying the relationship between water-specific typologies and the agriculture corridor through manual design gave additional insight into how spatial interactions surrounding the tradeline would function.

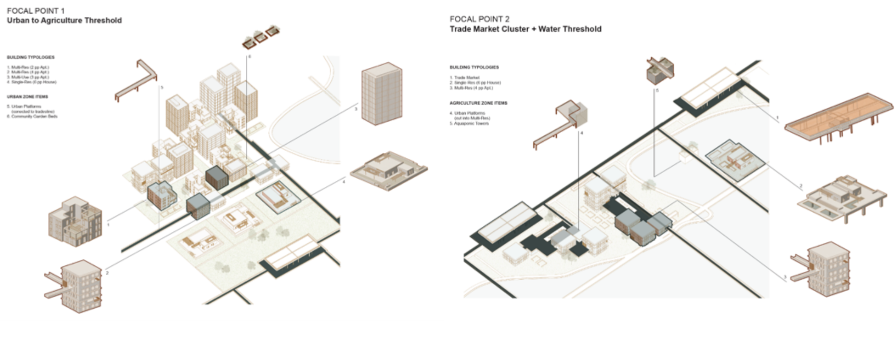


Fig.6: (left) Urban to Agriculture Threshold (right) Tradeline to Waterbody Threshold

6. Conclusion and Future Work

The case study of Oran Park examines housing development trends in a climate of rapid urban growth and demand. While the presented model targets the agricultural and flood-prone characters of its context, there is an opportunity to make site-specific adaptations for alternate locations. Through the use of evolutionary methods, a relationship between functional parameters and design goals was constructed, allowing the investigation to address social, environmental and urban challenges within and beyond the site. There is a consistent emphasis on the need for environmental connection to an agrarian, flood-prone site and the expansion of user demographic through variation in building typology and access. The integration of these two distinct systems evolved and strengthened at all stages; starting at the scale of broader site analysis, through the construction of the phenotype and into manual design. The analysis of the two focal thresholds in the selected solution exhibited a newly developed discourse between social and environmental contrasts through an assimilation of accessibility and irrigation networks (both visually and physically). Moving forward with the research, further work is needed to test the developed model on alternative urban sites in Western Sydney, exploring the adaptability of the model to solve similar problems experienced in various suburbs in the region. Moreover, future work will incorporate real-world government policies within the developed model, to transition the presented research towards real-world applications.

References

ABS. (2021). "2021 Oran Park, Census All persons QuickStats." *Australian Bureau of Statistics.*

Deb, K., Agrawal, S., Pratap, A., Meyarivan, T. (2000). "A Fast Elitist Non-Dominated Sorting Genetic Algorithm for Multiobjective Optimization: NSGA-II." In: International Conference on Parallel Problem Solving From Nature. Springer, France, pp. 849–858.

DPIE (2019). "Projections." Viewed 28 September 2021.

Farid Uddin, K., & Piracha, A. (2023). "Urban planning as a game of power: The case of New South Wales (NSW), Australia." *Habitat International, 133*, 102751.

Forster, C. (2006). "The challenge of change: Australian cities and urban planning in the new millennium." *Geographical Research, 44* (2), 173-182.

Gardiner, W. (1987). "How the West was lost: urban development in the western Sydney region." *Australian Quarterly, 59* (2), 234-244.

Johnson, L. (2016). "Western Sydney and the desire for home." *Australian Journal of Social Issues, 32* (2), 115-128.

Lawton, A., & Morrison, N. (2022). "The loss of peri-urban agricultural land and the state-local tensions in managing its demise: The case of Greater Western Sydney, Australia." *Land Use Policy, 120*, 106265.

Lewins, D. (2023). "Why Western Sydney is feeling the heat from climate change more than the rest of the city?" *The Conversation.*

Makki, M., Showkatbakhsh, M., & Song, Y. (2018). Wallacei: An evolutionary and Analytic Engine for Grasshopper 3D. Wallacei. https://www.wallacei.com/

Mee, K. (2002). "Prosperity and the Suburban Dream: quality of life and affordability in western Sydney." *Australian Geographer, 33* (3), 337-351.

Minns, C. (2023). "More planning resources to focus on the delivery of new houses and infrastructure." *Press release.*

Minns, C. (Premier & Minister for Education and Early Learning). (2023, September 11). "Minns Labor government invests $3.5 billion to tackle school infrastructure backlog for Western Sydney." *Press release.*

Morgan, G. (2005). "A city of two tales: Distinction, dispersal and dissociation in Western Sydney." In "*After Sprawl: Post-Suburban Sydney: E-Proceedings of the 'Post-Suburban Sydney: The City in Transformation*, pp. 22-23.

Morrison, N., Cmielewski, C., O'Mara, A., O'Neill, P., Pfautsch, S., & Power, E. R. (2022). "Wicked Urban Challenges in Western Sydney: Researchers Respond."

NSW. (2005). "City of cities: A plan for Sydney's future."

Showkatbakhsh, M., & Makki, M. (2022). "Multi-Objective Optimisation of Urban Form: A Framework for Selecting the Optimal Solution." *Buildings, 12* (9), 1473.

Trinder, J., & Liu, Q. (2020). "Assessing environmental impacts of urban growth using remote sensing." *Geo-Spatial Information Science, 23* (1), 20-39.

Wikramanayake, M. (2023, November 6). "Minns Reveals Medium-Density Overhaul for NSW." *The Urban Developer*.

RULE-BASED GENERATION OF INTERWOVEN ASSEMBLIES IN ARCHITECTURAL DESIGN: A COMPUTATIONAL APPROACH INTEGRATING ATTRIBUTE GRAMMAR

KEVIN HARSONO[1], SHEN-GUAN SHIH [2], YEYINT AUNG [3], FELICIA WAGIRI [4] and TSUNG-WEI CHENG [5]
[1,2,3,4,5]Department of Architecture, National Taiwan University of Science and Technology, Taipei, Taiwan.
[1]d11113803@mail.ntust.edu.tw, 0000-0003-1336-3380
[2]sgshih@mail.ntust.edu.tw, 0000-0001-7108-9960
[3]d11013803@mail.ntust.edu.tw, 0009-0008-3716-1891
[4]felicia_wagiri@hotmail.com, 0000-0001-8414-2599
[5]mike861104@gmail.com, 0009-0001-6620-2073

Abstract. This paper explores the potential of procedural modelling in generating design and assembly sequences for interwoven panel systems. The primary objective is to establish a computational framework that enables architects and designers to conceptualize and construct interwoven assemblies through rule-based generation, employing attribute grammar. The research delves into how procedural modelling, can effectively generate design and assembly sequences for interwoven panels in architectural design. The methodology unfolds in two parts: firstly, the framework for generative design for individual interwoven panel; and secondly, the definition of production rules for the assemblies, accompanied by practical applications of attribute grammar. The results indicate that using attribute grammar is an effective way to handle rule-based generated assemblies, incorporating attributes as semantics to manage transformations and block placement. This approach is expected to aid architects and designers by providing a robust computational toolkit for the creation of rule-based generation and assembly of interwoven panels. This research is expected to contribute to enhancing the development of complex geometric designs within architectural practice.

Keywords. Interwoven Structures, Procedural Modelling, Architecture Design, Computational-aided Design Tools, Attribute Grammar.

1. Introduction

Complex geometric structures, including interwoven designs, have undergone significant development over the years. However, their full potential in architectural design remains largely unexplored. In architectural design, interwoven assemblies involve the intricate weaving of various elements within buildings. Interwoven

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 363-372. © 2024 and published by the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA), Hong Kong.

assemblies are a type of interlocking system of blocks that enhance structural stability. These components also promote modular design, where rule-based generation can effectively represent their configuration and topological relationships. The ability to create panels from these interwoven assemblies allows designers to engage in procedural design. Despite dedicated research efforts on interwoven assemblies (Casado et al., 2021; Muslimin, 2010, 2011), architects have predominantly drawn inspiration from manual techniques and materials rooted in traditional basketry to craft elements such as furniture, facades, walls, and roofs. However, the widespread application of interwoven principles remains primarily within traditional practices. In the field of computer science, there has been extensive research on rule-based generation based on grammar techniques. However, there is a notable absence of systematic discussion in architectural research regarding the systematic and automated generation of designs and assembly sequences in real-world construction, particularly in the context of interwoven systems.

This research proposes a systematic framework for generating interwoven assemblies. The study is structured into two main parts. The first part focuses on the generation and design of the interwoven panel. The goal is to create a set of simplified blocks that can serve as components in more intricate assemblies within the interwoven assemblies. This phase involves the initial generation of foundational building blocks for the interwoven assemblies. In the second part, the research will extend to generating rules and grammar for the blocks created in the first phase. Additionally, it defines rules for hierarchical assemblies, providing a structured organization for assemblies based on different levels. This step aims to establish a systematic set of rules and grammar governing the construction of both individual blocks and their assembly into more complex structures.

In this paper, a computational framework designed for creating and assembling interwoven masonry systems is proposed. This study attempts to maximize the possibilities of interwoven masonry systems in architectural design, with a focus on their integration into the architectural building component. This work aims to revolutionize how architects and designers approach the concept of interwoven assemblies by merging computational design tools, procedural methodology, and utilization of attribute grammar and parsing techniques.

2. Interwoven in Architecture Realm

Previous research (Muslimin, 2010) has asserted that employing an interwoven configuration plays a crucial role in balancing stress factors like tension, bending, and shear on surfaces and cells. Moreover, the interweaving of neighbouring cells in both axes within the interwoven configuration serves to prevent yarn deformation caused by vertical loads. This underscores the significance of the interwoven concept not only as a cultural and artistic expression but also as a practical and functional solution in addressing structural stresses.

Transitioning into architecture, the term "interwoven" assumes a specific design connotation. In architectural contexts, it refers to a technique that intricately intertwines or locks different elements to shape a unified and visually compelling structure, exemplified by structures like the Imai Hospital Daycare Center in Japan designed by Shigeru Ban. Upon deeper examination, the significant characteristic of interwoven assemblies lies in their repetition, allowing for modularization into smaller, more manageable parts. This inherent quality facilitates procedural design and prompts consideration of the rules and procedures involved in generating and defining these complex structures.

3. Interwoven Block Design

To prepare the design of the interwoven building block, a comprehensive understanding of its components is imperative. As clarified in Section 2, the interwoven assemblies can be derived from smaller modular panels, allowing for the dissection of each part based on its function. Through the literature review and analysis of case studies, it becomes obvious that developing a modularized interwoven structure for architectural building components involves the integration of two fundamental elements. The first element, the horizontal component, is embodied by sturdy columns that furnish structural support and stability. On the other hand, the vertical component comprises interwoven panels strategically integrated with the columns. This not only enhances aesthetic appeal but also provides functional advantages such as environmental control or privacy. The right side of Figure 1 illustrates the initial prototyping for interwoven assemblies.

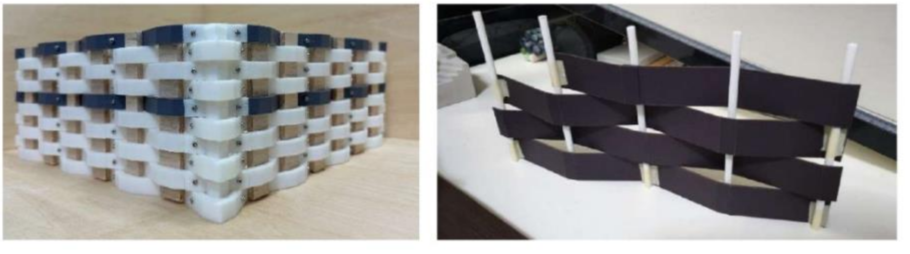


Figure 1. Left: The 3D printed model of the interwoven panel. Right: Prototype of interwoven block.

From this experiment, we establish a design framework for producing interwoven panels. The parameters involved include the profile of the weaving section and the distance between columns. The weaving angle is dependent upon the distance between the columns, while the profile dictates the height of the column panels. In this case, the formation of modular interwoven panels produces five distinct types, each assigned a specific function. Broadly categorized, these panels can be classified into two main components, shown in Figure 2, which are weaving and column panels. The weaving part further diverges into two categories. First, there is the Straight Weave (SW), shown in red colour, regarding panels characterized by a straightforward weaving pattern. Second, the corner weave category introduces two types: Inner Corner (IC), shown in green colour, indicating panels weaving through the inner part of the column, and Outer Corner (OC), shown in blue colour, signifying panels weaving through the outer part of the column. Parallelly, the column part presents two primary types: Corner Column (CC), shown in white colour, representing columns with a turning configuration, and Straight Column (SC), shown in yellow colour, where panels exhibit a straightforward column configuration.

The parameters for designing the interwoven assemblies include the distance between the vertical part and the profile of the weaving part. The weaving part's profile is rectangular, measuring 8x12 cm, with a column distance of 48 cm. The column height is maintained at twice the profile height for modularity, specifically 24 cm. The interlock between panels is meticulously designed to preserve the connection position, ensuring the maintenance of XY displacement. To establish connections between panels, a bolt join method is employed.

The modular approach guarantees the generation of a varied collection of panels, each customized for specific functionalities. This versatility is essential in constructing interwoven structures, facilitating nuanced and purposeful designs. To validate the functionality of the assemblies, a scaled model is presented using 3D printing, as illustrated on the left side of Figure 1.

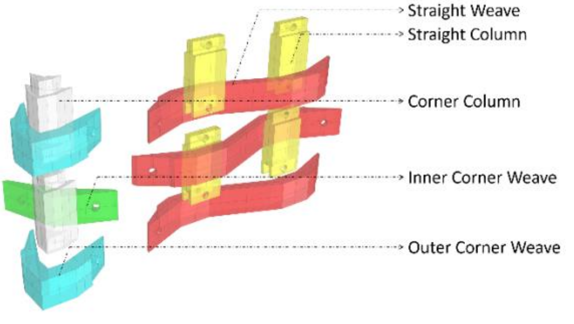


Figure 2. Block design and assemblies.

4. Grammar and Rule-Based Generation

The concept of incorporating grammar into design draws parallels between visual and natural languages. In natural language, grammar plays a vital role in facilitating communication by providing a structured framework for organizing and conveying meaning. Designers can represent complex visual information with the assistance of formal grammar defined by specific rules (Bruton, 1997). Rule-based generation involves computational tools and a predefined set of rules to generate architectural designs. These rules, affiliated with grammar in computer programming languages, provide a structured framework for design creation. By defining rules, designers can systematically generate designs that adhere to predetermined standards. Grammar serves as a systematic set of principles, functioning as an architect's toolkit to articulate the syntax and structure of the design language.

In this context, Context-Free Grammar (CFG) plays a prevalent role in rule-based generation research due to its simplicity and efficiency in interpretation for parsing and language generation. CFG necessitates being context-free and composed of a single non-terminal (Bikaun et al., 2022). However, CFG can only deal with strings of grammar rules, but not necessarily executable. Addressing the challenges posed by CFG, Attribute Grammar (AG) serves as an extension to CFG. By assigning attributes to symbols and defining how to compute these attributes, attribute grammar incorporates additional information (Prasad, 2009) into more meaningful and executable grammar.

Ultimately, to translate grammar into semantics, parsing is essential. This implies that to convert syntactic rules into meaningful and executable assembly sequences, the input must undergo analysis and interpretation by a parser. In architectural design, parsing involves dissecting input into components such as nouns (representing objects), verbs (depicting methods), and associated attributes or options. Guided by established grammar, this process takes a token string as input and transforms it into the corresponding representation (Teboul, 2021). In this study, the SLR parser is utilized. The parser ensures adherence to language grammar rules, thereby providing a logical and organized framework for the design investigation.

5. Attribute Grammar for interwoven Assemblies

In this study, the attribute grammar (AG) for a syntax-directed definition (SDD) is denoted as AG with a triple (G, A, AR), where G represents the context-free grammar for the language, A is associated with each grammar having attributes, and AR is semantic linked with production rules. Specifically, for interwoven assemblies, Token$\langle s, TS(s)\rangle$ and Token$\langle a, ID(a)\rangle$ are addressed to signify the terminal symbols with transformation action and the semantic of block placing action, respectively.

Take the production rule $X \rightarrow s\ a\ s\ a\ s\ a$ as an example; it implies moving straight and placing the block three times. Here, s represents TS or transformation of the block, and a signifies the $ID(a)$, the semantic aspect, indicating the placement. Since CFG alone cannot handle the semantics, the property $ID()$ is treated as an attribute inside the production rule in attribute grammar. Thus, the appropriate representation of the grammar with syntax and semantics is:

$$X ::= s\ s\ s\ \{X.val = TS(s)\ ID(a)\ TS(s)\ ID(a)\ TS(s)\ ID(a)\} \tag{1}$$

In this context, there are two primary rules governing interwoven assemblies: S, utilized for straight assemblies, and T, employed for turning configurations. Bellow CFG and AG grammar rule set S' show a language pattern with translation $TS(t)$, rotation $TS(u), TS(v)$ and placing block $ID(a)$:

$$\begin{array}{ll} S' ::= S & \{S'.val = S.val\ \} \\ S ::= E\ S \mid \epsilon & \{S.Val = E.Val\ S.Val \mid ID(a)\} \\ E ::= S \mid T & \{E.Val = S.Val \mid T.Val\} \\ S ::= t & \{S.val = ID(a)\ TS(t)\ \} \\ T ::= U \mid V & \{T.Val = U.Val \mid V.Val\} \\ U ::= u & \{U.val = TS(u)\ ID(a)\ TS(t)\} \\ V ::= v & \{V.val = TS(v)\ ID(a)\ TS(t)\} \end{array} \tag{2}$$

To establish the production rule of the interwoven assembly system, it is crucial to define the transformation between blocks. The terminal $\langle t, TS(t)\rangle$is made to define the transformation of the blocks. This step is essential because the grammar can only be formulated when the assemblies are explained in a geometric manner. The configurations of these interwoven structures find geometric explanations, particularly concerning rotational and translational transformations.

The designation of the terminal $\langle a, ID(a)\rangle$ acts as a semantic descriptor for the interwoven panel block placement. This flexibility allows for dynamic adjustments to the block based on predefined geometries or groups of geometries. The initial configuration of interwoven block IDs introduces two distinct types: $\langle corner\rangle$ and $\langle straight\rangle$, as shown in Figure 3. The $\langle straight\rangle$ type functions as a straight arrangement, incorporating 4 types of components: 2 panels of Straight Weave (SW) and 2 panels of Straight Column (SC). On the other hand, the $\langle corner\rangle$ type serves as a corner or turn function, comprising 3 types of components: Inner Turn Weave (IC), Outer Turn Weave (OC), and Corner Column (CC). Here, the height of the assemblies is determined by the H, as a branch of the sentence.

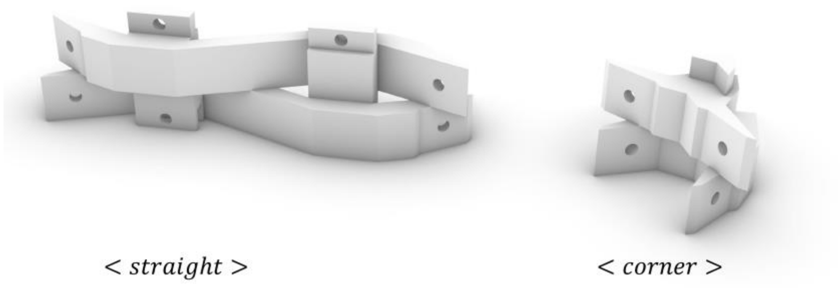


Figure 3. Block arrangement.

To address branching, L-systems are employed to manage the hierarchy of the sentence structure. In L-systems, branching is achieved by using the push ('[') operation and the pop (']') operation following the rule "first in last out", storing an action onto the top of a stack and retrieving it out of the stack, respectively. The concept was derived from turtle graphics and push-down automata (Prusinkiewiczy et al., 1996; Sipser, 2012). For example, if and only if n is a natural number, it could represent the height is H' token concatenating for n times and $\mathrm{ID}(i)$ places a block, the appropriate way to express this in production rules is:

$$\begin{aligned} &A \ ::= [\, n\, H' \,] \quad \{A.Val = \ H'.Val^{N(n)}\} \\ &H' \ ::= h \qquad \{H' = TS(h)\ ID(i)\} \end{aligned} \tag{3}$$

The transformation of this interwoven configuration will have three distinct forms of a natural number n with the non-terminal token in the right of rules: h, s, and t. Each attachment type serves a specific purpose, h is employed for the height, which means it will determine the number of interwoven blocks placed vertically. This s is employed for straight assemblies, and t represents turn actions. The transformation associated with these terminals is illustrated in Figure 5. Each type corresponds to a specific block type, along with a corresponding transformation, which is rotation $R_{(x,y,z)}$ and translation $T_{(x,y,z)}$ as proposed by previous research (Shih, 2018).

The rule regulating the interwoven panels demands that the turn is executed by flipping it. However, prescribing a rule to be flipped can introduce ambiguity and inefficiency. To address this challenge, the rule set for this interwoven panel is streamlined to only two options: go straight $\langle S\rangle$ and turn $\langle T\rangle$. To ascertain the direction, the parser matches terminal patterns with the transformation matrix derived from the segments of the input polyline, determining whether it aligns with the given token. To obtain the semantic patterns for matching the height configuration, the attributes are modified with $[\langle H\rangle.Val^{N(n)}]$. Figure 4 illustrates the transformation matrix for each $TS()$. Additionally, the $\langle T\rangle$ offers two choices depending on the direction, whether it involves a left turn or a right turn, this means the $\langle T\rangle$ can be either $\langle U\rangle$ or $\langle V\rangle$, corresponding to the matching direction transforming actions. This approach enhances clarity and efficiency in the rule-based generation of interwoven structures.

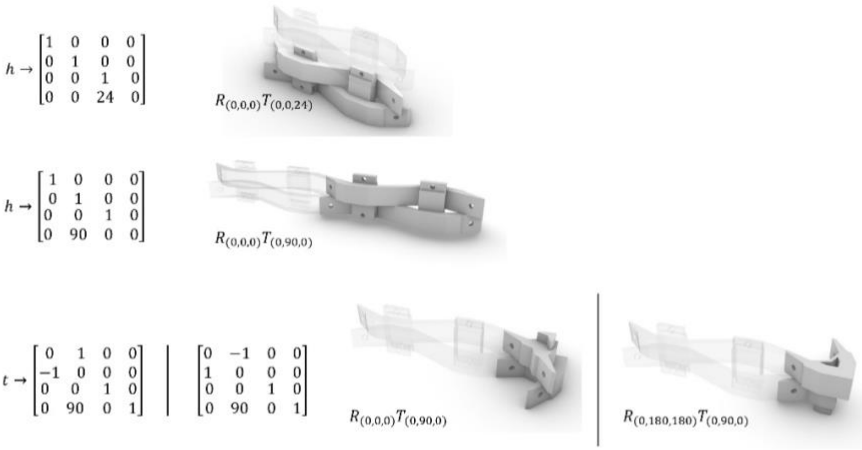


Figure 4. The transformation matrix for the non-terminal of the interwoven panel.

Therefore, after defining the terminal tokens for transformation $TS()$ and the $ID()$ of each block, the production rules can be specified. The syntax and semantic structure of the production rules are represented as follows:

$$\begin{array}{ll}
\langle Main\rangle ::= \langle Stmt\rangle & \{\langle Main\rangle.Val = \langle Stmt\rangle.Val\} \\
\langle Stmt\rangle ::= \langle E\rangle\langle Stmt\rangle \mid \epsilon & \{\langle Stmt\rangle.Val = \langle E\rangle.Val\ \langle Stmt\rangle.Val \mid \epsilon\} \\
\langle E\rangle ::= \langle corner\rangle \mid \langle straigth\rangle & \{\langle E\rangle.Val = \langle corner\rangle.Val \mid \langle straigth\rangle.Val\} \\
\langle straight\rangle ::= \langle S\rangle & \{\langle straight\rangle.Val = \left[\langle H\rangle.Val^{N(n)}\right]\langle S\rangle.Val\} \\
\langle corner\rangle ::= \langle T\rangle & \{\langle straight\rangle.Val = \left[\langle H\rangle.Val^{N(n)}\right]\langle S\rangle.Val\} \\
\langle H\rangle ::= h & \{\langle H\rangle.Val = TS(h)ID(i)\} \\
\langle S\rangle ::= s \mid \epsilon & \{\langle S\rangle.Val = ID(i)TS(s) \mid ID(i)\} \\
\langle T\rangle ::= \langle U\rangle \mid \langle V\rangle & \{\langle T\rangle.Val = \langle U\rangle.Val \mid \langle U\rangle.Val\} \\
\langle U\rangle ::= u & \{\langle U\rangle.Val = TS(u)ID(i)TS(s)\} \\
\langle V\rangle ::= v & \{\langle V\rangle.Val = TS(v)ID(i)TS(s)\}
\end{array} \quad (4)$$

To employ attribute grammar in a parser, it's necessary to build a parse tree. This tree is subsequently navigated in a specific sequence through Syntax-Directed

Translation (SDT). The creation of the parse tree relies on the token string and the production rules outlined in the preceding section, illustrated in Figure 5. At every node of the parse tree, the SDT rules is implemented, utilizing attributes linked to the grammar symbols to calculate semantic information (Aho et al., 2007).

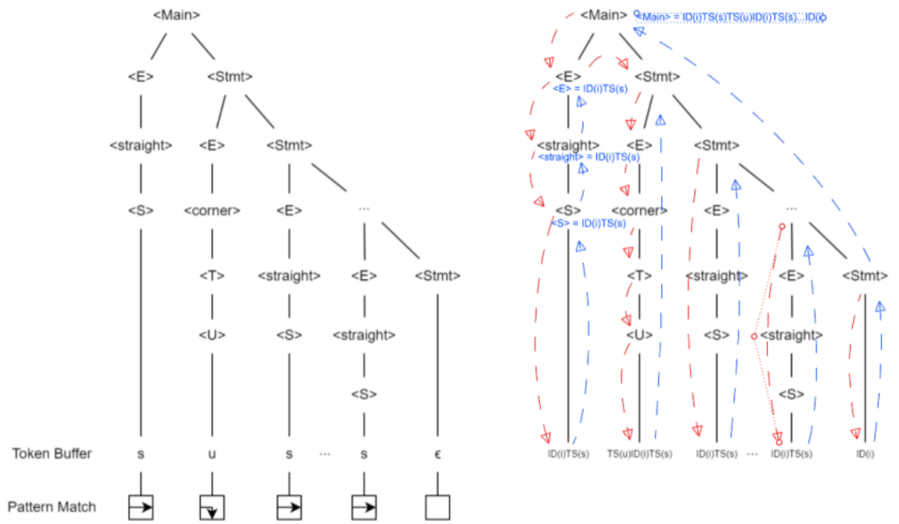


Figure 5. Left: Parse tree. Right: SDT.

6. Application of Grammar for Interwoven Assemblies

In this section, the practical application of grammar and parsing is demonstrated. The initial step involves inputting data, represented by curves in this case, into the parser as a token string. This input curve can represent a building plan, which architects can utilize for designing an interwoven panel wall.

Expanding on the rules delineated in the prior section for interwoven assemblies, the subsequent phase entails parsing the curve into interwoven assemblies. In this scenario, the height is configured to be H concatenating 8 times, approximately 2 meters. By defining H, we address the branching issue in the production rule. The applied grammar in this context can be observed as follows:

$$
\begin{array}{ll}
\langle straight\rangle ::= \langle S\rangle & \{\langle straight\rangle.val = [\langle H'\rangle.Val^{8}]\ \langle S'\rangle.Val\ \} \\
\langle corner\rangle ::= \langle T\rangle & \{\langle corner\rangle.val = [\langle H'\rangle.Val^{8}]\ \langle T'\rangle.Val\ \} \\
\langle H'\rangle ::= h & \{\langle H'\rangle.val = TS(h)\ ID(i)\} \\
\langle S\rangle ::= s \mid \epsilon & \{\langle S\rangle.val\ =\ ID(i)\ TS(s) \mid ID(i)\} \\
\langle T\rangle ::= \langle U\rangle \mid \langle V\rangle & \{\langle T\rangle.Val = \langle U\rangle.Val \mid \langle U\rangle.Val\} \\
\langle U\rangle ::= u & \{\langle U\rangle.Val = TS(u)ID(i)\ TS(s)\} \\
\langle V\rangle ::= v & \{\langle V\rangle.Val = TS(v)ID(i)\ TS(s)\}
\end{array}
\tag{5}
$$

Figure 6 shows the application for the input curve. The curves serve as tokens for parsing the grammar. Consequently, adjustments are made to the curve to ensure the proper spacing between blocks. The original curve, highlighted in black, is modified to align with the block's dimensions. The resulting adjusted curve, shown in green, emerges as a product of this process.

From the token buffer, the initial string is derived as

$$Token\ buffer\ \rightarrow\ \#s^4\ u\ s^{23}\ u\ s^4 u\ s^{12} v s^{12} \epsilon \#$$

Applying the production rule to the string results in

$$\langle Main \rangle ::= 4\ \langle straight \rangle\ \langle corner \rangle\ 23\ \langle straight \rangle\ \langle corner \rangle\ 4 \langle straight \rangle$$
$$\langle corner \rangle 12\ \langle stright \rangle \langle corner \rangle 12\ \langle straight \rangle\ \epsilon$$

Therefore, after traversing the syntax tree, we can gain the semantic:

$$\{\langle Main \rangle . Val = ([\ (TS(h)\ \ ID(i))^8]\ \ ID(i)\ \ TS(s))^4 \ldots ID(i)\} \tag{6}$$

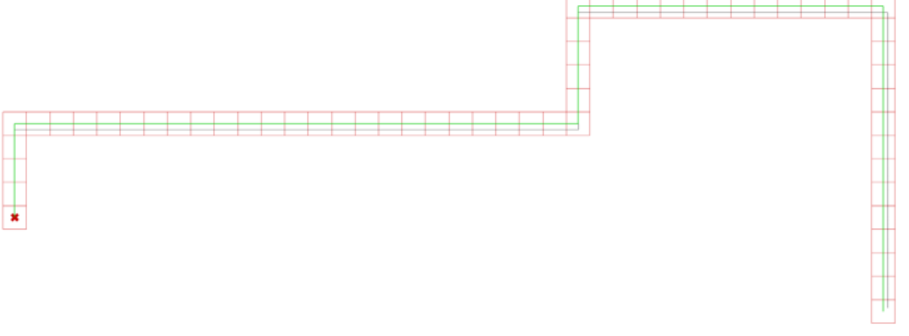


Figure 6. Curve adjustment.

This production rule is then incorporated into the parse tree, and the result from SDT is depicted in Figure 7, showcasing the interwoven panel wall.

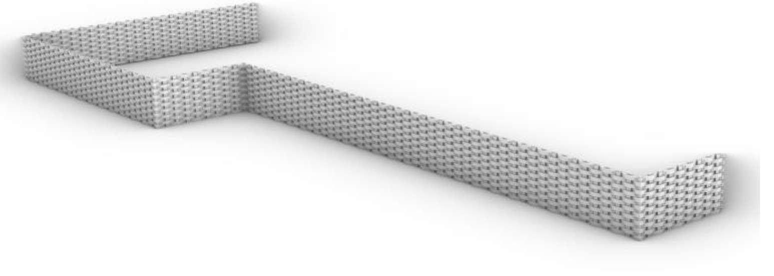


Figure 7. Result from SDT

7. Discussion

The utilization of this rule-based generation applies to architectural components, such as landscape furniture, walls and room dividers. By employing grammar, the designer becomes capable of managing the complexity associated with architectural elements. Furthermore, comprehending the production rule and its grammar provides designers with increased flexibility in designing the interwoven block and the aggregation of it.

This study emphasizes the potential of rule-based generation in creating interwoven assembly systems. By taking curves as token input, the system can generate both the interwoven assemblies and the corresponding syntax representing these assemblies. Furthermore, the integration of attribute grammar enhances the parsing process, particularly concerning geometric information. Attribute grammar facilitates the formal inclusion of semantic information, utilizing attributes, with a specific focus on symbols representing geometric objects in this study.

However, this study has certain limitations. Firstly, it does not currently support branched token input. Another limitation is its restriction to working only with curves at perpendicular angles. This implies that the software lacks the capability to generate interwoven assemblies at angled intersections. Lastly, the scope of this research is limited to flat surfaces. In future research, the integration of L-systems will be explored to overcome the branch limitation. Additionally, the application of interwoven assemblies on curved surfaces will be a focus of further investigation.

Reference

Aho, A. V., Lam, M. S., Sethi, R., & Ullman, J. D. (2007). Compilers: Principles, techniques, & tools (2. ed., Pearson internat. ed). Pearson Addison-Wesley.

Bikaun, T., Stewart, M., & Hodkiewicz, M. (2022). Using Context-Free Grammar to Generate Synthetic Technical Short Texts. In H. Aziz, D. Corrêa, & T. French, in *AI 2022: Advances in Artificial Intelligence* (Vol. 13728, pp. 325–338). Springer International Publishing.

Bruton, D. (1997). Grammars and Art. In R. Junge, *CAAD futures 1997* (pp. 71–82). Springer Netherlands.

Casado, A., Sánchez, A., Marieta, C., & Leon, I. (2021). Use of Flat Interwoven Wooden Strips in Architecture and Construction. Simulation and Optimization Using 3D Digital Models. *Sustainability*, 13(11), 6383. https://doi.org/10.3390/su13116383

Muslimin, R. (2010). Interweaving Grammar: Reconfiguring Vernacular Structure through Parametric Shape Grammar. *International Journal of Architectural Computing*, 8(2), 93–110. https://doi.org/10.1260/1478-0771.8.2.93

Muslimin, R. (2011). One-piece weaving: Reconfiguring folding and knotting algorithm in computational design. In *International Conference on Computer-Aided Architectural Design Research in Asia, CAADRIA 2011* (pp 9–18).

Prasad, T. K. (2009). *Attribute Grammars*. Encyclopedia of Information Science and Technology, Second Edition. http://dx.doi.org/10.4018/978-1-60566-026-4.ch046

Prusinkiewiczy, P., Hammely, M., & Hananz, J. (1996). L-system: from the theory to visual model of plants. In *2nd CSIRO Symposium on Computational Challenges in Life Sciences (pp. 1-32)*.

Shih, S.-G. (2018). The Art and Mathematics of Self-Interlocking SL Blocks. In *Bridges 2018 Conference* (pp. 107–114).

Sipser, M. (2012). *Introduction to the theory of computation (3rd Ed)*. Course Technology Cengage Learning.

Teboul, O. (2021). *Shape Grammar Parsing: Application to Image-based Modeling*. Ecole Centrale de Paris.

SIMFORMS

A Web-Based Generative Application Fusing Forms, Metrics, and Visuals for Early-Stage Design

BAIZHOU ZHANG[1], YICHEN MO[2], BIAO LI[3], YANYU WANG[4], CHAO ZHANG[5] and JI SHI[6]
[1,2,3,4,5,6]*School of Architecture, Southeast University, Nanjing, China*
[1]*zhang_baizhou@seu.edu.cn, 0000-0003-3153-2264*
[2]*moyichen@seu.edu.cn, 0000-0002-2844-7678*
[3]*jz_studio@seu.edu.cn, 0000-0002-8713-4954*
[4]*wangyanyu_jz@seu.edu.cn, 0009-0000-4123-1049*
[5]*zhangchao_hust@seu.edu.cn, 0009-0009-3172-6203*
[6]*shiji74454@gmail.com, 0009-0009-8322-9938*

Abstract. At the current moment when digital technology is gradually spreading, various generative design methods and the rapid iteration of AI models have exceeded public expectations. Nevertheless, in the early stages of architectural design, there are still numerous challenges in connecting forms, metrics, and visuals. When facing conceptual design, designers often need to strike a balance between the visual effects of form and quantitative metrics, lacking an efficient tool to quickly control different design choices and obtain comprehensive feedback. To address this issue, this paper introduces SIMForms, a web-based application aimed at fusing architectural forms, quantitative metrics, and visualization. SIMForms integrates rule-based parametric modelling, metric computation and feedback, along with AI-assisted conceptual image generation. Through SIMForms, designers can generate diverse architectural forms with simple operations in the early design stages, obtaining crucial quantitative metrics and conceptual image as feedback. This multi-module integrated application not only provides a more intuitive and efficient tool for the design process but also offers concept innovation and guidance for designers, driving further development in digital design tools.

Keywords. Early-stage Architectural Design, Web-based Application, Metrics Feedback, Rule-based Modelling, Conceptual Image Generation.

1. Introduction

In the era of rapid advancement in generative design and generative models, the early-stage architectural design process is gradually shifting towards interactive feedback and

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 373-382. © 2024 and published by the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA), Hong Kong.

real-time decision-making (Singh and Gu, 2012). For architects and urban designers, diverse inspiration guidance and sufficient data feedback are crucial criteria for evaluating the usability of digital design tools. From another perspective, facing the iterative process of "design test - metric verification", it is crucial to enable architects to quickly obtain architectural form suggestions through basic metric control requirements. This approach serves as a key means to simplify the difficulty of conceptualizing and refining architectural ideas. Therefore, establishing a digital bridge between "architectural form - quantitative metrics - creative visualization" is a practical challenge for the design process.

Rule-based parametric modelling is the most direct and fundamental way to connect architectural style prototypes with quantitative metrics. It has been extensively explored in interdisciplinary fields such as urban procedural modelling (Parish and Müller, 2001). By analysing architectural style prototypes and design logics, multiple architectural forms can be rapidly generated through the definition of algorithms and parameters, providing creative inspiration references (Lee et al., 2014). In recent times, with the widespread enthusiasm for AI large models, discussions on architecture generation and conceptual design based on image generation have gained significant attention (Ploennigs and Berger, 2023). Generating conceptual images that match architectural representation effects based on text prompts or images has become a potential development direction (Albaghajati et al., 2023; Cheng et al., 2023). An early-stage design platform incorporating different modules is a crucial solution for advancing the integration of digital design resources and addressing practical issues.

For integrated design-assist applications described above, user-friendliness has been a significant concern in recent years. This allows designers to obtain generated model or analysis results without the need to master complex algorithmic principles. Some application-oriented research has opted for local software or plugins (He and Yang, 2019; Fink and Koenig, 2019; Song et al., 2023). However, there is also a growing number of lightweight, cross-platform, web-based design applications showing potential in urban design and architectural form design problems (Zhang et al., 2022; Ortner et al., 2023; Xu et al., 2023). As an approach of visualization and interaction, web platforms serve as ideal tools for design promotion and collaborative communication. Through the separation of core computational modules and user interfaces, web platforms can effectively enhance the ease of use for practitioners.

Focusing on the above-mentioned issues, this paper aims to explore the potential for mutual feedback among forms, metrics, and visuals in the early-stage architectural design. The effort is made to integrate rule-based modelling, conceptual image generation, and web-based platforms, unifying "form-metrics-visualization" into an interactive application. To achieve this goal, "SIMForms" is introduced as a web-based conceptual design assistant application that integrates parametric modelling with AI-based conceptual image synthesis (Figure 1). Through this web application, designers and stakeholders can, during the conceptual design phase, obtain diverse architectural style models based on simplified control parameters. Furthermore, the application allows the synthesis of rendering images through AI models, enabling the examination of generated results and providing design references.

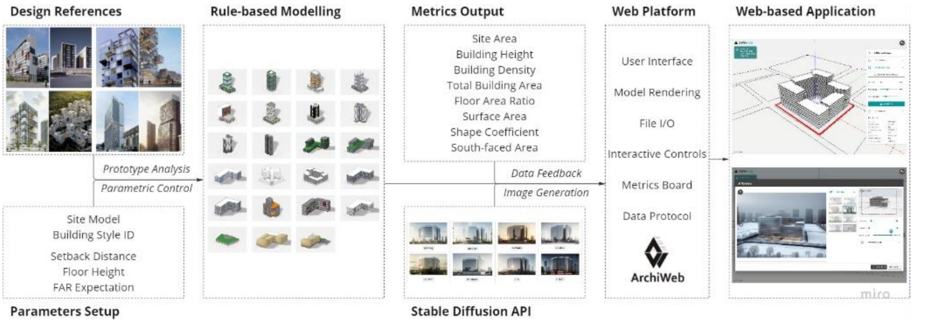


Figure 1. Workflow of SIMForms

2. Methods

2.1. PROTOTYPE ANALYSIS & PARAMETRIC MODELING

The parametric styles in SIMForms are derived from the analysis of real-world architectural design cases. The research team chose several examples of multi-story and high-rise public buildings from publicly available sources such as the internet and datasets, deconstructing the rules and geometric principles behind their forms. The current selection of cases aims to maximize diversity in form and revolves around different configurations of architectural elements. In terms of overall layouts, they include central enclosure, linear arrangement, and unit dispersion. In terms of operational techniques, they encompass strategies like volumetric displacement, chamfering, and stacking.

Inspired by prototype analysis, the basic generation rules for each architectural style were encoded through computer programming. All architectural styles treat storeys as basic volumetric units, considering them as stacked blocks regardless of their specific form. Facade or detailing rules are then refined based on the stacked blocks. Three fundamental parameters are employed in the generation rules to control the resulting forms while preserving their stylistic essence: storey height, setback distance, and FAR expectation.

- Storey height controls the height of each storey in the generated model.
- Setback distance simulates the architectural setback from the street and city boundaries. As the setback distance increases, the building footprint decreases, resulting in an increase in the total number of storeys, given the FAR expectation.
- FAR expectation is a parameter controlling the scale of the building. Using this value and the site area, the total building area can be calculated, and, based on the form rules of each building, a rough estimate of the total number of storeys can be deduced.

The selection of these three fundamental parameters reflects the objective of enabling designers to obtain different architectural models by adjusting values in a simple and intuitive manner. Behind these parameters, adaptive modelling rules have been defined for each style's generation process, including:

- In some styles, hidden parameters with random values are introduced, enabling control over aspects such as random wall insertions or random openings in facade grids. This feature results in variations in each generation under the same set of parameters.
- In certain high-rise building styles, the tower's bounding box is derived from the maximum inscribed rectangle of the site outline, utilizing optimization algorithms for iterative searches. This ensures the generation of effective and rational tower forms in various site shapes.

In conclusion, the goal from analysing design prototypes to parametric modelling is to incorporate the generation rules of different architectural styles into a straightforward control parameter framework. This ensures the preservation of style characteristics, diversity, and adaptability. The same architectural style can guide adaptive generation results under different parameters and site shapes (Figure 2).

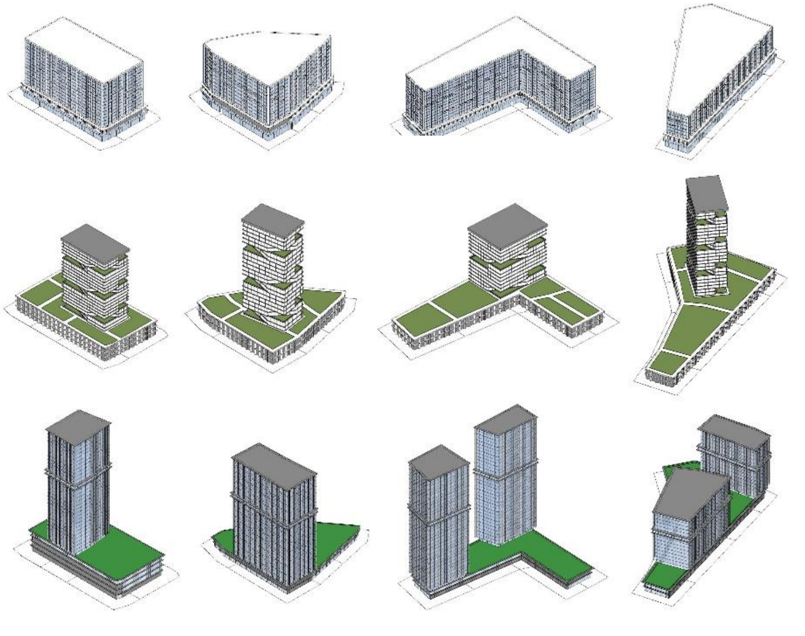


Figure 2. The results of the same style generated by different sites and parameters

2.2. METRICS CALCULATION

During the early-stage design, a crucial step is the rapid verification of basic technical indicators. As the creation rules of 3D models are encoded through computer programming, fundamental metrics can be obtained through simple geometric operations. In the parametric modelling phase, all created 3D elements include "elementType" attribute, allowing for the straightforward differentiation of the element's type. This attribute serves to filter whether an element is marked as a volumetric block or as a non-spatial component, such as windows, floor slabs, facade grids, or decorative elements. Given that the logic of parametric modelling is based on storeys, the volumetric block encompasses both volume and floor area information. These elements marked as volumetric block are then extracted separately for metrics calculation, fulfilling the approximate control requirements during the early design phase (Figure 3). SIMForms currently employs a set of 8 metrics for calculation and display:

- Site Area: The area of the site, representing the polygonal area of the user-inputted site.
- Building Height: The total building height, calculated based on the expected storey count derived from the FAR expectation and the user-inputted standard storey height.
- Building Density: The union area of the base polygons of all volumetric blocks projected onto the plane.
- Total Floor Area: The cumulative area of the base of each volumetric block.
- FAR (Floor Area Ratio): The floor area ratio, calculated as the ratio of the total floor area to the site area. The FAR at this stage represents the computed result after model generation and may differ from the user's inputted expectation, although they are theoretically expected to be similar.
- Surface Area: The external surface area of the building. It is calculated by obtaining the union of all volumetric blocks, forming the volumetric primitive, and summing the areas of each surface on the primitive.
- Shape Coefficient: The shape coefficient, calculated as the ratio of the external surface area to the total volume. The total volume is obtained by multiplying the base area of each volumetric block by the inputted floor height.
- South-facing Area: The area of south-facing facades. In certain architectural functions and specific regional requirements, the number of rooms facing south is a crucial indicator. The south-facing area is determined by calculating the angle between the normal vectors of each surface of the primitive and the vector pointing in the true south direction (0, -1, 0). Angles in the range of [-π/2, π/2] are marked as facing south.

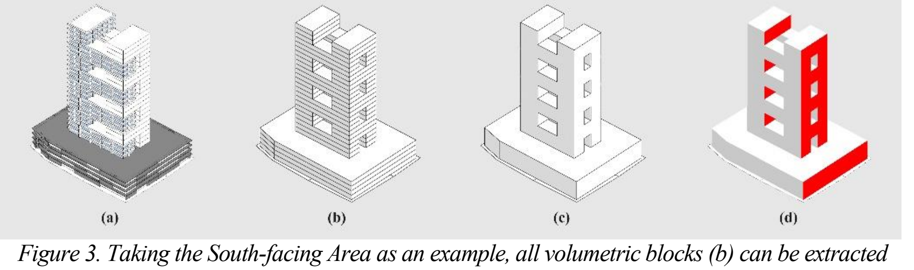


Figure 3. Taking the South-facing Area as an example, all volumetric blocks (b) can be extracted from the original generated model (a). The union of these blocks results in the total volumetric primitive (c). The specific faces can be identified and filtered from the primitive (d).

2.3. IMAGINARY SYNTHESIS

The AI Render backend employs diffusers to create a comprehensive pipeline. Operating on an NVIDIA RTX 2080Ti, it efficiently renders an image with a maximum edge of 1200 pixels in approximately 8 seconds before transmitting it back to the browser. Figure 4 depicts the complete workflow, showcasing how the browser captures preview frames from the canvas of the user-generated model. These frames are then relayed to a Python module in the backend where OpenCV's Canny algorithm retrieves their edges. These edge images function as control parameters, and together with user-defined generation parameters such as Render Style, Sampling Method, and Sampling Step, they integrate with the outputs generated by the pretrained Stable Diffusion model.

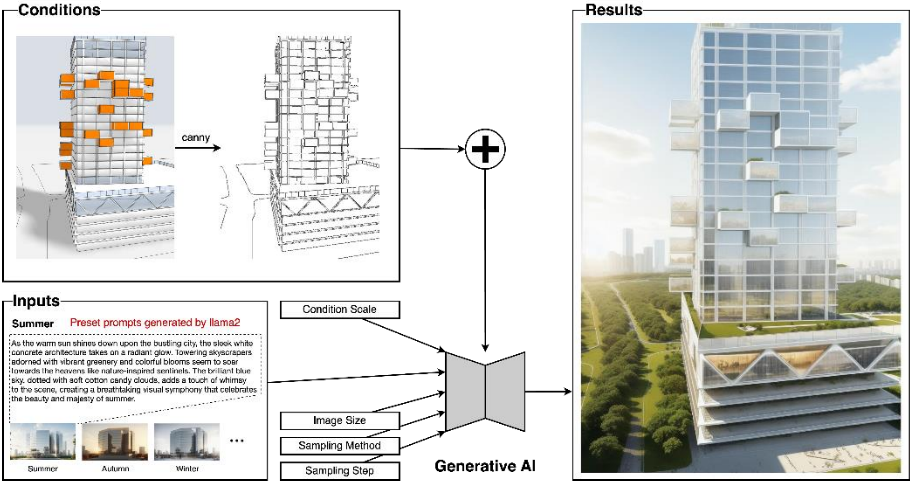


Figure 4. Backend workflow of imaginary synthesis with generative AI

The intervention of AI-generated images has accelerated the visual representation of parametric modelling, allowing users to evaluate and adjust the generated architectural forms based on conceptual images. It in itself represents a fascinating creative avenue for architecture.

3. SIMForms: An Integrated Web-based Application

SIMForms, as a web application, adopts the classical Browser-Server architecture, consisting of 1 browser module and 2 server modules: a web page based on the Vue framework, a modelling and metric calculation module compiled by the Java environment, and an image generation module deployed through Python. To facilitate data interconnection between modules and interactive control on the web page, the team developed the ArchiWeb platform, first released in 2021 (Mo, 2021). The SIMForms application has been deployed on a cloud server (Figure 5): https://web.archialgo.com/simforms/. The application comprises three operational modules: Site Configuration, Model Generator, and AI Renderer (Figure 6).

Figure 5. Main interface of SIMForms

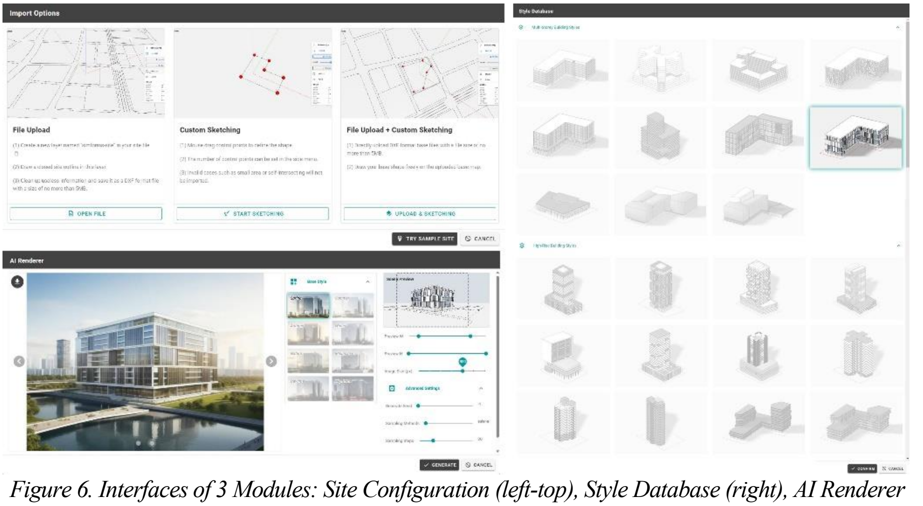


Figure 6. Interfaces of 3 Modules: Site Configuration (left-top), Style Database (right), AI Renderer (left-bottom)

The Site Configuration module requires users to choose one mode for defining the design base. Mode 1 allows users to upload a DXF file (a common AutoCAD file format) of the design site, which should include a layer named "simforms-site" with a polyline placed on this layer, identifiable in the application as the site outline. Mode 2 involves manual drawing, providing 3-8 control points that users can drag to define the desired site shape. Mode 3 is a combination of mode 1 and mode 2, allowing users to use the uploaded file as a base map and draw the desired site on it. Additionally, the application provides test case for users to quickly experience.

The Model Generation module consists of 23 selectable styles and three basic control parameters. The 23 styles are categorized into two groups based on multi-storey and high-rise buildings. A style database panel is provided for selection. After selecting one style, users can adjust three parameters: storey height, setback distance, and FAR expectation, generating the corresponding model. Each time a model is generated or updated, eight metrics are simultaneously calculated and updated on the data panel area. During the model generation stage, aggressive site shapes or parameters may result in the inability to generate effective models or excessively long generation times. The application provides feedback to users and suggests reconfiguring the shape or parameters.

The AI Renderer module includes camera parameters, rendering area, and rendering style. The camera parameters inherit the native camera adjustment functions of ArchiWeb, providing three functions: pre-set camera angle, isometric view, and camera FOV. After determining the related camera parameters, SIMForms uses the current scene as the original layout, allowing users to further select the rendering frame and image size. The application sends the rendering range, set by the user, to the AI Renderer backend. Current render parameters include eight base styles for convenient use: spring, summer, autumn, winter, nature, minimalism, daytime, and nighttime. Four advanced parameters are also introduced for fine-tuning the image generation effects: random seed, sampling method, samp ling step, and ControlNet conditioning scale. The generated images will be displayed in the preview box, supporting a maximum of the latest five concept images for comparison and download.

Figure 7 displays some of the building models and corresponding conceptual images generated through SIMForms. These examples are extracted from three semi-public tests conducted since SIMForms went online. Participants in the testing phase included students and professional architects, with a total of approximately 500 individuals. From October 17, 2023, to December 19, 2023, the server backend has recorded over 1700 sessions and 4100 operations, including model generation and image rendering actions. The purpose of the tests was to assess the server's stability, generation speed, and bug identification. The showcased results demonstrate the application's adaptability under diverse user input conditions, providing users with flexible creative inspiration.

Figure 7. Website usage data and results of various building models with corresponding images

4. Conclusion

This study introduces a comprehensive application for early-stage architectural design, named SIMForms, which combines rule-based parametric modelling, AI image generation, and the web framework based on ArchiWeb. In this application, users can interactively define design sites through a lightweight web interface, choose architectural styles, control 3D model generation parameters, and quickly obtain visual representations of architectural concepts along with metrics feedback. The main contributions of this research include:

- Bridging the gap between architectural form, quantitative metrics, and visualization. In the conceptual design phase, this application aids architects in evaluating various architectural forms under specified metrics (volume combinations, height variations, storey count controls, etc.), providing rapid feedback on corresponding effects, and serving as decision support for subsequent design stages.
- Integrating rule-based generative algorithms with the Artificial Intelligence Generated Content (AIGC), enhancing the visual presentation effects of generative design and parametric modelling. This introduces a new approach, connecting algorithmically driven computational design results rather than static 3D models or hand-drawn sketches with image generation engines.
- The application on a web platform not only facilitates the promotion, dissemination, and popularization of different architectural styles but also supports architects in sharing and communicating with clients and the public.

Planned future work includes expanding and categorizing the building style database, automatic matching of concept images with building styles, and expanding performance-related measurements for more design scenarios. Advances in computational hardware, software, and AI models will further reduce the complexity and barriers to entry for computational design applications, making it possible to apply digital technologies to comprehensive and diverse projects.

Acknowledgements

The study is funded by The Carbon Emission Peak and Carbon Neutrality Innovative Science Foundation of Jiangsu Province “The key research and demonstration projects of future low-carbon emission buildings” (No. BE2022606) and National Natural Science Foundation of China (NSFC) “Research on Generative Design Method of Architectural Space Based on Morphological Analyses and Typological Combination” (No. 52378008).

References

Albaghajati, Z. M., Bettaieb, D. M., & Malek, R. B. (2023). Exploring text-to-image application in architectural design: Insights and implications. *Architecture, Structures and Construction*, 3(4), 475–497. https://doi.org/10.1007/s44150-023-00103-x

Cheng, K., Neisch, P., & Cui, T. (2023). From concept to space: A new perspective on AIGC-involved attribute translation. *Digital Creativity*, 34(3), 211–229. https://doi.org/10.1080/14626268.2023.2248103

Fink, T., & Koenig, R. (2019). Integrated Parametric Urban Design in Grasshopper / Rhinoceros 3D Demonstrated on a Master Plan in Vienna. In *Proceedings of eCAADe / SIGraDi 2019* (pp. 313–322). Education and Research in Computer Aided Architectural Design in Europe (eCAADe) and Iberoamerican Society of Digital Graphics (SIGraDi). https://doi.org/10.5151/proceedings-ecaadesigradi2019_398

He, W., & Yang, X. (2019). Artificial Intelligence Design, from Research to Practice. In *Proceedings of the 2019 DigitalFUTURES* (pp. 189–198). Springer. https://doi.org/10.1007/978-981-13-8153-9_17

Lee, J., Gu, N., & Williams, A. (2014). Parametric Design Strategies for the Generation of Creative Designs. *International Journal of Architectural Computing*, 12, 263–282. https://doi.org/10.1260/1478-0771.12.3.263

Mo Y. (2021, August 25). *ArchiWeb*. Inst. AAA. Retrieved October 18, 2023, from https://web.archialgo.com/

Ortner, F. P., Yenardi, A. C., & Zhi, T. J. (2023). Urban Design Optimizer: A Comparative User Interface Study for a Web-Embedded Tool. In *2023 Annual Modeling and Simulation Conference (ANNSIM)* (pp. 616–627). The Society for Modeling and Simulation International (SCS).

Parish, Y. I. H., & Müller, P. (2001). Procedural modeling of cities. In *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques* (pp. 301–308). Association for Computing Machinery. https://doi.org/10.1145/383259.383292

Ploennigs, J., & Berger, M. (2023). AI art in architecture. *AI in Civil Engineering,* 2(1), 8. https://doi.org/10.1007/s43503-023-00018-y

Singh, V., & Gu, N. (2012). Towards an integrated generative design framework. *Design studies*, 33(2), 185-207.

Song, Z., Tang, P., & Song, Y. (2023). Digital Application of Typo-morphology in the Conservation and Renewal of Historic Areas. In *HUMAN-CENTRIC - Proceedings of the 28th CAADRIA Conference* (pp. 545–554). The Association for Computer-Aided Architectural Design Research in Asia (CAADRIA). https://doi.org/10.52842/conf.caadria.2023.1.545

Xu, Q. M., Garcia Del Castillo Lopez, J. L., & Samuelson, H. W. (2023). Towards a Decision Framework Integrating Physics-Based Simulation and Machine Learning in Conceptual Design. In *HUMAN-CENTRIC - Proceedings of the 28th CAADRIA Conference* (pp. 371–380). The Association for Computer-Aided Architectural Design Research in Asia (CAADRIA). https://doi.org/10.52842/conf.caadria.2023.2.371

Zhang, Q., Li, B., Mo, Y., Chen, Y., & Tang, P. (2022). A Web-based Interactive Tool for Urban Fabric Generation: A Case Study of Chinese Rural Context. In *POST-CARBON - Proceedings of the 27th CAADRIA Conference* (pp. 625–634). The Association for Computer-Aided Architectural Design Research in Asia (CAADRIA). https://doi.org/10.52842/conf.caadria.2022.1.625

TURNING EMPIRICAL KNOWLEDGE INTO DIGITAL ASSETS

A workflow assisting intern architects in high-rise conceptual design

CHENGYU SUN[1], YILIN LI[2] and YUELI WANG[3]
[1,2,3]Tongji University; Shanghai Key Laboratory of Urban Renewal and Spatial Optimization Technology.
[1]ibund@126.com, 0000-0002-5686-5957
[2]erica175134@gmail.com, 0009-0000-1172-642X
[3]499630837@qq.com, 0009-0009-4903-0909

Abstract. Experienced architects can quickly formulate high-rise conceptual designs that align with both the clients' needs and building codes through estimation and logical deduction. Usually, such empirical knowledge takes architects years of practice to accumulate. This research aims to externalize this empirical knowledge into generative design method, thereby aiding intern architects in quickly gaining the skills needed to conceive high-rise conceptual designs. We have developed a prototype program and a corresponding workflow, which can assist architects in generating and presenting high-rise designs featuring a rectangular plan. Additionally, when architects make adjustments to the building mass, the program can perform some preliminary verification on the new design.

Keywords. High Rise, Conceptual Design, Empirical Knowledge, Design Automation, Knowledge Transmission, Generative Design.

1. Introduction

Due to safety, economy, and cityscape protection factors, clients and urban planning departments have strict design criteria on high-rise buildings. Often during the conceptual design phase, experienced senior architects can quickly formulate feasible solutions through estimation and logical deduction, which is something most intern architects aren't capable of. Therefore, current conceptual design process of high-rise buildings is always led by a senior architect who first determines the basic architectural form and spatial layout. Mid-level and intern architects then work together to create detailed drawings and 3D models to complete the design. This common workflow creates a scenario where intern architects are often too engaged in drawing work to have the opportunity to learn the advanced experiences of senior architects in high-rise conceptual design. This hinders the progress of intern architects in terms of design capability and also obstructs the passing-on of empirical knowledge from senior architects.

This research aims at using generative design to bridge this gap. By translating the

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 383-392.

high-rise conceptual design strategies of senior architects into generative rules, this research turns empirical knowledge into digital assets that can be reused to guide intern architects in formulating a feasible design concept independently.

In the following sections, this paper presents a concise review of related studies on generative design and analyses whether these studies are helpful for guiding intern architects. Then, we introduce our generative-design prototype program and a corresponding workflow, which aimed at assisting intern architects in high-rise conceptual design. The prototype is verified through a simulated design practice. The final part of the paper concludes the strengths and limitations of this research and outlines goals for future work.

2. Related Studies

Based on the logic that drives generation, generative design can be categorized into the following three types: rule-based generation, artificial intelligence (AI) algorithm-based generation and integration of rule-based and AI approaches. Rule-based generative design employs predefined design rules and principles as a structured framework to steer the design process. AI algorithm-based generative design leverages the power of artificial intelligence techniques such as genetic algorithms, neural networks, and evolutionary strategies, enabling the system to learn, adapt, and evolve solutions iteratively.

2.1. RULE-BASED GENERATIVE DESIGN

Donath and Lobos (2009) translated building codes and clients' needs into Boolean operation rules that subtracted mass from the original box volume to generate the envelopes of high-rise buildings. The program automatically generated floor/ section/ elevation plans, industry foundation classes (IFC) model and indicator report for each design. Abdullah and Kamara (2013) demonstrated the capability of parametric design in generating a large number of alternatives, with generative rules based primarily on the aesthetics of building form and facade. Yavuz and Celik (2015) employed shape grammar to generate house units, which were then copied and combined according to specific rules to form the design of an entire high-rise residential building. Erhan and Shireen (2017) introduced a methodology that enabled designers to generate design alternatives derived from a base parametric model, facilitating the simultaneous parameter adjustment of multiple solutions by designers.

2.2. AI ALGORITHM-BASED GENERATIVE DESIGN

Rafiq et al. (2003) employed genetic algorithm to optimize both the total building area and structural material cost, determining the optimal structural layout plan for rectangular, frame-structure high-rise buildings. Elnimeiri and Nicknam (2011) employed genetic algorithm generates the final form of the building through multi-objective optimization of energy conservation, structural performance and morphological constraints. Gane and Haymaker (2012) proposed a methodology called Design Scenarios (DSs) to take into account to as many design goals from various stakeholders as possible during the conceptual design phase. The program optimized the weighted average of these goals. As et al. (2018) presented a graphically based automatic learning system to generate conceptual designs. Deep neural networks were trained to evaluate existing house designs, decomposing them into subgraphs to extract semantically rich building blocks, and then merge the blocks into new compositions. Lee et al. (2023) used algorithms to move and scale the several original volumes based on grid points, generating a large number of building site planning alternatives.

2.3. INTEGRATION OF RULE-BASED AND AI APPROACHES

Sydora and Stroulia (2020) introduced a generative design method that utilized a BIM model representing an empty space, a predefined set of furnishings to be placed, and a set of layout rules as input. Through multiple iterations of adjusting furniture placement and rule-checking, the optimization algorithm progressively completed the interior layout. Abrishami et al. (2021) presented a prototype named G-BIM for residential building layout generation. The building model was constraint by several rules such as number of spaces, site boundary and outline dimensions. Leveraging spatial topology, internal space layouts were iteratively adjusted and optimized by genetic algorithm.

2.4. IMPLICATIONS DRAWN FROM PREVIOUS STUDIES

Some rule-based method research expects to directly generate feasible architectural designs by adding abundant constraints. However, in most practices, constraints primarily serve to prevent extreme situations, and most design decision-making still relies on the architect's analysis and judgement. Other rule-based method research first generates multiple design alternatives, and then uses artificial intelligence algorithms to identify the relatively optimized solution. The algorithm would grow more and more complex as the architect increases the number of optimization goals and adjustable parameters. Consequently, existing research has primarily focused on single-objective and bi-objective optimization problems (Ekici et al., 2019), which falls short in aiding intern architects to comprehensively accomplish the design goals.

In recent years, there has been a growing interest in utilizing deep learning for the generation of architectural designs. However, due to the intricacies of the data collection process and challenges in computer image recognition, designs generated by these algorithms are not yet mature enough for practical use. Depending solely on algorithms to generate designs may also hinder the development of design capabilities for intern architects.

Our research suggests that previous research faced limitations in generating designs comparable to those of senior architects, which can be attributed to the omission of “empirical process” and “approximate values” when developing the generative method. Senior architects would use approximate values to estimated the rough volume and layout of the building, and gradually add detail to the design through a certain empirical process. This saves them time and effort when optimizing multiple design goals. We believe that externalizing this empirical knowledge is not only beneficial for intern architects in passing on the experience of the senior architects, but also enhances the efficiency of current high-rise conceptual design process.

3. Methodology

3.1. FIVE STAGE HIGH-RISE CONCEPTUAL DESIGN WORKFLOW

To recreate the senior architects' empirical design process, the design workflow is segmented into five stages, as illustrated in Figure 1. In Stage One, the program receives two key indicators from the architect: gross floor area and height limit of the building. It then immediately estimates sub-indicators for the architect. Number of floors is calculated by dividing height limit by a preset rough floor-height value. Typical floor area is equal to gross floor area divided by number of floors. The area of each core component, such as elevators and staircases, is estimated separately using the formulas provided by TOSPUR Real Estate Consulting Co. Ltd (2016). These sub-indicators will guide the architect in the following design process. In Stage Two, intern architects engage in the primary parametric design process. The program guides them with input hints, and shows them if the design aligns with the building criteria. In Stage Three, the program automatically generates drawings and a model of the designed building. In Stage Four, architects have the freedom to modified the building's mass using the 3D modelling software which they are familiar with. In Stage Five, the program assists architects in evaluating the modified building mass from Stage Four. The program cuts the mass according to the building section designed in Stage Three, generates new floor plan sketches and compares them with the original floor plan. All the generated drawings can be automatically exported to PowerPoint, facilitating the creation of a brief presentation.

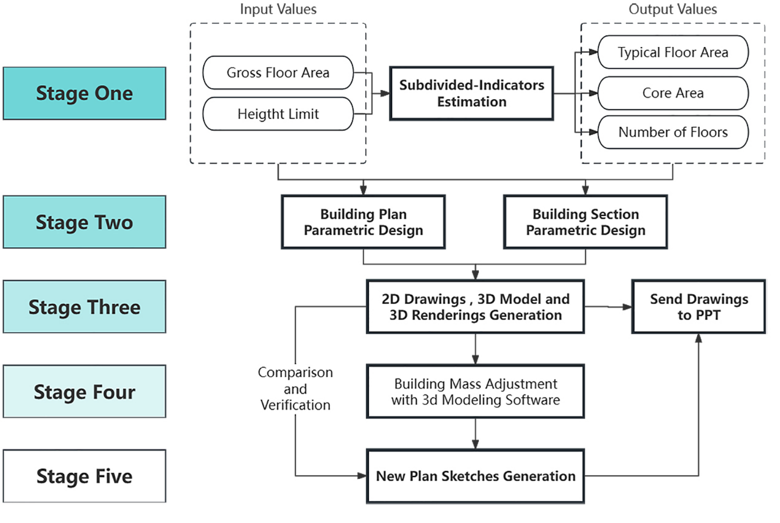


Figure 1. Five Stage High-rise Conceptual Design Workflow

3.2. CONTRIBUTION

3.2.1. Provision of Approximate Values

The program provides plenty of approximate values, which are gathered from literature reviews and research on design institutes in Shanghai, to intern architects through input prompts or predefined options. For instance, in plan parametric design procedure, there are three most commonly used lease span value— 10.5m, 12m, and 13.5m (Mirniazmandan et al., 2018) —for the architects to choose from. However, they are also granted the flexibility to input other values.

3.2.2. Design Check Based on Building Codes

Since intern architects are often not familiar enough with various building codes, they may enter parameters which generate designs violating those codes, for instance, planning only one staircase in a fire compartment. These mistakes are fatal to the design. Once the architect enters wrong parameters, our program will give an instant warning. This not only prevents intern architects from working in vain on invalid designs, but also helps them become familiar with the codes.

3.2.3. Immediate Calculation of Indicators

To ensure both the economic efficiency and safety of the high-rise, design criteria impose strict regulations on numerous building indicators, such as height limit, gross floor area of the building, and so on. During the conceptual design stage, architects dedicate significant effort to adjust the building design, ensuring those indicators comply with the criteria. Usually, these indicators are calculated manually, and an architect has to do lots of recalculation whenever adjustments are made to the design. This scenario not only leads to substantial time wastage, but also diminish the confidence of intern architects. Our research addresses this issue by providing architects with real-time indicator feedback through automated calculations. This enables architects to promptly grasp the outcome of their decisions, empowering them to make adjustments more decisively and efficiently.

3.2.4. Automatic Generation of Drawings and Presentations

In the traditional workflow, architects have to dedicate significant time to manually create drawings and presentations of their designs. In our proposed workflow, once the architect completes parametric design process and clicks "Generate", the program automatically generates plans, sections, axonometric drawings, and renderings of the design. It can also insert these drawings, along with corresponding building information text, into a PowerPoint presentation. This automation alleviates the drafting workload for intern architects, enabling them to invest more time in the creative aspects of design.

4. Prototype

Using two-dimensional drawings to represent architectural designs is still the mainstream in China. Consequently, the prototype program was developed mainly based on the AutoCAD platform. In this section, the prototype would be demonstrated through a simulated design.

According to the design brief, the architect was assigned to design a high-rise building with a height limit of 180 meters and an anticipated gross floor area of 92,000 square meters.

In Stage One, the architect entered 'Gross Floor Area' and 'Height Limit' parameters (highlighted in red in Figure 2), then clicks the 'Run Estimation' button. The program automatically filled in other sub-indicators. The estimated ratio of core area to typical floor area is 21.95%, which falls within the feasible range (Li & Gao, 2020). The program also reminded that according to the building code, current design needed a sprinkler system.

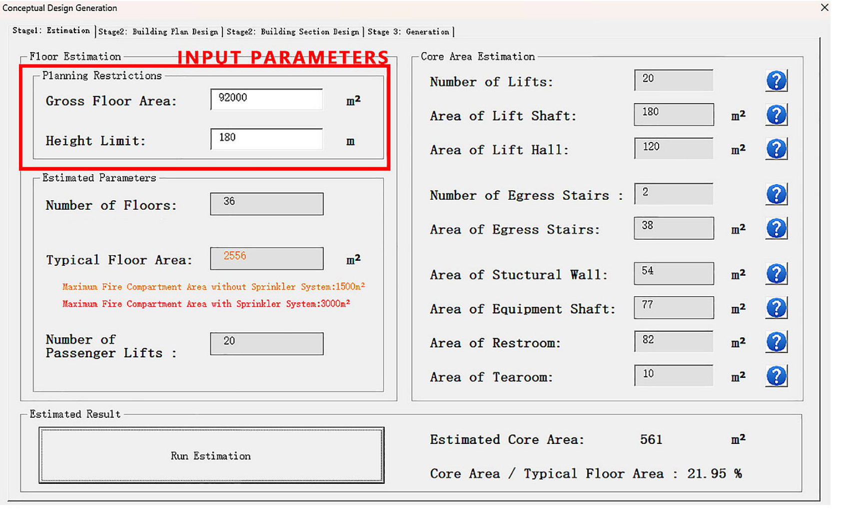


Figure 2. Stage One Interface Demonstration

In Stage Two, the architect adjusted plan and section parameters, as illustrated in Figure 3 and Figure 4. The 'Indicator Check' panel displayed the area deviation between the architect's design and the goal.

To design building plan, the architect adjusted the column grid and the building outline dimension by modifying the column spacing and column number. The "Core Width" was calculated by subtracting twice the "Lease Span" from the "Building Width", ensuring the space between the exterior wall and the core wall was economically efficient for office layout. The upper-right window in Figure 3 provided a preview of the building plan, allowing the architect to know whether some columns were too close to the core before generation of CAD drawings.

As for building-section design, the architect vertically zoned the building by editing the floor list, assigning different floor heights and numbers of floors to different zones.

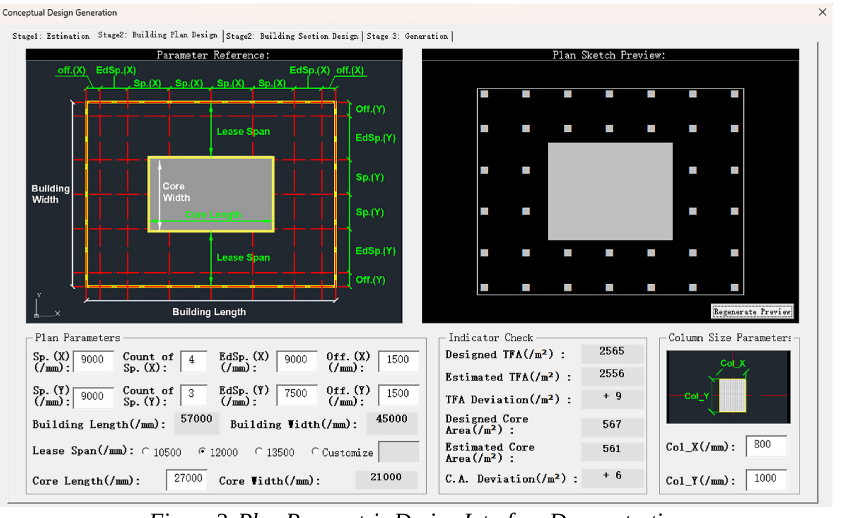


Figure 3. Plan Parametric Design Interface Demonstration

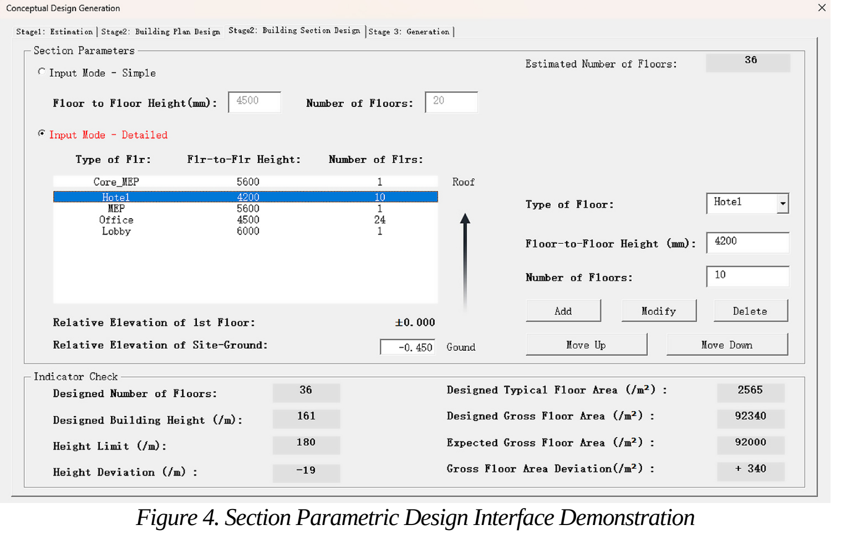


Figure 4. Section Parametric Design Interface Demonstration

With the design of plan and section settled, the architect clicked "Generation", and the program automatically drew various drawings (some of which are shown in Figure 5). The renderings are generated by the POV-Ray engine, using a pre-set site template.

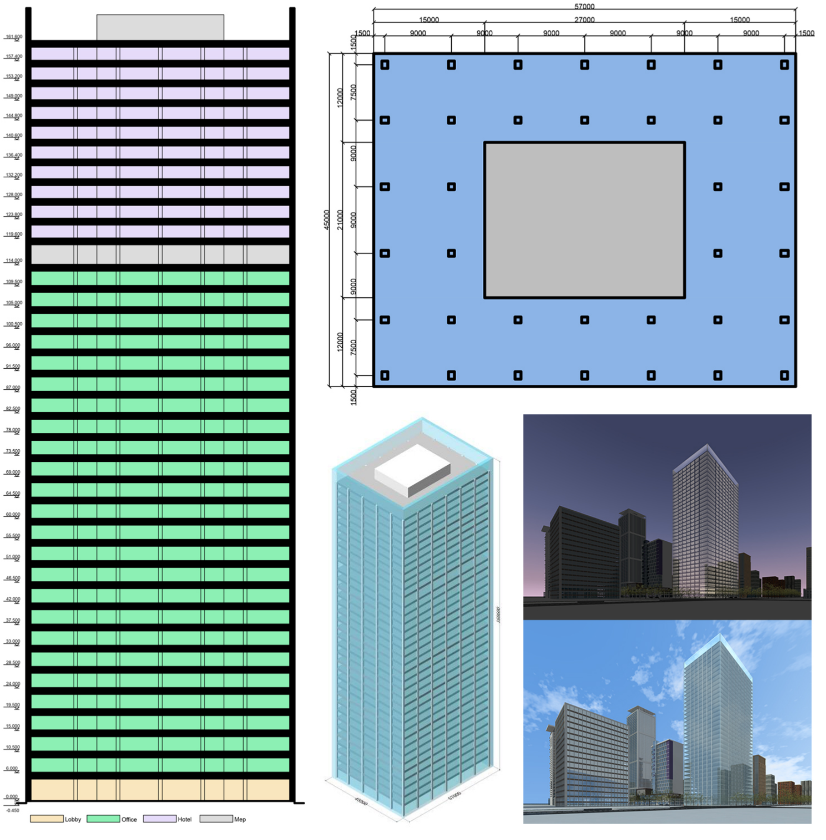


Figure 5. Design Drawings Generated by the Prototype

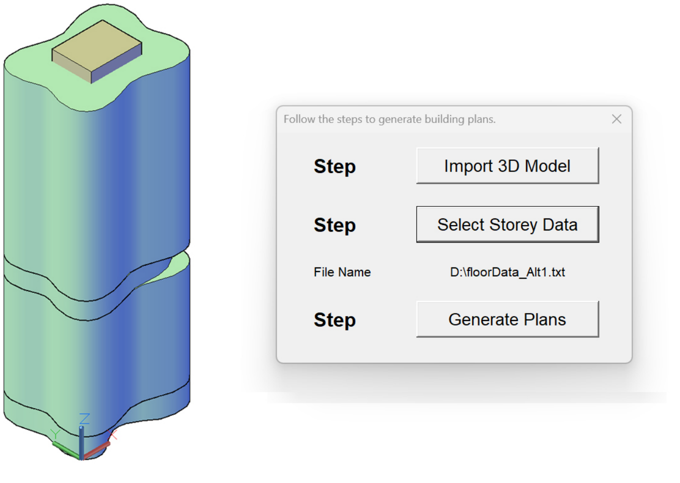


Figure 6. Stage Five Interface Demonstration

In Stage Four, the architect operated on the building's mass using Rhino platform, adjusting the building's outline to a curved shape and making some mass subtraction. The architect then returned to the AutoCAD platform to execute Stage Five. The architect imported the new building mass model into a new CAD drawing, selected the building floor information file of the original design generated in Stage Three, and clicked 'Generate Plans' (see Figure 6). The program generated floor plans of the new mass and compared them with the original structural layout. The area of each floor and the gross floor area were calculated, and the columns that appeared outside of the floor outline were marked out in red (see Figure 7).

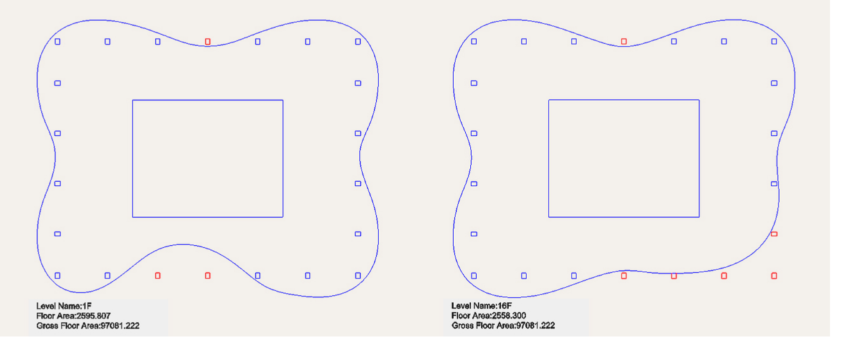


Figure 7. New Plan Sketches Generated in Stage Five

5. Conclusion and Future Work

The senior architects collaborating on this study believed that the prototype can greatly reproduce their usual high-rise-design process. Provision of approximate values, immediate calculation of indicators and quick generation of design drafts spare architects great time when exploring alternatives, enabling intern architects to independently and methodically design high-rises. This proves the feasibility of the prototype in externalizing senior architects' design empirical knowledge.

However, the prototype was mainly developed based on the spatial organization logic of the high-rise building itself, and the site of the building has not yet been taken into account. Also, since the prototype was developed based on the AutoCAD platform, it is mainly good at generating two-dimensional drawings. When architects need to modify the 3D building model, they have to use other modelling software, and there will be information loss during the export and import process.

According to the deficiencies in the above two aspects, future work on the one hand will look for ways to link two-dimensional and three-dimensional building information, and on the other hand, will continue to externalize senior architects’ empirical knowledge on high-rise site planning.

Acknowledgements

We would like to express our sincere gratitude to the Shanghai Key Laboratory of Urban Renewal and Spatial Optimization Technology for their generous support and assistance. With their expertise, the research was enabled to explore in-depth the

complex principles of high-rise design.

References

Abdullah, H. K., & Kamara, J. M. (2013). Parametric design procedures: a new approach to generative-form in the conceptual design phase. In *AEI 2013: Building Solutions for Architectural Engineering* (pp. 334-343).

Abrishami, S., Goulding, J., & Rahimian, F. (2021). Generative BIM workspace for AEC conceptual design automation: Prototype development. *Engineering, Construction and Architectural Management,* 28(2), 482-509.

As, I., Pal, S., & Basu, P. (2018). Artificial intelligence in architecture: Generating conceptual design via deep learning. *International Journal of Architectural Computing,* 16(4), 306-327.

Donath, D., & Lobos, D. (2009). Plausibility in early stages of architectural design: a new tool for high-rise residential buildings. *Tsinghua science and technology*, 14(3), 327-332.

Ekici, B., Cubukcuoglu, C., Turrin, M., & Sariyildiz, I. S. (2019). Performative computational architecture using swarm and evolutionary optimisation: A review. *Building and environment, 147,* 356-371.

Elnimeiri, M., & Nicknam, M. (2011, October). A Design Optimization Workflow for Tall Buildings Using Parametric Algorithm. In *Proceedings of the CTBUH 2011 Seoul Conference, Seoul, Korea* (pp. 10-12).

Erhan, H., & Shireen, N. (2017). Juxtaposed designs models: A method for parallel exploration in parametric CAD. In G. Çagdas, M. Özkar, L. F. Gül, & E. Gürer (Eds.), *Future trajectories of computation in design: Proceedings of the 17th International Conference on Computer-Aided Architectural Design Futures (CAAD Futures 2017)* (pp. 402-412). ISBN 978-975-561-482-3.

Gane, V., & Haymaker, J. (2012). Design Scenarios: Enabling transparent parametric design spaces. *Advanced Engineering Informatics,* 26(3), 618-640.

Lee, J., Cho, W., Kim, S., Sohn, D., & Lee, J. (2023). Conceptual design algorithm configuration using generative design techniques. *KIEAE Journal,* 23(1), 5-12.

Li, Y., & Gao, Y. (2020). Chaogaoceng bangong jianzhu biaozhunceng shiyong mianji xishu yanjiu[Study on the floor area ratio of standard floors in super high-rise office buildings]. *Architecture and Culture*, 2020(07), 264-265.

Mirniazmandan, S., Alaghmandan, M., Barazande, F., & Rahimianzarif, E. (2018). Mutual effect of geometric modifications and diagrid structure on structural optimization of tall buildings. *Architectural science review*, 61(6), 371-383.

Rafiq, M. Y., Mathews, J. D., & Bullock, G. N. (2003). Conceptual building design—Evolutionary approach. *Journal of computing in civil engineering,* 17(3), 150-158.

Sydora, C., & Stroulia, E. (2020). Rule-based compliance checking and generative design for building interiors using BIM. *Automation in Construction, 120,* 103368.

TOSPUR Real Estate Consulting Co. Ltd. (2016, June 10). *Chaogaoceng bangong jianzhu hexintong qianxi[A brief analysis of the core of super high-rise office buildings].* Winshang. Retrieved February 15, 2024, from http://file2.winshang.com/file3/2016/6/10/%E8%B5%A2%E5%95%86%E5%85%B1%E4%BA%AB-0113%E8%B6%85%E9%AB%98%E5%B1%82%E5%8A%9E%E5%85%AC%E5%BB%BA%E7%AD%91%E6%A0%B8%E5%BF%83%E7%AD%92%E6%B5%85%E6%9E%9013573394.pdf.

Yavuz, A. O., & Celik, T. (2015). A Rule Based Computer Aided Model Suggestion Developed for Design Process of High-Rise Housing. *Journal of Academic Research,* 7(3), 533.

Performance-based Design & Analytics

A KNOWLEDGE GRAPH MODEL FOR PERFORMANCE-BASED GENERATIVE DESIGN AND ITS APPLICATIONS IN ACCELERATED DESIGN

ZHAOJI WU[1], ZHE WANG[2], JACK C. P. CHENG[3] and HELEN H. L. KWOK[4]

[1,2,3] Department of Civil and Environmental Engineering, The Hong Kong University of Science and Technology, Hong Kong SAR, China.
[2] HKUST Shenzhen-Hong Kong Collaborative Innovation Research Institute, Futian, Shenzhen, China.
[4] Institute for the Environment, The Hong Kong University of Science and Technology, Hong Kong SAR, China.
[1] zwubz@connect.ust.hk, 0000-0002-6739-7909
[2] cezhewang@ust.hk, 0000-0002-2231-1606
[3] cejcheng@ust.hk, 0000-0002-1722-2617
[4] hlkwokab@connect.ust.hk, 0000-0002-7179-9281

Abstract. Data integration and information enrichment pose significant challenges to the advancement of Performance-based Generative Design (PGD). One potential solution to these challenges is the utilization of Knowledge Graph (KG). However, the implementation of KG in PGD, particularly in leveraging expert knowledge to accelerate the process, remains an area that has not been thoroughly explored. In this research, we propose a PGD-KG schema to capture and represent the topological relationships and functionalities within PGD. We also introduce a method for automatically generating PGD-KG models from parametric design models enriched with semantic information. Additionally, we develop reasoning algorithms based on expert knowledge of sustainable design to facilitate automated performance evaluation. The effectiveness of the PGD-KG approach was demonstrated through its implementation in a design project, where the reasoning algorithms proved capable of significantly reducing the solution space in PGD by 88.50%, while still ensuring the inclusion of an adequate number of Pareto optimal solutions. This research contributes to the design acceleration by integration of expert knowledge, particularly sustainable design strategies, into PGD.

Keywords. Performance-based generative design, Knowledge graph, Reasoning algorithm, Building performance evaluation, Sustainable building design.

1. Introduction

The demand for green buildings has gained significant momentum, driven by the urgent need to alleviate the environmental impact of the built environment. As shown

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 395-404.

by The MacLeamy Curve (Davis, 2011), the earlier stage in a project, the less costly it is to make a change. Hence, it is important to incorporate green performance during the early design stage of buildings. Early design is a complex task which involves multiple factors, e.g., building programs, constraints, parameters, and objectives. Performance-based generative design (PGD) (or performative computational architecture (Ekici et al., 2019)) is a powerful tool to mitigate the complexity of early design with the assistance of computational methods and artificial intelligence. In the current practices of PGD, architects create and manipulate a geometric model in parametric design tools (e.g., Rhino and Grasshopper). However, to conduct building performance evaluation, more information (e.g., materials and constructions) is needed beyond geometry (Negendahl, 2015). Although some tools (e.g., Ladybug Tools) have been developed to facilitate the information enrichment in the process of PGD, architects still need to collect and input the information manually or semi-automatically, which requires expert knowledge in building evaluation. Consequently, data integration and information enrichment are two important issues that hinder the acceleration process of PGD.

A knowledge graph (KG) represents knowledge in a structured and interconnected manner, capturing the relationships between entities and properties among multiple domains (Ji et al., 2022). By providing a unified schema and semantic framework, KG enables seamless integration of diverse data sources (Pauwels et al., 2017). In addition, the graph-based representation of knowledge in KG enables automated reasoning, which can uncover implicit relationships and enrich the semantic information of the KG model (Pauwels et al., 2017). With these advantages, KG could provide a solution to the issues of data integration and information enrichment for PGD. In addition, several studies (Machairas et al., 2014; Su & Yan, 2015) have stated that expert knowledge can help accelerate PGD. For example, design knowledge from experts or architects could help minimize the size of the solution space (Machairas et al., 2014) and customize the optimization process (Su & Yan, 2015). However, how to implement KG to PGD and accelerate PGD with expert knowledge has not been fully studied yet.

The purpose of this research is to develop a KG model for PGD and apply it to the process of PGD. Our specific goals are: 1) to design the KG schema for PGD (PGD-KG) that can integrate data from multiple domains, 2) to propose the PGD-KG generation method from parametric design models with enriched semantics, 3) to integrate expert knowledge to accelerate PGD by developing PGD-KG-based reasoning algorithms for automatic performance evaluation. Finally, one detached house design project is selected as the illustrative example of the implementation of the PGD-KG, and performance of the PGD-KG in accelerated design is evaluated.

2. Methodology

2.1. PGD-KG SCHEMA

Previously, we proposed an ontology model for building energy modelling (Wu et al., 2023), which integrates four domains (weather, building, internal heat gain and Heating, Ventilation, and Air Conditioning (HVAC) system) based on the existing models, Brick Schema (Balaji et al., 2018) and Building Topology Ontology (BOT)

(Rasmussen et al., 2020). The KG model schema for PGD is developed by leveraging the previous ontology model (Wu et al., 2023), with necessary extensions implemented to comprehensively capture and represent the topological relations and functionalities in PGD. Figure 1 shows the overview of the PGD-KG schema that covers five key

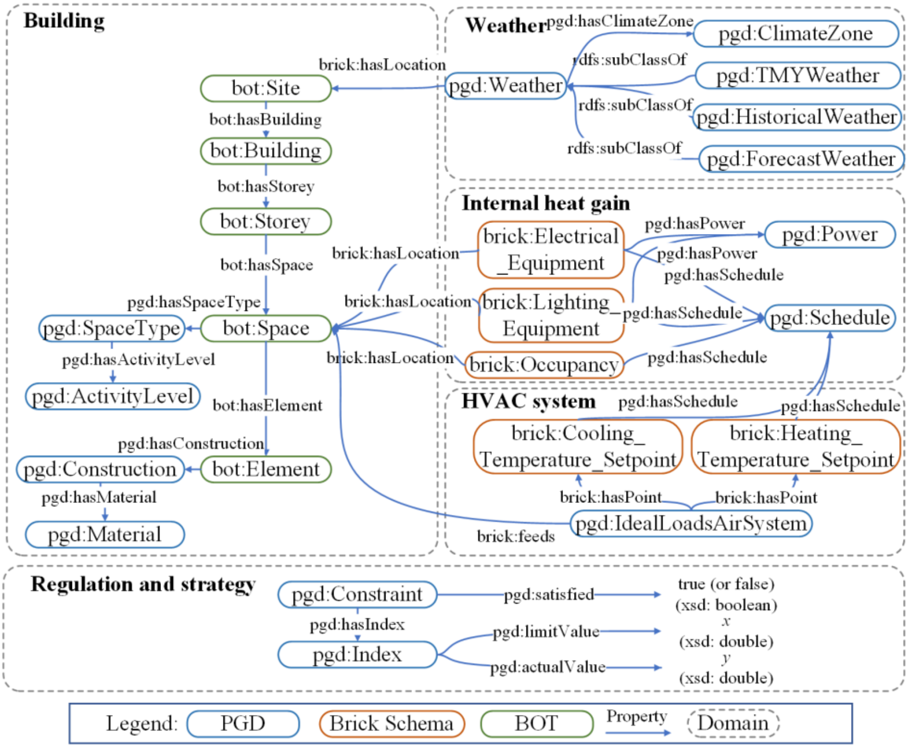


Figure 1. PGD-KG schema

domains in PGD, including building, weather, internal heat gain, HVAC system, and regulation and strategy. The components beyond the existing ontology models (Brick Schema and BOT) are assigned a prefix pgd.

Building is the important domain in the PGD-KG schema. The class bot:Site is assigned with the property pgd:has Area_m2 to support the calculation of site-area-related indices, e.g., plot ratio. The class bot:Building stores the information on building types and building directions. The class bot:Space describes spatial zones, and stores the geometric information (ceiling heights, floor areas, volumes, etc.). The topology property pgd:hasAirFlow is assigned to the class bot:Space to describe the topology relation of airflow among spaces. Each bot:Space is connected with the class pgd:SpaceType that indicates the space's function and the activity level of occupants in the space. The class bot:Element describes surfaces including building surfaces (walls, floors, roofs, ceilings, air walls, etc.), fenestration surfaces (windows, doors, etc.), and shading surfaces. Topology relations in a building surface contain the information on surface types, geometries (surface areas, vertex coordinates, normal vectors, etc.), boundary conditions (whether the surface is exposed to the outdoor environment), constructions and materials (thermal properties, optical properties, etc.). Air walls are commonly adopted to describe large openings between two spaces in modelling software, e.g., EnergyPlus. In the PGD-KG schema, air walls are regarded as a kind of building surfaces with the construction AirBoundary, which is aligned with the setting in EnergyPlus. The topology relations in a fenestration surface and a shading surface are similar to the ones in a building surface but the material properties are different. Minor revisions of the components in the domains of weather, internal heat gain and HVAC system are made compared with the previous ontology model (Wu et al., 2023) and details of these components can be referred to Section 3.1 in Ref. (Wu et al., 2023). The graphs and files of the PGD-KG model can be found on https://github.com/GeorgeZWu/PGD_KG_Schema.

2.2. PGD-KG MODEL GENERATION

Figure 2 depicts the workflow of the PGD-KG model generation. Firstly, spaces and surfaces in parametric design models are translated into individuals under the classes bot:Space and bot:Element in the PGD-KG model. Secondly, properties of the spaces and surfaces, including volumes and floor areas of spaces, types, areas, vertex coordinates and normal vectors of surfaces, etc., are identified or calculated. Corresponding KG properties (pgd:hasVolume_m3, pgd:hasFloorArea_m2, pgd:hasSurfaceType, pgd:hasArea_m2, pgd:hasVertex1X_m, pgd:hasNormalVectorX_m, etc.) are generated. Thirdly, semantic information including weather, space type, etc. is translated into individuals under the classes pgd:Weather, pgd:SpaceType, etc. of the PGD-KG model. The semantic information is derived from default templates which could be reused in various projects, or the semantic information is manually input by designers for specific cases. Finally, knowledge inference is conducted based on the generated PGD-KG model, and boundary conditions and topology relations of airflow are supplemented to the PGD-KG model.

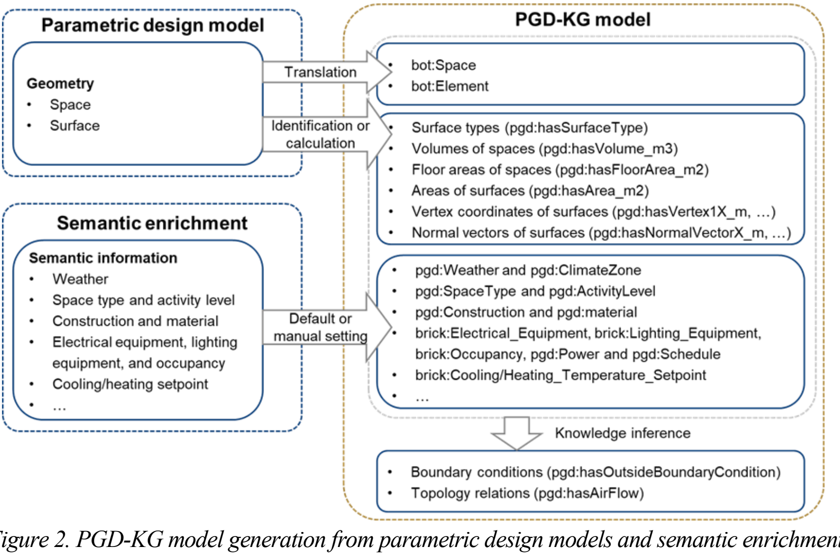


Figure 2. PGD-KG model generation from parametric design models and semantic enrichment.

2.3. REASONING FOR PERFROMANCE EVALUATION

Figure 3 elaborates the extraction workflow of design strategies based on expert knowledge from text-based literature to machine-readable reasoning algorithms. Firstly, design strategies are summarized from the literature on sustainable building design. In this research, three design strategies (summarized in Table 1) are extracted from literature (Brown & DeKay, 2013; Heywood, 2019; Kwok & Grondzik, 2018). Secondly, the design strategies are translated into machine-readable algorithms that support cross-domain reasoning of the PGD-KG models. Table 2 shows one example of the reasoning algorithm in SPARQL for the number of nodes in the overall airflow path, the indicator of Strategy 2.

Following the execution of the algorithms, the metrics for the three indicators can be derived. The sustainable performance of design cases can be evaluated based on the rankings of these indicators. To determine the number of selected cases, a hyperparameter called the threshold is introduced. For example, if the threshold is 40%, the cases with indicators ranking within the top 40% for all strategies will be chosen.

Table 1. Design strategies derived from literature.

	Interpretation	Indicator	Performance
1	The area of exterior envelopes should be as less as possible.	Shape factor	Cooling and heating
2	The ventilation path should be sufficiently short, and the resistance should be sufficiently low.	Number of nodes in overall airflow path	Cooling
3	The light-transmitting windows should be mainly directed to the south or north, and the light-transmitting windows on the south and north side should be sufficiently large.	Projection area of light-transmitting windows on the south and north side	Daylighting and cooling

Table 2. Reasoning for the number of nodes in overall airflow path.

Algorithm 1 (in SPARQL): Reasoning for the number of nodes in overall airflow path.

```
SELECT (COUNT(DISTINCT ?intermediateNode) AS ?nodeCount)
WHERE {
    ?window1 rdf:type bot:Element .
    ?window1 pgd:hasSurfaceType "Window" .
    ?window1 pgd:hasOutsideBoundaryCondition "Outdoors" .
    ?window2 rdf:type bot:Element .
    ?window2 pgd:hasSurfaceType "Window" .
    ?window2 pgd:hasOutsideBoundaryCondition "Outdoors" .
    ?window1 pgd:hasAirFlow* ?intermediateNode .
    ?intermediateNode pgd:hasAirFlow* ?window2 .
    }
```

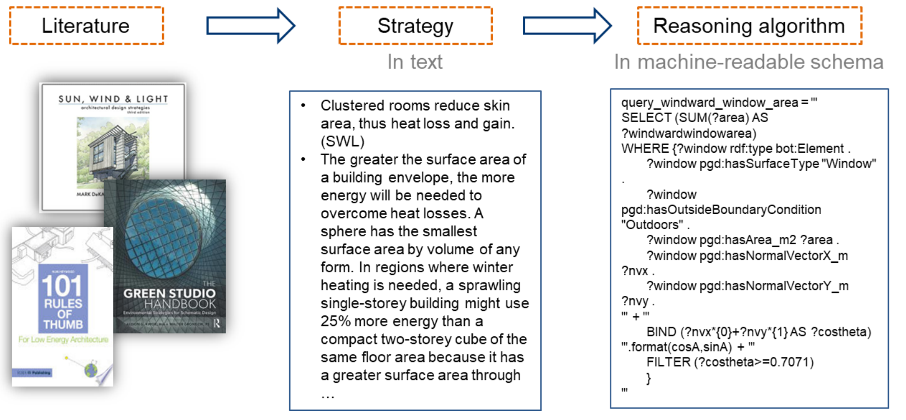


Figure 3. Extraction of design strategies from text-based literature to machine-readable reasoning algorithm.

3. Illustrative Example

3.1. INTRODUCTION OF THE DESIGN PROJECT

A 2-storey detached house using prefabricated modular techniques in a pilot design study was selected as the illustrative example of the implementation of the PGD-KG. Figure 4 illustrates the modelling rule of the prefabricated modular detached house. The sizes and locations of the staircase and living rooms were settled according to the design constants (Table 3), while the positions and orientations of other prefabricated modular units varied among five slots (A, B, C, D and E). There were two directions (along the x-axis and y-axis of the building coordinate) of the prefabricated modular units in slot A, C and D, while the prefabricated modular units in slot B and E were only directed along the y-axis. Five kinds of module combination were allocated in the five slots according to the design variables (Table 3). The goal of the design project was to generate optimal layouts of the module combinations that minimize the cooling energy consumption and maximize the daylighting performance of the detached house.

The implementation was programmed by Python. Firstly, the parametric design models of the detached house were developed using Python library rhino3dm, which offers geometric manipulation functions that can create and modify parametric Rhino 3D models in Python. Secondly, the PGD-KG models were generated based on the parametric design model and enriched semantics. Thirdly, reasoning for performance evaluation based on the three design strategies was conducted. The Python library rdflib was adopted to construct the PGD-KG models and conduct reasoning for the performance evaluation. We studied the sizes of the solution spaces selected by the reasoning algorithms for performance evaluation under different thresholds (40%, 30%, 20% and 10%) with three metrics: the ratio of selected solutions (the number of solutions selected by the reasoning algorithms divided by the total number of cases), the ratio of selected Pareto optimal solutions (the number of selected Pareto optimal solutions divided by the number of Pareto optimal solutions in the ground truth) and computational time reduction (compared with simulations of all the cases). In addition, all the cases (7680 in total) were simulated by EnergyPlus and Radiance to obtain their performance metrics Annual Cooling Load (ACL) and Useful Daylight Illuminance (UDI) (Nabil & Mardaljevic, 2006) as the ground truth of the performance.

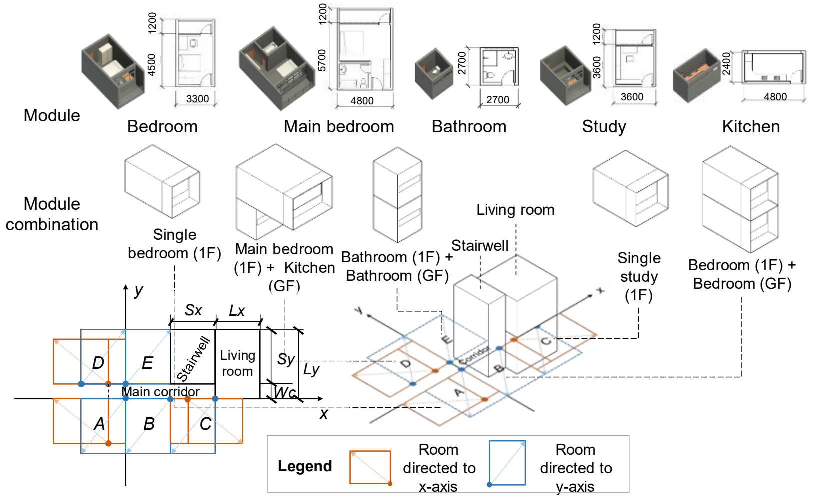


Figure 4. The modelling rule of the prefabricated modular detached house.

Table 3. Design parameters of the prefabricated modular detached house.

Parameter		Description	Value
Constant	Wc	Corridor width	1.2 m
	H	Floor height	3.3 m
	Lx	Living room width	3.0 m
	Ly	Living room depth	4.8 m
	Sx	Stairwell width	1.8 m
	Sy	Stairwell depth	3.6 m
Decision variable	Sa	Combination in slot A	Range: [0, 4], step: 1
	Sb	Combination in slot B	0: Single bedroom (1F);
	Sc	Combination in slot C	1: Main bedroom (1F) + Kitchen (GF);
	Sd	Combination in slot D	2: Bathroom (1F) + Bathroom (GF);
	Se	Combination in slot E	3: Single study (1F);
			4: Bedroom (1F) + Bedroom (GF).
	Da	Direction of the combination in slot A	Range: [0, 1], step: 1
	Dc	Direction of the combination in slot C	0: along y-axis;
	Dd	Direction of the combination in slot D	1: along x-axis.
	O	Orientation	Range: [0,360°), step: 45°

3.2. RESULTS AND DISCUSSION

The PGD-KG models of all the design cases could be generated from their parametric design models with enriched semantics. Fig. 13 shows the classes, properties, individuals, and metrics of the PGD-KG model of one example (Case 6725) that is visualized in Protégé.

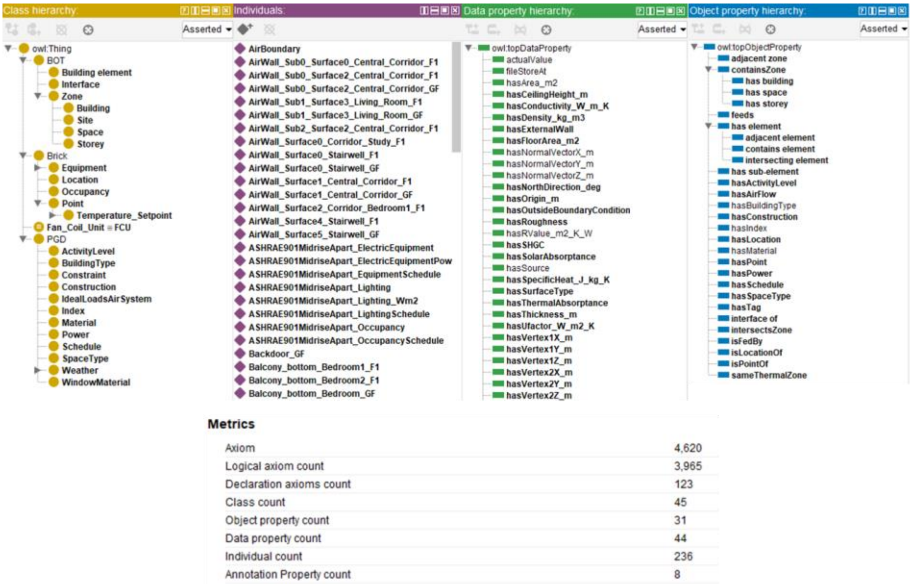
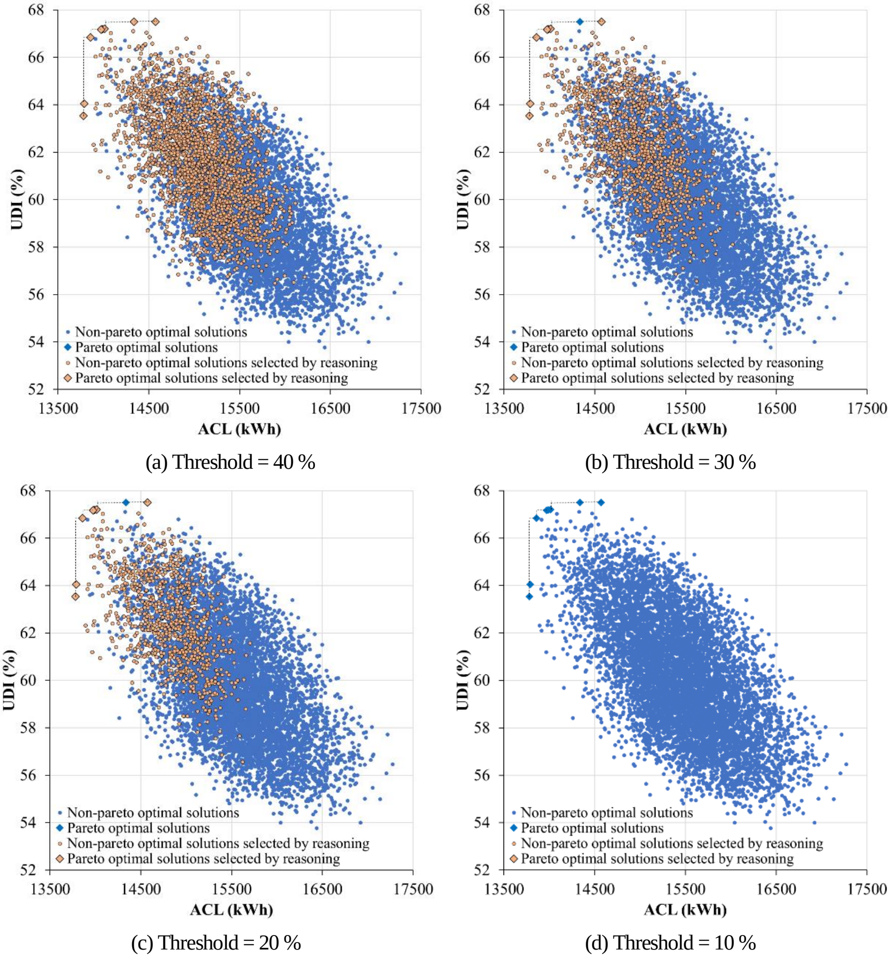


Figure 5. Classes, properties, individuals, and metrics of the PGD-KG model of Case 6725, visualized in Protégé.

Table 4. Performance of the PGD-KG reasoning under different thresholds.

Threshold (%)	Ratio of selected solutions (%)	Ratio of selected Pareto optimal solutions (Selected / ground truth)	Computational time reduction (%)
40	31.04	8/8	64.73
30	16.33	7/8	75.25
20	11.50	7/8	75.21
10	0	0/8	0

Table 4 shows the performance of the PGD-KG reasoning and Figure 5 shows the selected solution spaces by the reasoning algorithms for performance evaluation under different thresholds (40%, 30%, 20% and 10%). As the threshold was more stringent (i.e., smaller in magnitude), the ratio of the selected solutions became smaller and the selected solution space contracted, orienting towards lower ACL and higher UDI values. When the threshold was equal to 40%, all the Pareto optimal solutions were included in the selected solution space. When the threshold was 30% and 20%, one Pareto optimal solution was not included. When the threshold was 10%, no solution was selected, showing that no case satisfies all the strategies if the threshold was too stringent. In the illustrative example, it took averagely 11.0 s to conduct performance evaluation reasoning while it needed averagely 135 s to conduct performance simulation for each case. The reasoning algorithms can effectively narrow down the solution space in PGD while ensuring the inclusion of an adequate number of Pareto optimal solutions (≥ 87.5% in the illustrative example). The narrower solution space leads to less cases to be simulated and analysed, and hence time for simulations in PGD is reduced.



(a) Threshold = 40 %

(b) Threshold = 30 %

(c) Threshold = 20 %

(d) Threshold = 10 %

Figure 6. Selected solution spaces under different thresholds.

4. Conclusions

In this research, we propose a PGD-KG schema that effectively captures and represents the topological relationships and functionalities within Performance-based Generative Design (PGD). Furthermore, we introduce a generation method that automatically generates PGD-KG models enriched with instances from parametric design models, incorporating semantic information. Additionally, we develop PGD-KG-based reasoning algorithms for automated performance evaluation. To evaluate the effectiveness of the PGD-KG approach, we implemented it in a design project and observed that the reasoning algorithms successfully narrowed down the solution space in PGD, while ensuring the inclusion of a satisfactory number of Pareto optimal solutions. This research realizes the unsolved ideas proposed by the previous studies including minimizing solution spaces (Machairas et al., 2014) and customizing optimization processes (Su & Yan, 2015) by the automated integration of expert knowledge into PGD. However, only one design case is illustrated in this research. In the future, more design projects should be studied to further evaluate the efficacy and reliability of the proposed method.

References

Balaji, B., Bhattacharya, A., Fierro, G., Gao, J., Gluck, J., Hong, D., Johansen, A., Koh, J., Ploennigs, J., Agarwal, Y., Bergés, M., Culler, D., Gupta, R. K., Kjærgaard, M. B., Srivastava, M., & Whitehouse, K. (2018). Brick: Metadata schema for portable smart building applications. Applied Energy, 226, 1273–1292. https://doi.org/10.1016/j.apenergy.2018.02.091

Brown, G., & DeKay, M. (2013). Sun, wind, and light (The 2nd edition). Wiley.

Davis, D. (2011). The MacLeamy curve. The MacLeamy Curve. https://www.danieldavis.com/macleamy/

Ekici, B., Cubukcuoglu, C., Turrin, M., & Sariyildiz, I. S. (2019). Performative computational architecture using swarm and evolutionary optimisation: A review. Building and Environment, 147, 356–371. https://doi.org/10.1016/j.buildenv.2018.10.023

Heywood, H. (2019). 101 rules of thumb for low energy architecture. Riba publishing.

Ji, S., Pan, S., Cambria, E., Marttinen, P., & Yu, P. S. (2022). A Survey on Knowledge Graphs: Representation, Acquisition, and Applications. IEEE Transactions on Neural Networks and Learning Systems, 33(2), 494–514. https://doi.org/10.1109/TNNLS.2021.3070843

Kwok, A. G., & Grondzik, W. (2018). The green studio handbook: Environmental strategies for schematic design. Routledge.

Machairas, V., Tsangrassoulis, A., & Axarli, K. (2014). Algorithms for optimization of building design: A review. Renewable and Sustainable Energy Reviews, 31, 101–112. https://doi.org/10.1016/j.rser.2013.11.036

Nabil, A., & Mardaljevic, J. (2006). Useful daylight illuminances: A replacement for daylight factors. Energy and Buildings, 38(7), 905–913.

Negendahl, K. (2015). Building performance simulation in the early design stage: An introduction to integrated dynamic models. Automation in Construction, 54, 39–53. https://doi.org/10.1016/j.autcon.2015.03.002

Pauwels, P., Zhang, S., & Lee, Y.-C. (2017). Semantic web technologies in AEC industry: A literature overview. Automation in Construction, 73, 145–165. https://doi.org/10.1016/j.autcon.2016.10.003

Rasmussen, M. H., Lefrançois, M., Schneider, G. F., & Pauwels, P. (2020). BOT: The building topology ontology of the W3C linked building data group. Semantic Web, 12(1), 143–161. https://doi.org/10.3233/SW-200385

Su, Z., & Yan, W. (2015). A fast genetic algorithm for solving architectural design optimization problems. Artificial Intelligence for Engineering Design, Analysis and Manufacturing, 29(4), 457–469. https://doi.org/10.1017/S089006041500044X

Wu, Z., Cheng, J. C. P., Wang, Z., & Kwok, H. H. L. (2023). An ontology-based framework for automatic building energy modeling with thermal zoning. Energy and Buildings, 296, 113267. https://doi.org/10.1016/j.enbuild.2023.113267

A PARAMETRIC METHOD IN ROOM ACOUSTICS SIMULATION WITH PERFORMANCE VERIFICATION

Real-time Ray Tracing Techniques in Parametric Modeling

BO-YOUNG LIN[1] and WEI LIN[2]
[1,2] School of Architecture, Feng Chia University
[1]bob230198@gmail.com, 0009-0004-0682-7057
[2]wlin@fcu.edu.tw, 0000-0001-9682-3316

Abstract. Since not all architectural projects can be involved in performance design calculations or optimizations, this study aims to reduce user technical barriers and enhance the visualization of the evaluation process. This paper introduces a parametric approach to room acoustics simulation, offering an interactive interface for modeling and simulation. By using Rhino-Grasshopper as a real-time ray tracing interface, both the Image Source Method (ISM) and Early Scattered Method (ESM) are employed for ray particle rebound tracking and analysis. The feasibility and analysis of spatial acoustic assessment tools are explored through the combination of data visualization and open-source parameter design meth. This research diversifies spatial models using various modeling tools like Revit, ArchiCAD 25, Rhinoceros 3ds, and SketchUp. This approach enables the evaluation of the effects and disparities caused by different modeling systems on simulation calculations. Furthermore, this paper also highlights the potential optimization of the Building Information Modeling (BIM) design workflow process. The results underscore the relationship between improved simulation accuracy and the utilization of numerical calculations, referencing benchmark simulation software ODEON for data comparison and review.

Keywords. room acoustics simulation, parametric modeling, rhino-grasshopper, real-time analysis, ray tracing methods, SDG 9

1. Introduction

This research conducts the scrutinization of the rationale and system framework methods employed in room acoustic and ray-tracing computations. Simultaneously, the research addresses the nonlinear workflow challenges faced by designers, users, and researchers in their respective tasks (Tan et al., 2019). Furthermore, the previous intricate logic demands complex settings have come to access. Additionally, plug-in software for room acoustics and design optimization with Graphic User Interface (GUI) based workflow has eased the use and performance simulation works.

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 405-414.

1.1. ROOM ACOUSTICS SIMULATION

The interplay between material absorption and acoustic performance is pivotal. Both sound and material hold a crucial position and can yield divisive effects contingent upon the properties of the materials in question (Brandão et al., 2022). Studies dedicated to appraising and contrasting software configurations tailored for divisive room acoustic settings (Lombardo et al., 2020) and the progress in Computation Aided Design (CAD) and Algorithm Aided Design (AAD) provide access to material and simulation databases, previous studies, processes, and achievements as shown in Figure 1.

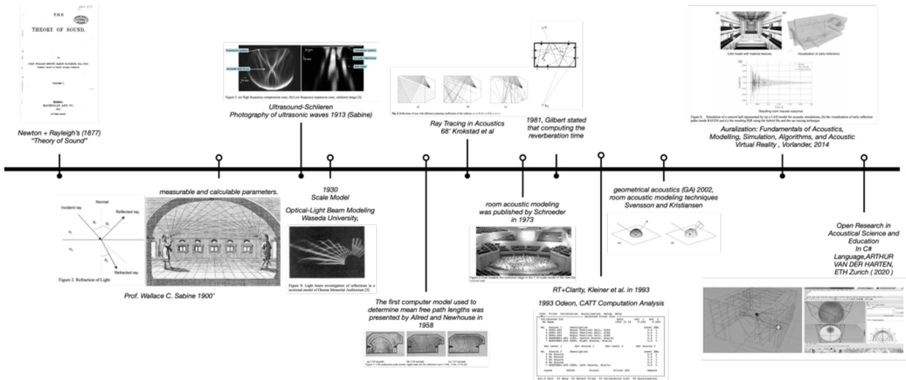


Figure 1. Studies in Room Acoustic Simulation and Measurement Methods

This research delves into the examination and development of boundaries and objective acoustic indices, as well as evaluation and computation methods. This exploration spans from the analogical era to contemporary algorithm-based calculations. The intention is to dissect the components of acoustic assessment and experimental approaches, providing insight and workflow into procedures underpinning the simulation of acoustic performance.

2. Background

2.1. GEOMETRY ACOUSTICS

The introduction of the Non-Uniform Rational B-splines (NURBS) algorithm (Echenagucia et al., 2013) provides a significant mark in room acoustics simulation. When assessing the impacts between sound and multidimensional surfaces (mesh-based 3D files), imperative the move beyond 2D Geometry-based computation. We re-examined and depicted the effects of NURBS operations on energy reflections from surfaces constructed with three-dimensional files. The scattering in three-dimensional variation states and energy intersection points disperses in a multi-dimensional pattern. NURBS algorithm has become a pivotal development in acoustics, particularly in the calculation of sound interactions with multidimensional surfaces, as highlighted in the work in Geometry Acoustics (Savioja et al., 2015). This necessitates departing from the previous confined 2D Geometry and employing three-dimensional surfaces composed of knots to influence energy reflection and utilization of the Rhino-Grasshopper modeling tool in room acoustic calculations (Bassuet et al., 2014), streamlines subsequent operations and research thresholds, enabling the calculation of 3D projections through the Rhino-Grasshopper plug-in method (Van der Harten et al., 2019), employed for ray-tracing geometry calculations in 3D Boundary Representation (BREP) based environment.

2.2. ROOM ACOUSTICS SIMULATION WORKFLOW

The process of performance-oriented design and simulation, followed by a nonlinear workflow, requires a complete rerun for adjustments in design. In the revamped design workflow, using a Python-based plug-in introduces a preliminary performance graph in the pre-modeling stages. Additionally, a genetic algorithm is employed for material optimization during software correction calculations. This approach allows the material search to be conducted from a performance-oriented perspective, avoiding the need to search for materials in the database individually. Significantly streamlines the entry into the Detailed Design (DD) phase, saving hours of repetitive searching and operation simulation by simplifying the searching of data and results. Workflow Optimization can be achieved by utilizing the INSUL Software system database and manufacturers' laboratory databases. The primary tools for acoustic performance at this stage include EASE, CATT, and ODEON Software. In this research, ODEON Room Acoustics Simulation Software 15 is employed for calculations based on geometrical acoustics. A combination of the ISM and the ray-tracing method is utilized for Algorithm-based acoustic simulations, incorporating visualization capabilities. Design changes within the indoor reflection, material absorption, and interface scattering coefficient also address various uses, musical performances, speech intelligibility, and loudspeaker system inputs, in which the physical parameters measured predominantly focus on sound rays, as shown in Figure 2.

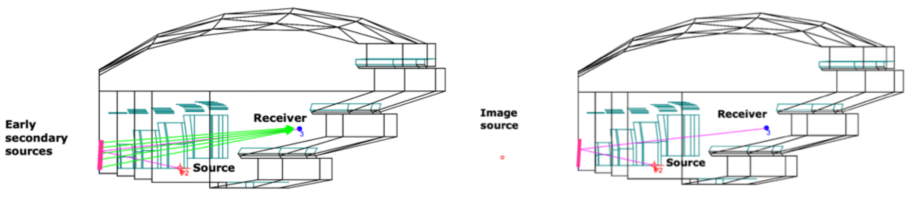


Figure 2. Ray-Tracing Methods in Geometry Acoustics (ODEON User Manual)

3. Research Methodology

Software for simulating and evaluating geometric acoustic performance has progressively advanced and matured. Utilizing computation simulations and the advantage of reducing the time and labor costs associated with constructing large-scale models for scaled tastings, thereby making performance design more accessible. Presently, room acoustics simulation relies on Geometry acoustic controls, which manage indoor reflection, material absorption, and interface scattering coefficients for various purposes such as musical performances, speech clarity, and input for speaker systems. The primary physical parameters measured pertain to sound energy. ODEON software stands out as a professional tool for audio performance simulation grounded in geometrical acoustics. Computer-based sound field calculations involve a combination of both ISM and ESM ray-tracing methods, with outcomes in sound dispersion that are then visualized for data visualization.

3.1. PERFORMANCE ORIENTED DESIGN AND MATERIAL

This research centers on spaces where speech intelligibility stands as the primary performance requirement, focusing on the exploration and establishment of Performance-Oriented Design (POD) principles for space dimensions and the treatment of indoor sound-absorbing materials such as ceilings, walls, and seating. Additionally, we presented the workflow and cycle of the BIM process. In this system, computer-aided simulation software, specifically ODEON and its derivative parameter tools, is utilized for simulation results, serving as a reference for data. This aids in understanding the optimization process of building planning design cycles and performance design, particularly for numerical comparisons related to voice clarity. Six distinct room types are scrutinized concerning verbal spaces. This research involves simulation and the impact of diffusion on the clarity of sound energy distribution between users and seats, particularly focusing on reflection's impact on performance design. It encompasses testing spatial design variations, evaluating their effects, and optimizing and enhancing the acoustic performance of verbal spaces.

Ceilings in interior design serve not only an aesthetic role but also play a crucial part in incorporating sound-absorbing materials into key areas of modern design vocabulary. This is achieved through a unitized design approach using a lightweight steel frame and sound-absorbing materials. The production design adheres to imported specifications, employing 35mm Wood Wool Cement Board (WWCB) with air gap. This strategy is devised to fulfill formal conditions, optimizing spatial performance in terms of the material's specifications and performance. The replacement of partition walls with multi-layered glass-based facades emphasizes the importance of wall performance and acoustic treatment. In the design of walls and facades, the upper wall is retained for accommodating electrical wiring, air conditioning, and equipment piping channels above door height. The material selection focuses on localized sound-absorbing treatment, utilizing aluminum-based acoustic panels to mitigate reverberation and sound reflection effects, enhancing speech clarity. The selection of seating arrangements is pivotal to the acoustic environment's performance. In this study, where the sound-absorbing material area was relatively limited, the visual presentation of the seat design significantly influenced its performance. The study relied on laboratory data from sound-absorbing seat manufacturers to comprehensively assess room characteristics and seat specifications, as illustrated in Figure 3.

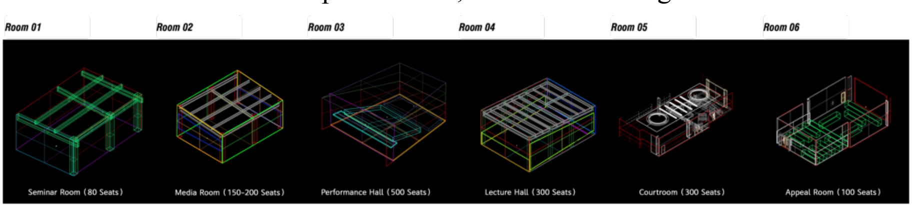


Figure 3. Room Typology and Characteristics for Acoustics Simulation

3.2. PARAMETRIC RAY-TRACING METHOD

Raytracing involves calculating the impacts of energy and particle paths using the ray method, with the energy particles diminishing after five impacts, as in the experimental setting. The utilization of derivatives significantly enhances the connection between performance and spatial design at the initial stage. This is achieved by promptly providing users with feedback data and visual patterns, aiding in their comprehension of the relationship between spatial configurations and performance. Additionally, it presents the numerical state of the content. The calculation results were sequentially presented at 63Hz, 125Hz, 500Hz, 1000Hz, 2KHz, 4KHz, and 8KHz. For initial data, algorithm-based & data-based simulation was employed. Through derivative tools, the relationship between performance and spatial design at the outset was markedly strengthened, offering users immediate feedback on data and visual patterns. This facilitated a clearer understanding of the correlation between spatial configurations and performance.

4. Room Acoustics Simulation

The use of the Rhino-Grasshopper interface has led to the modularization of algorithmic tools through plug-in tools, make easy for designers and firms to operate. By breaking down and reassembling the compositional steps of acoustic algorithms, it becomes feasible to sequentially acquire spatial design and acoustic performance graphics. This is achieved through the projective line method and surface calculation. The process allows for simulation and adjustments in parametric design mode, facilitated by the Grasshopper plug-in. This feature enables real-time modifications to the design, serving as a helpful reference during the Conceptual Design (CD) phases, simulation, and cross-interface workflow as shown in Figure 4.

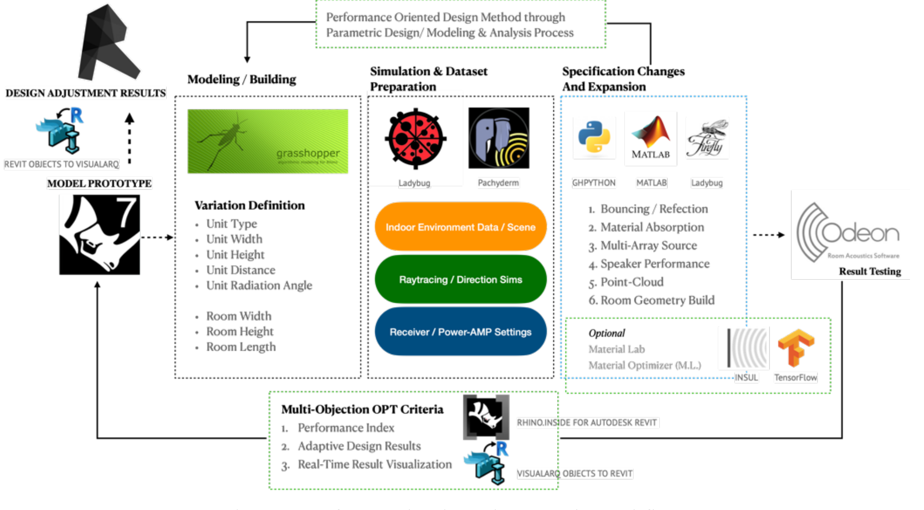


Figure 4. Software plug-in and Research Workflow

Reverberation Time Index (T30) and Clarity Index (C80) present calculated results with data visualization in the interaction between various software and testing methods for each room type. Comparing data from different software conducts use of standard deviation, revealing remarkably close standard deviation values in larger room types. One factor contributing to this was the modeling of seating in larger spaces to generate a greater surface area of sound-absorbing material. This aspect was thoroughly investigated in subsequent studies to validate the connection between modeling methods and the obtained data. In larger spaces, seats were specifically modeled to increase the surface area of sound-absorbing materials. The computational logic indicated that Rhino-Grasshopper-based surface calculation closely aligned with ODEON surface calculation when standard deviation was utilized as the accuracy metric for data comparison. The study found variations in data accuracy under different modeling tools and types, emphasizing the impact of the state and number of open-source tools on simulated values, settings, and methods as shown in Figure 5.

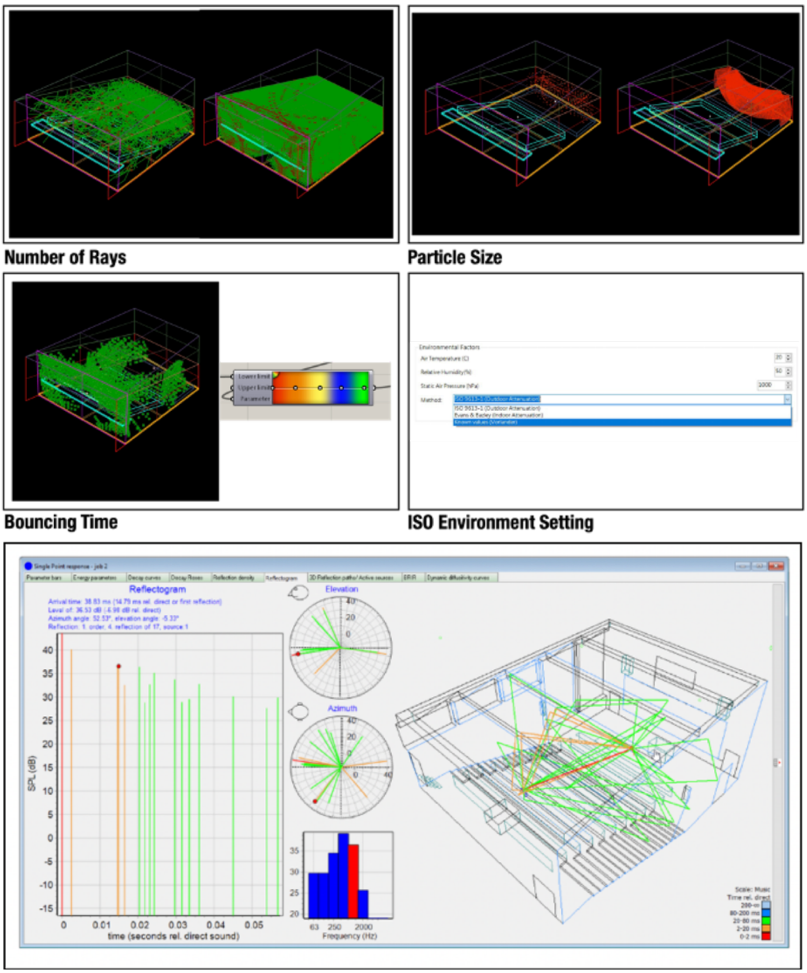


Figure 5. Ray-Tracing Acoustics Simulation in Rhino-Grasshopper and ODEON Acoustics Software

5. Results and Discussion

The software for simulating and evaluating the acoustic environment through geometrical acoustics. The ease of use in simulations offered the advantage of time and cost reduction and also physically scaled models and tests, thereby making performance design more accessible. In this study, the credibility of the tool and the results were established through comparison and examination of data between this tool and the ODEON performance software after presenting performance indexes. The ODEON software algorithm's simulation data and material library were derived from the material performance coefficient database. To standardize the control and influence of variables on reverberation time, the study compares the differences in reverberation time and clarity values between the two tools. This involves minimizing variable variations in control and parameter acquisition while focusing on the spatial performance of different chambers. Under identical material performance settings and configurations, the study assesses the differences in Reverberation Time Index (T30) performance, with data obtained from the interactive adjustment of different software and test methods for each chamber type. Comparisons of standard deviation in the data from different software revealed extremely close values, particularly in medium and large chamber types, results shown in Figure 6.

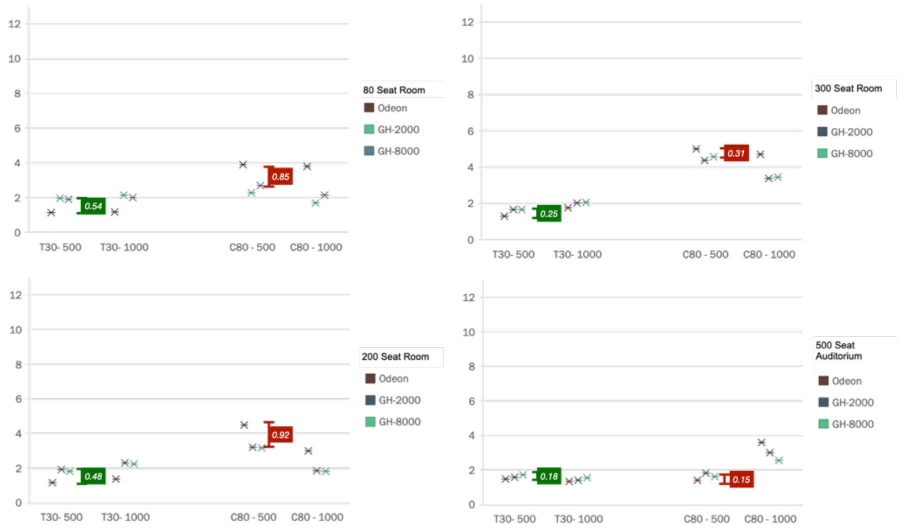


Figure 6. Simulation Results in Comparison

Linked with modeling tools such as Revit and Rhino-Grasshopper 3D enables the system to conduct real-time analysis and calculations. This functionality assists designers and users in promptly comprehending the consequences of modifications to spatial functionality. Consequently, in the intermediate phase of this study, our focus shifted towards enhancing the accuracy of the simulation. We achieved this by utilizing ODEON, a numerical and metric simulation software, as a reference benchmark. The spatial models generated through various modeling tools, including Revit, ArchiCAD 25, Rhinoceros 3D, and SketchUp, served as the basis for our examination of the impacts and disparities in the simulation operations introduced by different modeling systems.

5.1. CONCLUSIONS

This paper introduces a parametric approach to room acoustics simulation using Rhino-Grasshopper allowing interactive modeling and real-time acoustic analysis through ray tracing methods. The parametric workflow allows rapid acoustic analysis during early design stages. The Image Source Method (ISM) and Early Scattered Method (ESM) are used for ray particle tracking. Various modeling tools like Revit, ArchiCAD, Rhino, and SketchUp are used to create spatial models. The paper demonstrates a parametric methodology for interactive room acoustics simulation and analysis. It optimizes the building design process by integrating performance simulation. The parametric workflow may assist participants in getting closing comments in the very early stage of the SD phase.

In the current non-linear workflow of architectural design and construction, this research tool can intervene and streamline the design process and work stages. Through the optimization of the cross-interface/system data modeling shown in Figure 7.

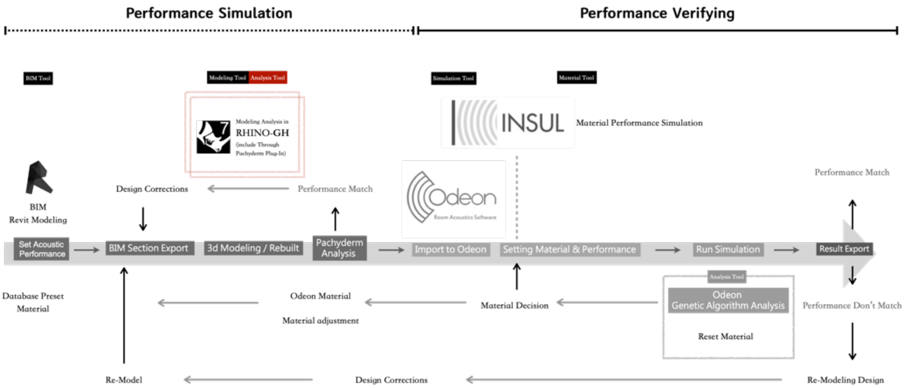


Figure 7. Acoustic Simulation Workflow

Integration coupled with suitable corrections and fine-tuning, it becomes feasible to achieve data results that closely align with those of benchmark simulation software.

The indoor acoustic design/simulation workflow is structured as a bidirectional process involving parametric design and metrics software. This design aims to reduce simulation time and enhance the effectiveness of each tool in the simulation process. For room modeling and research, a singular omnidirectional sound source from ODEON simulation software is employed to simulate and assess the teaching environment. The workflow, impacts, and data acquisition are analyzed. Additionally, the teaching environment is simulated with multiple lines of fire. The following are the conclusions and recommendations derived from this research.

Reverberation time and clarity index (C80) are simulated and compared between tools. The other parametric like IACC and LF which are related to the spatial factors may be discussed for the next stage.

5.1.1. Achievements

- Within the study, sound undergoes multiple reflections within the modeled sandwich thickness. However, in the real environment, the discrepancy arises from the fact that the reflection coefficient of sound traversing through the material doesn't align with the diffusion coefficient of sound absorption.
- The layer-based material setup proves more advantageous for calculating area and specifying materials. Consequently, not all models can be directly imported; some require fine-tuning post-import.
- Number of emission lines in Rhino-Grasshopper, it's observed that 2000 line of rays is closer to the ODEON's performance than 8000 line of rays, with the standard difference of the closest value being 0.04.
- Constructing models with seating, reflective panels, and other objects using linear modeling as opposed to surface structures such as BREP or Mesh files can further reduce the standard deviation of residence numbers, thereby enhancing the accuracy of simulation values.

5.1.2. Limitations and Future Works

- In the BIM workflow, exporting the model in CAD format (.dwg) provides easier analysis than other formats, ensuring more accurate calculations through the logic of linear modeling.
- Given correct modeling, indoor acoustic analysis values can be rapidly obtained, saving significant time in workflow execution. This flexibility allows operators to make corrections compared to software such as ODEON.
- The study employs a single omnidirectional sound source for simulation, without the addition of an electroacoustic system or directional sound source.
- Settings and fine-tuning: In configuration and refinement, this study delves into critical aspects, notably, the calibration of ray numbers, with the ongoing inquiry into whether restricting rebounds to five instances might be overly conservative. Moreover, the incorporation of multiple sound sources is indispensable for a nuanced and comprehensive examination, where both performance and source diversity are pivotal considerations. Additionally, advancements in input/output mechanisms and settings not only contribute to the sophistication of testing environments but also underscore a commitment to refining experimental precision and methodology.
- Calculation optimization: Our research initiates an extensive examination across diverse modeling platforms, emphasizing the need to align these platforms under more analogous conditions. This entails meticulous fine-tuning of file import and transcoding settings. Simultaneously, to enhance the linearity of our simulation methodology, dedicated efforts are directed toward refining Grasshopper-Rhino-based plug-in tools. The intricate comparison between on-site and scaled model results serves as a pivotal focus, representing a significant stride in our ongoing academic research endeavors. This approach not only contributes to the robustness

of our simulations but also holds implications for the broader field of [insert relevant field] by establishing a foundation for more standardized and comparable modeling practices.

Acknowledgments

We express our gratitude to the Room & Building Acoustics Lab at Feng Chia University for their technical support and extend our appreciation to the Construction Planning and Operations Management Office for providing the educational version of the ODEON software, which greatly contributed to the success of our research.

References

Bassuet, A., Rife, D., & Dellatorre, L. (2014). Computational and optimization design in geometric acoustics. *Building Acoustics*, 21(1), 75-85. DOI: 10.1260/1351-010X.21.1.75

Brandão, E., Santos, E. S., Melo, V. S., Tenenbaum, R. A., & Mareze, P. H. (2022). On the performance investigation of distinct algorithms for room acoustics simulation. *Applied Acoustics*, 187, 108484. https://doi.org/10.1016/j.apacoust.2021.108484

Echenagucia, T. M., Shtrepi, A. A. M. S. L., & van der Harten, A. (2013, March). NURBS and mesh geometry in room acoustic ray-tracing simulation. *In Proceedings of the AIA-DAGA 2013 conference Merano, Merano-Italy*.

Kim, H., Remaggi, L., Fowler, S., Jackson, P. J., & Hilton, A. (2020). Acoustic room modeling using 360 stereo cameras. *IEEE Transactions on Multimedia*, 23, 4117-4130. DOI: 10.1109/TMM.2020.3037537

Lombardo, A., Shtrepi, L., & Astolfi, A. (2020, December). Applicability of multi-objective optimization in classroom acoustics design using analytical and geometrical acoustic models. *In Forum Acusticum* (pp. 1605-1608). DOI: 10.48465/fa.2020.0782

Savioja, L., & Svensson, U. P. (2015). Overview of geometrical room acoustic modeling techniques. *The Journal of the Acoustical Society of America*, 138(2), 708-730. https://doi.org/10.1121/1.4926438

Tan, Y., Fang, Y., Zhou, T., Gan, V. J., & Cheng, J. C. (2019). A BIM-supported 4D acoustics simulation approach to mitigating noise impact on maintenance workers on offshore oil and gas platforms. *Automation in Construction*, 100, 1-10. DOI: 10.1016/j.autcon.2018.12.019

Van der Harten, A. W., & Goodman, C. (2019). Novel applications of re-tooled open-source acoustic simulation algorithms. *The Journal of the Acoustical Society of America,* 146(4_Supplement), 2801-2801. https://doi.org/10.1121/1.5136701

Van der Harten, A. (2011, May). Customized room acoustics simulations using scripting interfaces. *In Proceedings of Meetings on Acoustics* (Vol. 12, No. 1). AIP Publishing. DOI: 10.1121/1.3587656

Xu, Q., & Wang, H. (2022). Sound field modeling method and key imaging technology of an ultrasonic phased array: A review. *Applied Sciences*, 12(16), 7962. https://doi.org/10.3390/app12167962

A PHYGITAL FORM-FINDING METHOD FOR BODY-SCALE SOUNDSCAPE INSTALLATION WITH FLEXIBLE MATERIALS

SIJIA GU[1], CHAO YAN*[2], TIANYI GAO[3], YUFEI WU[4], RUISHU WANG[5], GUOLIANG LIAN[6], RIXIN WANG[7], YUFEI PENG[8], FANGSHUO MO[9], and PHILIP F. YUAN[10]
[1,2,4,5,6,7,8,9,10]*College of Architecture and Urban Planning, Tongji University*
[3]*Shanghai Research Institute for Intelligent Autonomous Systems, Tongji University*
[1]*2230081@tongji.edu.cn, 0000-0003-4469-8925*
[2]*yanchao@tongji.edu.cn*
[3]*gerryvito@outlook.com, 0000-0001-8974-9234*
[4]*1851583@tongji.edu.cn, 0009-0004-3111-9968*
[5]*1753856@tongji.edu.cn, 0009-0001-7245-2896*
[6]*1952466@tongji.edu.cn, 0009-0009-4521-0443*
[7]*1953338@tongji.edu.cn, 0009-0005-6233-5484*
[8]*1952182@tongji.edu.cn , 0009-0003-9518-6337*
[9]*mfs@tongji.edu.cn*
[10]*philipyuan007@tongji.edu.cn, 0000-0002-2871-377X*

Abstract. Body-scale spatial installations with flexible materials effectively shape the soundscape in micro-environments. Design directives for such installations primarily rely on digital simulations and physical experiments. However, common acoustic software often fails to authentically replicate real experiential effects in compact installations. In contrast, physical experiments offer precise insights into the relationship between material forms and acoustic outcomes. Thus, a combined approach using virtual simulations and physical experiments across various scales is essential to determine the morphology and distribution of the flexible material. This research proposes a progressive design method for prototypes of soundscape installation, amalgamating phygital experimentation-based reasoning. It aims to validate the pivotal role of diverse experimental data and explore their integration within architectural design processes. The methodology involves software simulations, scale-down physical experiments, and full-scale tests, providing incremental design conclusions and facilitating the gradual realization of architectural design solutions. This approach was successfully applied to the "Natural Speaker" installation, presenting an innovative, shared experiential dataset and fostering fresh insights for soundscape design.

Keywords. Urban Soundscape, Phygital Experiment, Flexible Material, Form-finding Method, Architectural Design

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 415-424.

1. Introduction

1.1. BACKGROUND

The concept of soundscape was introduced by R. Murray Schafer in the 1970s (Schafer, 1993) to encompass the amalgamation of all sounds within a specific location or environment. Subsequently, Kang Jian proposed the concept of urban soundscapes in 2007, focusing on the formation of soundscapes within cities(Jian, 2007). A well-crafted soundscape can elevate spatial experiences and shape people's perceptions and emotional responses to space. Conversely, unfavorable soundscapes might have adverse effects on the health and psychological well-being of humans and other organisms(Fritschi et al., 2013). Consequently, numerous researchers have endeavored to shape favorable soundscapes through the design of sound propagation.

Materials possessing unique acoustic properties have garnered attention, with particular emphasis on flexible materials, such as textiles(Fridh et al., 2019). Besides, to enhance the acoustic characteristics of specific spaces, designers commonly employ body-scale spatial installations for acoustic design(Fusaro et al., 2018). These installations, characterized by their small size, adjustability, and mobility(Steele et al., 2019), can adapt to diverse urban environments, actively engaging in the shaping of urban soundscapes(Cerwén, 2016).

However, the design of body-scale spatial installations encounters challenges. Due to their compact size, body-scale installations often overlap with the spatial areas influenced by user behavior. It's necessary to consider the differences in acoustic performance caused by body positions. This complexity demands a more refined simulation and validation concerning soundscape. However, for the common acoustic design software, such as EASE (Enhanced Acoustic Simulator for Engineers), they often focus on larger spaces, making it challenging to accurately reflect the real experiential effects of body-scale spatial installations. Hence, in the process of optimizing the form and materials of body-scale spatial installations, a combined approach using different scales is necessary to determine the form and distribution of flexible materials, thereby enhancing the efficacy of body-scale spatial installations.

1.2. LITERATURE REVIEW

Numerous scholars have extensively researched the design methods for soundscape installation, with a primary focus on digital simulations. The integration of APBD (Acoustic Performance-Based Design) (Badino et al., 2020)allowed for the early identification of dependencies between design features and acoustic performance. Projects like the Smithsonian Institution Courtyard Enclosure and Manufacturing Parametric Acoustic Surfaces (Peters, 2010) demonstrate the potential of using sound simulation parameters. Genetic algorithms and APBD have been repeatedly employed in prior research projects for the form-finding of acoustic shells. The Aalborg Acoustic Pavilion (Foged et al., 2014), utilizing evolutionary computational strategies, achieves direct feedback between form, material and acoustic performance through comprehensive simulations. Similarly, projects like Soundforms (Foged et al., 2012) and RES (Resonant Strings Shell) (Mirra et al., 2018) exhibited researches akin to exploring diverse topological structures using acoustic predictions. In these cases, the optimization are usually based on Grasshopper plugins like Galapagos.

Contrary to digital simulations, physical experiments have also been incorporated into acoustic performance design considerations. An exemplary case is RES(Mirra et al., 2018), notable for its cyclic mechanism involving yearly fabrication experiments and on-site measurements during outdoor music festivals. This approach leverages machine learning for acoustic design, compensating for the lack of utilization of actual measurement data in the design process and validating the accuracy of the Aeolus plugin (Mirra et al., 2023) in early reflection predictions for musical scenarios.

However, for soundscape installations with flexible materials, numerical simulation or physical experimentation alone may not be applicable to the design of small soundscape installations with flexible materials. Research by Alvise Morandi and Carol Monticelli (Morandi & Monticelli, 2023) isolated flexible materials from secondary features of acoustic reflection installations, initially validating the effectiveness of membranes and their forms in noise control. Previous acoustic simulations have used an approximation in which the surface of the installation was considered to be a plane that reflects sound waves. The impact of this approximation on the simulation of flexible structures still needs to be verified Therefore, although previously physical experiments were mostly used for validation and localized testing in the later stages of design. However, it is particularly needed in the pre-design of flexible materials where the behavior in the acoustic environment is not yet known.

Moreover, prior practices largely focused on music performances and conference scenarios, abstracting most receivers as defined-height receiving planes (Foged et al., 2014) or defining spherical ranges(Morandi & Monticelli, 2023). The Aeolus plugin (Mirra et al., 2023) enabled rapid generation of 2DSPL diagrams. However, for the research related to body-scale soundscape installations, consideration of receiver bodily behavior becomes essential. Thus, the extension of acoustic prediction tools into 3D spaces warrants real-world acoustic environment integration.

In summary, while computational power enables traditional digital software to offer a basis for spatial form through multi-objective optimization, real physical experiments intuitively and effectively establish relationships between specific materials and acoustic properties. Combining digital simulation and physical experiments in the design process of body-scale soundscape installations would be a valuable research direction.

2. Proposal

This research proposes a progressive design methodology for body-scale soundscape installations based on the integration of digital and physical experiments. It aims to validate the guiding role of multi-source experimental data in the design process and explore its alignment with architectural design process. This method deconstructs the design process, identifying specific design problems at each stage and translating them into digital or physical experimental methods by controlling parameters. Designers use the experimental data generated at each stage as a basis for design conclusions, applying them in subsequent discussions, gradually derive architectural design solutions.

Specifically, the method consists of three cyclical stages:

(1) Digital simulation. Designers simulate the initial form of the installation by using acoustic software to obtain rough topological and shape information.

(2) Scale-down physical experiments. To address uncertainties in digital software, designers conduct physical experiments on various derived forms of the installation, acquiring more accurate data and scientifically formulating acoustic strategies.

(3) Full-scale physical experiments. The data derived from scale-down experiments cannot be directly related to the human experience. Therefore, full-scale physical experiments are included in the design methodology as an important validation.

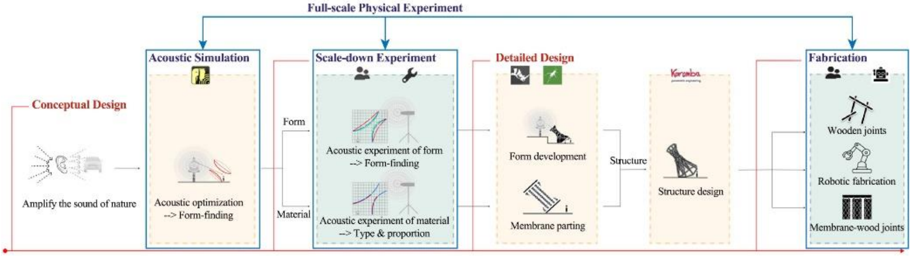


Figure 1. An illustration of the method proposed

3. Case study

This methodology was demonstrated in the design process of the "Natural Speaker", a body-scale soundscape installation. Within this project, the initial design process was divided into four stages: form generation, form elaboration, material selection, and material distribution. A combined approach involving software simulations and physical experiments was adopted to complete the design reasoning for each stage.

3.1. EXPERIMENTAL SETUP

This section predominantly focuses on the preparation of software, physical model and variable design for scale-down physical experiments. For the software, Grasshopper plugins for acoustic simulations were involved. Hardware requirements include: (1) components for physical models such as aluminum alloy frames, wooden boards, acrylic sheets, and hand tools. (2) equipment for measuring sound effects, including sound sources, a source controller, sound level meters, and tripods; (3) materials for testing, comprising fabrics and membranes. Materials and tools for full-scale physical experiments will not be discussed at this stage due to its complexity.

Figure 2. Equipment and materials for the experiment

In terms of physical model design, it consists the frame, the material under test, and the sound detection equipment. The frame was used to fix the material under test. A scaling ratio of 1:50 was used for the experiment, so an aluminum alloy frame with slots of 1300*900*1200mm dimensions was used. Meanwhile, wooden and acrylic boards were embedded in the frame to block the sound diffusion. The wooden and acrylic boards come with grids in order to fix the material in proportional segments. In order to focus on the experimental object, the profile morphology was extracted and used as the main object of study for the experiment. Thus, the tested materials are mainly the two parts that make up the profile. The material was fixed with the help of wooden strips placed between the wooden and acrylic boards.

The experiment comprises four types of variables: forms of section, types of the materials, proportions of the materials, and the location of acoustic equipment. Based on the digital simulation results. Nine profile forms are proposed. To simplify the experimental steps, forms A, B, and C are selected as representatives. The composition of the materials encompasses three types: pure fabric, pure membrane, and a composite of fabric and membrane material. There are three variations in the assembly position for both upper and lower materials, resulting in nine assembly methods denoted as A-1, B-1, C-1, A-2, B-2, C-2, A-3, B-3, and C-3. For the location of acoustic equipment, when testing the effects of bird sounds on people, the sound source was placed on the axis of the profile, simulating the location of the birds in the design. The sound level meter was placed at the human ear location. When testing the effect of human sound on the bird, the positions of the sound source and the sound level meter were switched.

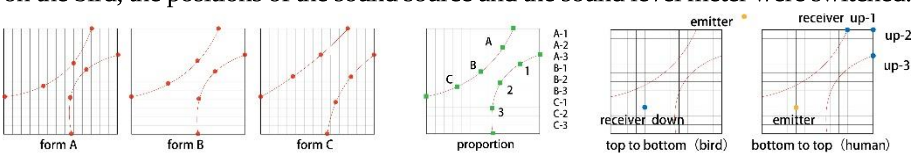


Figure 3. Definition of variables for the experiment

3.2. EXPERIMENTAL PROCESS

The experiment has two goals. On the one hand, the experiment can instruct on how to amplify the bird's voice, so that people can hear the birdsong more clearly. On the other hand, the experiment can instruct how to reduce the human voice so that the human voice does not disturb the bird's normal stay as much as possible. In terms of experimental arrangement, initially, multi-objective optimization of the overall form of the installation was conducted using acoustic-related digital plugins. Subsequently, the material distribution of the installation was obtained through a scale-down physical experiment. Lastly, an evaluation was conducted through full-scale experiments to validate the accuracy of the aforementioned stages.

3.2.1. Digital Simulation experiment

In the software simulation phase, the study utilized the Pachyderm Acoustic plugin in Rhino-Grasshopper for acoustic simulations. To replicate real-world scenarios within the design environment, the simulated sound source was positioned at the height of tree. Before the simulation process, an initial installation form was designated. Through continuous feedback from acoustic simulations of sound reflections, parameters controlling the form would be iteratively adjusted. Eventually, an optimal solution for parameters was obtained. It implied a best form of surface for sound propagation.

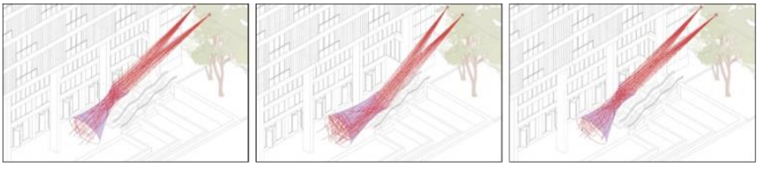


Figure 4. The process of digital simulation of acoustic performance

3.2.2. Scale-down physical experiment

Form	Material	Dividing Location	Receivers' Locations	Stage
/	none	/	up-1 up-2 up-3 bottom	stage 0
form A form B form C	membranes textile+membranes	none	bottom	stage 1
(3 types)	(2 types)	(1 type)	(1 type)	= (6 rounds)
form B	membrane(up)+membrane(bottom left)+membrane(bottom right) membrane(up)+textile(bottom left)+membrane(bottom right) membrane(up)+textile&membrane(bottom left)+membrane(bottom right) membrane(up)+membrane(bottom left)+textile(bottom right) membrane(up)+textile(bottom left)+textile(bottom right) membrane(up)+textile&membran(bottom left)+textile(bottom right) membrane(up)+membrane(bottom left)+textile&membrane(bottom right) membrane(up)+textile(bottom left)+textile&membrane(bottom right) membrane(up)+textile&membran(bottom left)+textile(bottom right)	× B-2	up-1 up-2 up-3	stage 2
(1 type)	(9 types)	(1 type)	(3 types)	=(27 rounds)
form B	membrane(up)+textile&membran(bottom left)+textile(bottom right)	A-1 A-2 A-3 B-1 B-2 B-3 C-1 C-2 C-3	up-1 up-2 up-3 bottom	stage 3
(1 type)	(1 type)	(9 types)	(4 types)	=(36 rounds)

Figure 5. Experimental process of the scale-down physical experiment

The previous step of digital simulation led to a theoretically optimal form. However, for body-scale installations, it demands more precise and reliable data guidance for sound effects. Thus, scale-down physical experiments were conducted. This process contained three progressive steps: (1) A best form would be selected according to the result of acoustic pressure testing with different forms. (2) The type of the material of the installation would be determined through acoustic pressure testing results. (3) The distribution of the material would be determined through another set of acoustic pressure testing experiments.

At the beginning, the experiment measured sound propagation in the absence of materials. Subsequently, the reasoning experiments for the form and material started. The aim of the experiments was to derive an optimal shape to maximize birdsongs reaching the human ear. The experimental tested the differences in acoustic performance at the human ear position with the 3 different forms of A, B, and C under identical material conditions. The materials used in experiments were unified as membrane or a composite of membrane and fabric. In the experiment, materials were set to predetermined lengths, fixed onto the aluminum frame. The sound source was mounted on a tripod fixed in the location of the axis of the section of the installation. Controlled by a sound source controller, it emitted pink noise. The sound level meter was positioned beneath the cross-section, simulating the location of the human ear.

Comparing data from two sets of experiments, the sound pressure level exhibited the best performance in LAF at Form B. Consequently, form B was selected as the morphology of the installation.

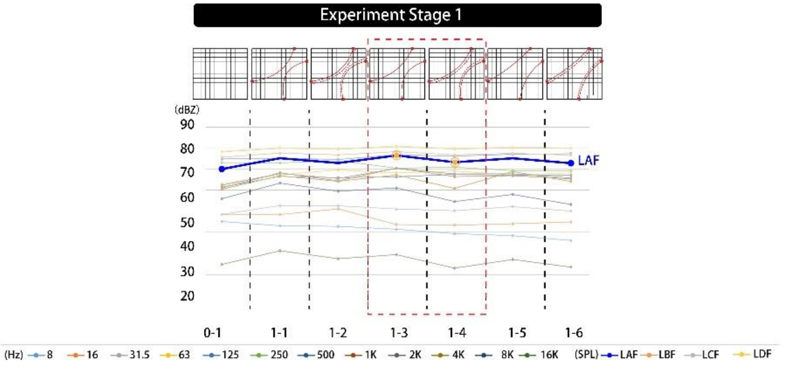


Figure 6. Results of the stage 1 of the scale-down physical experiment

After the morphology was determined, the experiment attempted to deduce the type of the material of the installation. The purpose of this set of experiments was to test under which material composition human voice could be maximally absorbed by the installation, thus not disturbing birds. In the experiments, the form remained a fixed parameter of form B, while materials were successively replaced with different types. Considering that fabrics do not have the ability to reflect sound in a way that would adversely affect the transmission of bird sound, the material above the two components of the cross-section was considered to be membranes. Conversely, the materials in the lower parts of the cross-section were individually replaced with three types (pure fabric, pure membrane, and a composite of fabric and membrane), yielding a total of 9 experimental scenarios. The sound source, simulating human voice, was fixed close to the top of the installation near where a user's mouth would be. The sound level meter was placed at 3 different positions at the top of the axis of the cross-section, mimicking the sound perceived by birds.

Comparing the results of the experiments, it was found that when the upper section of the installation was membrane and the lower section was a composite of fabric and membrane, the propagation of human voice was minimized for birds. Consequently, this material type was chosen for the installation.

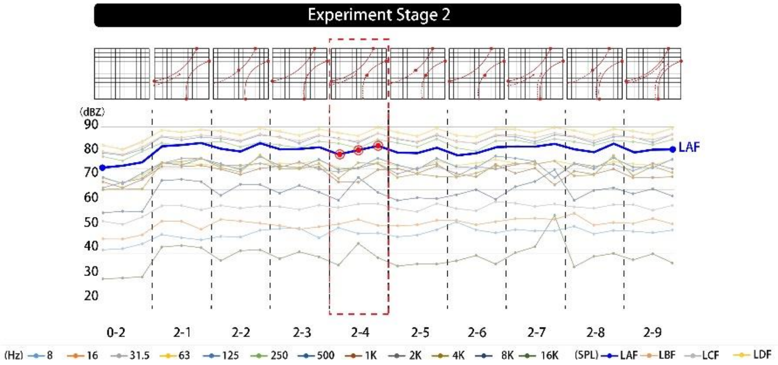


Figure 7. Results of the stage 2 of the scale-down physical experiment

In order to derive a more specific distribution of the material, the experiment conducted sound pressure measurements by varying the proportions of the material combinations based on the aforementioned materials. The purpose of this experiment was to obtain a material distribution ratio that would optimize the composite indicator of bird and human sounds. The experiment selected four equally spaced points on the cross-sections, totaling 9 sets of scenarios. The sound source was placed both up and down, measuring the acoustic performance of both birdsongs and human voices.

In the comprehensive comparison, the optimal effect in sound propagation was observed when the material junction points of the upper and lower cross-sections were respectively located at positions A and 2. Consequently, this proportion of the material was chosen for the material design of the installation.

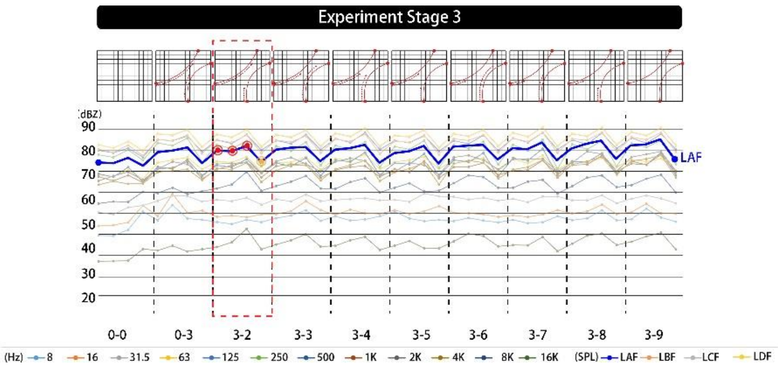


Figure 8. Results of the stage 3 of the scale-down physical experiment

3.2.3. Full-scale physical experiment

Although scale-down experiments provide highly intuitive data regarding sound propagation effects, they were insufficient to form reliable judgments regarding human experience. Therefore, a full-scale physical model becomes crucial as it allowed for the evaluation of whether previous experimental data and design strategies were truly effective and meaningful in terms of subjective experiences.

A full-scale experiment in a spacious outdoor location in Shanghai. The installation was positioned facing a tree frequently inhabited by birds, serving as the source of sound emission. Constructed primarily using pure membrane material as the enclosure structure and wood as the sup-port structure, the installation was divided into 50*50*50cm spatial grids during the experiment. Measurements and assessments of sound pressure at grid points within the spatial framework and human perception of bird sounds were con-ducted. Ultimately, it was observed that locations where bird sounds measured higher sound pressure levels also elicited stronger human perceptions of bird sounds (higher scoring in subjective assessments). This observation indicates that the outcomes guided by previous experiments have good guiding significance for real experiential perceptions.

Figure 9. Images of the process of full-scale physical experiment

4. Discussion

Despite the effectiveness of this general method, there remains improvement in the actual installation design process. For instance, in the scale-down experiments, cross-section was used to simplify design reasoning. However, in reality, the three-dimensional state of the physical installation would provide more data information than the two-dimensional profiles, enabling guidance for both the design of the listening position for humans and the resting location for birds.

5. Conclusion

The utilization of body-scale installation with flexible materials is considered an effective method for soundscapes. However, due to the higher demands of body-scale spatial installations for embodied experiences, digital acoustic simulation often fails to meet the data requirements for design.

This research explores the soundscape installation design process from an architectural perspective and introduces a progressive design method that combines experimentation with design. Through a three-stage process involving digital simulation, scale-down physical experiments, and full-scale physical experiments, the form and material design of the installation is derived. In this process, multisource sound data from both software simulation and physical experiments are integrated into the installation design process and serve as crucial foundations for design advancement.

This method is validated in the design process of the "Natural Speaker" soundscape installation. The results demonstrate the effectiveness and generalizability of the proposed method. Additionally, the study has generated experiential data that will be shared on GitHub. Furthermore, the research contemplates utilizing the experimental data of acoustic performance to establish a general correspondence among form, materials, and soundscape. This initiative aims to foster new inspiration and contemplation in soundscape architectural design for future reuse, potentially opening new avenues of thought in this field.

6. Acknowledgement

This study was funded by Tongji Architectural Design (Group) Co, Ltd. project (numbered 2022J-JB06 in the category of "Jie Bang Gua Shuai").

References

Badino, E., Shtrepi, L., & Astolfi, A. (2020). Acoustic Performance-Based Design: A Brief Overview of the Opportunities and Limits in Current Practice. Acoustics, 2(2), 246–278. https://doi.org/10.3390/acoustics2020016

Cerwén, G. (2016). Urban soundscapes: A quasi-experiment in landscape architecture. Landscape Research, 41(5), 481–494. https://doi.org/10.1080/01426397.2015.1117062

Foged, I. W., Pasold, A., & Jensen, M. B. (2014). Evolution of an instrumental architecture. Vol. 2 of Proc. 32nd eCAADe Conf, 365–372.

Foged, I. W., Pasold, A., Jensen, M. B., & Poulsen, E. S. (2012). Acoustic environments: Applying evolutionary algorithms for sound based morphogenesis. eCAADe 2012 Conference: Digital Physicality| Physical Digitality, 347–353.

Fridh, K., Zetterblom, M., & Femenias, P. (2019). Sound absorbing textile surfaces in the urban landscape-collaborative research in textile and architectural design. Proceedings of the e Associ-ation of Fashion & Textile Courses (FTC) Conference Futurescan 4: Valuing Practice.

Fritschi, L., Brown, L. A., Kim, R., Schwela, D., & Kephalopoulos, S. (2013). Burden of disease from environmental noise. Quantification of healthy life years lost in Europe. Geneva: World Health Organization 2011.

Fusaro, G., D’Alessandro, F., Baldinelli, G., & Kang, J. (2018). Design of urban furniture to enhance the soundscape: A case study. Building Acoustics, 25(1), 61–75. https://doi.org/10.1177/1351010X18757413

Jian, K. (2007). Urban soundscape. Journal of South China University of Technology, 35, 11–16.

Mirra, G., Mack, M., & Pugnale, A. (2023). Aeolus: A Grasshopper plugin for the interactive design and optimisation of acoustic shells. Proceedings of IASS Annual Symposia, 2023(6), 1–13.

Mirra, G., Pignatelli, E., & Di Rosario, S. (2018). An automated design methodology for acous-tic shells in outdoor concerts. Proceedings of the Euronoise, Crete, Greece, 27–31.

Morandi, A., & Monticelli, C. (2023). Textile Membranes Reused as a Tool for Noise Control. Buildings, 13(9), 2134. https://doi.org/10.3390/buildings13092134

Peters, B. (2010). Acoustic Performance as a Design Driver: Sound Simulation and Parametric Modeling Using SmartGeometry. International Journal of Architectural Computing, 8(3), 337–358. https://doi.org/10.1260/1478-0771.8.3.337

Schafer, R. M. (1993). The soundscape: Our sonic environment and the tuning of the world. Simon and Schuster.

Steele, D., Legast, É., Trudeau, C., Fraisse, V., & Guastavino, C. (2019). Sounds in the city: Improving the soundscape of a public square through sound art. Proceedings of the ISCV, 26(8).

AN APPROACH TO IMPROVE INDIVIDUAL THERMAL COMFORT BASED ON MOBILE MEASUREMENT OF BIOLOGICAL REACTION TO PROXIMATE THERMAL ENVIRONMENT

SHUN KIMURA[1] and YASUSHI IKEDA[2]
[1]*Keio University Graduate School of Media and Governance, Japan*
[2]*Tokyo University, Japan*
[1]*shun.k@keio.jp 0009-0008-9142-9342*
[2]*yasushi@arch1.t.u-tokyo.ac.jp 0000-0002-2016-4083*

Abstract. Thermal comfort is a critical research field because it affects people's health, productivity, and well-being in an architectural space. Numerous previous studies have attempted to find thermal environment conditions where average people feel comfortable. Recently, researchers have focused more on estimating and improving thermal comfort on an individual level to improve the accuracy of previous methods, leveraging recent developments in sensing technologies. However, there are still some remaining issues, such as limited scalability due to the use of in-place sensors and the assumption of a uniform thermal environment, making it unable to recognize locational differences. This study is fundamental research that proposes a human-based approach based on a mobile measurement system to overcome these issues. Firstly, the concept and the benefits are clarified and then, the methodologies and results of two fundamental experiments are explained. One was a field experiment to investigate its validity in estimating individual thermal comfort in real-life situations. The other one was to consider and examine ways to improve individual thermal comfort by the human-based approach. The results suggest that the human-based approach can be beneficial in a way never seen although some issues need to be addressed in future research.

Keywords. Individual Thermal Comfort, Mobile Sensing, HCI, Human-Centric Modeling, Interactive Environment

1. Introduction

Thermal comfort is the condition of mind that expresses satisfaction with the thermal environment. Previous research on thermal comfort has analyzed the environmental factors causing comfort or discomfort in humans and how humans perceive it. Conventional research attempted to tackle the question by empirical and physiological approaches and the most recognized achievements are PMV and SET*, which are accepted by international standards, ISSO and ASHRAE, to standardize architecture designs and HVAC control strategies to create thermal environments where average people feel comfortable. Recent research has shifted a focus to thermal comfort of

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 425-434. © 2024 and published by the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA), Hong Kong.

individuals, leveraging advancements in sensing technologies to enhance the precision of conventional average methods. Nevertheless, current research on individual thermal comfort also has some limitations. Firstly, it has limited scalability because many studies employed in-place sensors to monitor environmental thermal factors, which makes the estimation in another place impossible. Secondly, as a uniform thermal environment is often assumed, locational variations are neglected. Within the field of environmental engineering, there is growing interest in leveraging thermal environment unevenness to enhance occupants' comfort, necessitating their integration into research on individual thermal comfort. Thirdly, the only purpose of the research is the estimation of thermal comfort of individuals, and the network of individuals is often overlooked even though it could be useful to improve overall individual thermal comfort.

To address these issues, this paper proposes a novel human-based approach and introduces methodologies and results of our several fundamental experiments.

2. Previous Paper

The main characteristic of the research on individual thermal comfort is to attempt to estimate thermal comfort of individuals by monitoring the data of biological reactions of the human body. Initial studies about individual thermal comfort discovered the correlation between thermal comfort of the subjects and biological reactions, such as skin temperature, electroencephalography (EEG) and heart rate variability (HRV), through an experiment in a climate chamber (Wang et al., 2007)(Yao et al., 2009). Recently, with rapidly developing sensing technologies, researchers have focused more on making estimations under more diverse conditions and creating interactive environments with installed actuators to improve subjects' thermal comfort. One of the typical approaches to utilize developed sensing technologies is using infrared sensors to measure facial skin temperature to estimate individual thermal comfort (Ghahramani et al., 2016). Another study proposed a new approach using occupants' cooling and heating behavior to derive an individually tailored thermal comfort estimation model to deal with the subjectivity and the high cost of collecting user feedback (Kim et al., 2018). Lu et al. (2019) proposed a design of an interactive environment, using personal cooling devices to improve thermal comfort of each subject based on the estimation derived from infrared sensor data. These new studies indicated that the relationship between architectural space and users has become more dynamic and interactive with new technologies. Still, issues related to locational limitations and the assumption of uniformity in thermal environments have yet to be addressed.

Looking at other research fields, another approach taking advantage of the non-uniformity of thermal environments to improve occupants' thermal comfort is proposed in environmental engineering. It is suggested that a non-uniform thermal environment intentionally created by non-uniform ventilation modes could reduce Predicted Percentage of Dissatisfied (PPD) by 33% (Fan et al. 2022). Also, Taniguchi et al., (2023) illustrated that an appropriate combination of indoor thermal environment distribution and individual thermal preferences could improve occupants' thermal comfort by CFD and multi-agent simulation supported by a subject experiment. In healthcare, mobile sensing is becoming increasingly important in delivering personalized healthcare services based on monitored biological and behavioral data

(Baig et al., 2014).

This paper proposes a novel human-based approach to which the assumption of non-uniformity of thermal environments and mobile measurement are applied from the other research fields for the estimation and the improvement of individual thermal comfort.

3. Motivation

The motivation behind this study is to improve people's individual thermal comfort by introducing a human-based approach to solve the challenges mentioned in the preceding sections. In this study, a human-based approach is defined as one that is mobile and independent of any specific location or attached technologies. Employing a human-based methodology offers several advantages due to its independence. The first is the capacity to collect data at any location and time, even in real-life scenarios. The second is its ability to recognize locational differences in environmental factors, such as air temperature and relative humidity. The third is that data can be measured in the most efficient way because there is no need to spend time and money to introduce sensors to all places of interest. The last and foremost one is that a human-based approach inherently optimizes sensor distribution for humans, automatically placing necessary sensors in necessary locations. Fig 1 depicts a conceptual diagram of a society embracing a human-based approach to enhance individuals' thermal comfort.

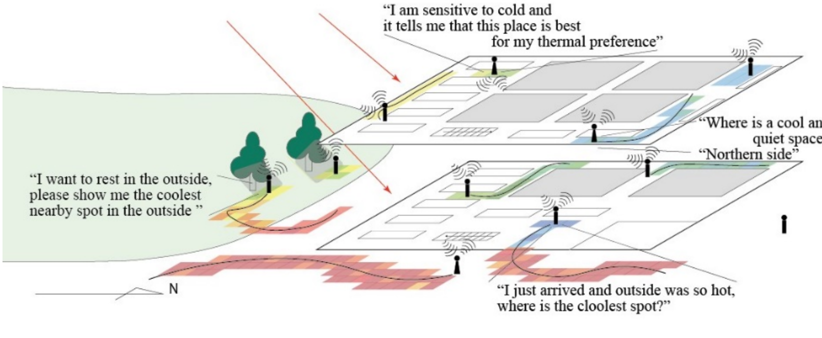


Fig 1. Concept of Human-based Approach for Thermal Comfort Estimation and Improvement

To validate the effectiveness of the human-based approach, our study comprises two fundamental experiments. The first one is a field experiment demonstrating that the human-based approach enables the estimation of individual thermal comfort at any location and time. This experiment was complemented by the development of a mobile thermal comfort measurement system that continuously measures environmental and biological data. At the end of this step, an estimation model to predict the author's individual thermal comfort was constructed. Based on the estimation model, a use case experiment was conducted as a second experiment to show that there is a high possibility that individuals can improve their own thermal comfort with minimal cost and labor by matching non-uniform thermal environment distribution and individual thermal preferences. These experiments were conducted only by the author since this study, unlike previous studies, focused on measuring much data about intrapersonal differences in biological conditions under given constraints of time and money. To reduce subjectivity, our study was divided into two experiments so that the subject's feedback did not affect the examination of the second experiments, and some additional measurements are applied in each step, which will be described in following sections.

4. Experiment

4.1. PURPOSE

The purpose of this experiment is to investigate the validity of a human-based approach in estimating individual thermal comfort in real-life and to create an estimation model for the second experiments. In the first stage, we developed a mobile thermal comfort measurement system that measures data about proximate thermal environment and the corresponding biological reactions. Subsequently, a field experiment was conducted to investigate the validity in a real-life situation.

4.2. DEVELOPMENT OF MOBILE THERMAL COMFORT MEASUREMENT SYSTEM

The first stage of the experiment is the development of a mobile thermal comfort measurement system. Before the hardware design, independent and dependent variables to measure are determined. As independent variables, we chose air temperature and air relative humidity as environmental factors and skin temperature of the wrist and chest, skin peripheral relative humidity of the wrist and chest, heart rate and clothing level as human factors to estimate individual thermal comfort. The dependent variable chosen was the 7-point thermal sensation scale accepted by ASHRAE (-3: cold - +3: hot) and was named as Subjective Thermal Comfort Vote (STCV) in this study. The requirement for the human-based measurement system was to be able to monitor these data robustly and precisely in a mobile way.

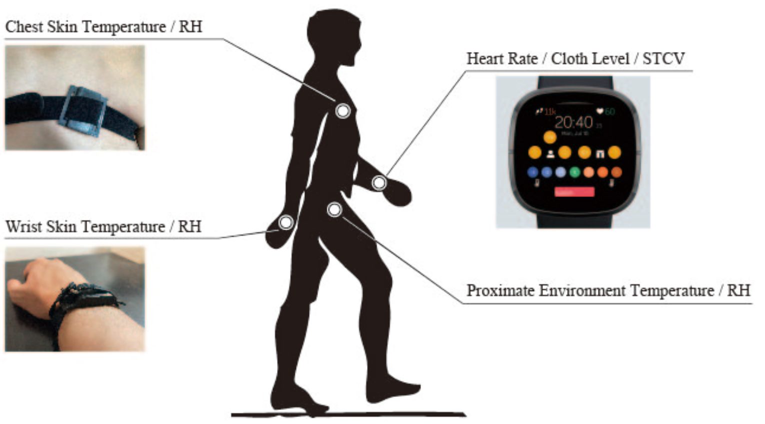


Fig 2. Mobile thermal comfort measurement system

To monitor temperature and relative humidity, iButton DS1923 was chosen. iButton DS1923 is a button-type data logger manufactured by Maxim Integrated for temperature and relative humidity. Its advantage is its ability to log data with higher resolution and sampling rate compared to other sensors within the same price range. The iButtons were affixed to each location, using holders fabricated by a 3D printer, with simple belts to secure them (Figure 2.). Data collected by the iButtons was stored in the sensor and could be retrieved to a laptop using the 1-wire protocol. This could make measured values invisible from the subject to deal with the subjectivity issue.

Another device to monitor the other variables was Fitbit Sense, one of the smartwatches manufactured and sold by Fitbit. The reason why it was chosen was because it is so customizable in terms of measuring data and interface design by utilizing Fitbit SDK that it provides flexibility to this study. For this experiment, the interface was designed like Figure 2. Data collected by Fitbit Sense was transferred and saved in a database in AWS.

4.3. METHODOLOGY

The data had been collected in real life by the author, a 25-year-old man, for 35 days from the second half of May (late spring) to the first half of August (summer). A sampling rate is set to 1 minute and a reminder is set every 10 minutes for a user to submit clothing level and STCV. The data collected less than 10 minutes after the start of the experiment was excluded from the dataset. The number of data points for each clothing level and each STCV is shown in Table 1 and Table 2.

Shorts & Short Sleeve T-shirt	Pants & Short Sleeve T-shirt	Pants & Long Sleeve T-shirt	sum
7509	3554	513	11576

Table 1. The number of data points for each clothing level

-1	0	1	2	3	sum
413	9657	1341	137	28	11576

Table 2. The number of data points for each STCV

4.4. RESULT

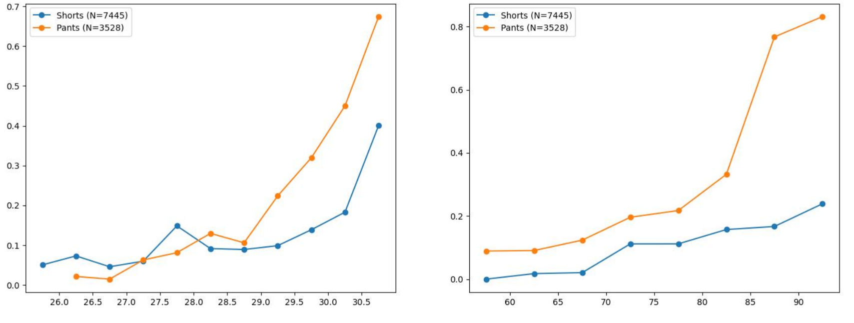


Figure 3. The probability of STCV≥1 for air temperature (Left) and 10-minute mean of HR (Right)

As a first step, the relationship between measured variables and STCV was analyzed. As expected, a strong relationship with STCV was confirmed in air temperature, air humidity and 10-minute mean of heart rate. Figure 3 shows two plots of probabilities of STCV≥1 for each data range of air temperature and 10-minute mean of heart rate when wearing pants and shorts. According to the plots, there is an over 40% chance of feeling warm/hot when air temperature is above 30.0°C and the probability is over 70% when 10-minute mean of heart rate surpasses 85%.

Figure 4 presents the same plots for wrist skin temperature and wrist peripheral skin relative humidity. For wrist skin temperature, the probability is higher when the skin temperature exceeds 35.0°C as expected, while it is also relatively higher around 34.0°C. The additional analysis shows that when feeling warm/hot (STCV≥1), the mean heart rate is about ten higher when the temperature is below 34.5°C than when the temperature is above it. Thus, it is presumed that feeling warm/hot with a low wrist temperature is mainly because of increased metabolism and, in contrast, feeling warm/hot with a high wrist temperature is mainly due to heat exchange between the outside and the body. The right side of Figure 4 shows that the probability is higher when the wrist peripheral relative humidity is around 80% although a simple positive correlation was expected. In real-life scenarios, it is normal for a person to take some measurements when feeling hot/cold, so the condition becomes different when the relative humidity reaches 90%. Thus, it is crucial to detect the initial increase in the relative humidity to identify when the subject starts to feel hot. In fact, Figure 5 shows the strong correlation between the 5-minute lag of the wrist skin peripheral relative humidity and STCV. These results suggest that the mobile human-comfort estimation system can identify a dynamic relationship between thermal comfort and some biological reactions.

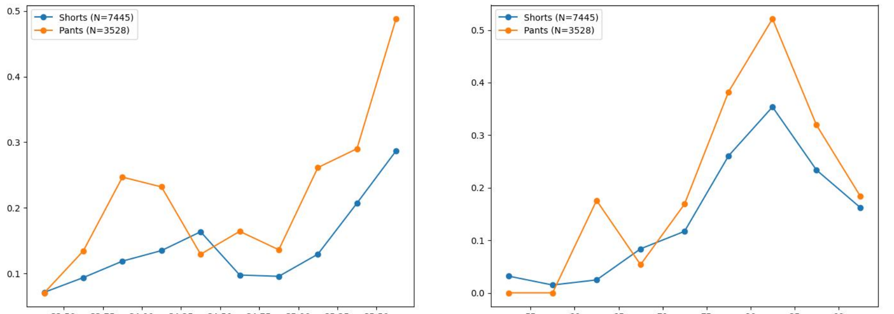


Figure 4. The probability of STCV≥1 for wrist skin temperature (Left) and peripheral relative humidity (Right)

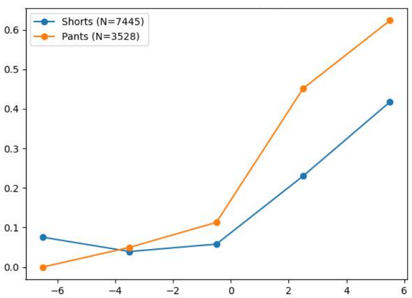


Figure 5. The probability of STCV≥1 for 5-minute lag of wrist skin peripheral relative humidity

Since the correlation between measured variables and thermal comfort expressed by STCV was confirmed, a Machine Learning (ML) model was constructed based on the dataset. SVC was chosen as a classifier model and all eight variables were used as independent variables. To split the dataset into train data and test data, the dataset is divided into sub-datasets based on a combination of clothing level and STCV, and the first 80% of the chronologically ordered sub-dataset is extracted for train data and, the rest is reserved for test data. Due to the unbalanced dataset, a f1 score is selected as an evaluation function, treating STCV=1,2,3 as the same label in the evaluation step due to the limited number of data points when STCV=2,3. The hyperparameters chosen by grid search were C=0.5 and gamma=0.1. F1, recall, precision, and accuracy for test data are 0.53, 0.51, 0.77, and 0.81, respectively. This ML model was used in the next experiment, Use Case, to detect discomfort of the user by calculating Predicted STCV, the average of STCV labels weighted by probabilities that one data point is estimated to belong to each label calculated by the Platt scaling method.

5. Use Case

5.1. PURPOSE

The purpose of a use case experiment is to conduct a demonstration experiment assuming one of the specific situations in which the invented system will be used to investigate whether and how it can help individuals improve their own thermal comfort when it is used by multiple users. In this study, the application of the system in a library in our campus was selected as a specific situation. It was because improving occupants' thermal comfort during desk work in an environment with unified air conditioning control is one of the most expected solutions, considering the advantages of human-based approach that can match thermal environment distribution and thermal preferences of occupants to improve their thermal comfort.

5.2. METHODOLOGY

Fig 6. Library

The experiment site was a main space in a 3-floor library building (32m*80m) in our campus (Figure 6). The biggest problem for this use case was time and money to prepare the invented system for multiple people and collect data for each person. In this use case, we pseudo-solved the problem by the author changing his location in a short period of time (3min) so that we could pseudo-measure data for around 20 individuals with thermal comfort preferences similar to the author in 1hour. The threshold of 3 minutes was chosen because it did not cause a measuring lag and allowed measurement for a sufficient number of locations. As GPS is ineffective indoors, indoor location estimation was made by pedestrian dead reckoning (PDR). PDR is one of the most famous and simplest ways to estimate a pedestrian's indoor locations based on an acceleration sensor to detect steps and a geomagnetic sensor to determine walking direction. WitMotion WT901BLECL BLE5.0 was selected as a device for PDR. Due to the limitation of PDR, data collection was conducted one floor by one floor, and four rounds of measurement were conducted in a single day, with one round consisting of measuring data for around 18 locations on 3 floors over 1 hour. This setting allowed us to compare data from 18 locations in the site in the analysis step under our hypothesis that occupants change their location in 1 hour on average. Measured variables was the subject location and all the variables in the experiment but STCV. Data had been collected for 9 days from August 17 to September 12. Figure 7 is a plot that shows the location of the subject and measured air temperature for each place in one day.

In the analysis, the room for the improvement for the subject's thermal comfort was investigated while the subject's thermal comfort was calculated by Predicted STCV to reduce subjectivity.

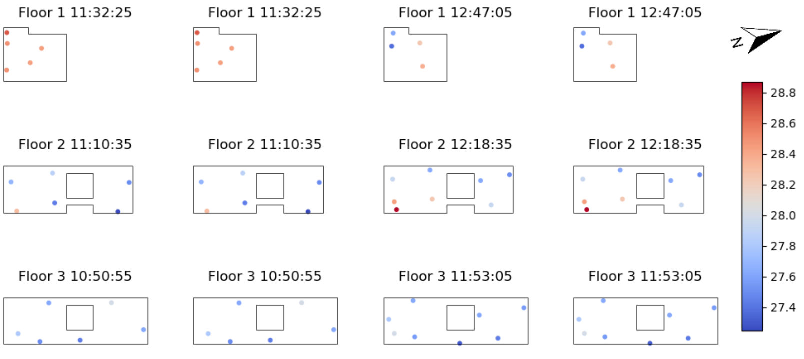


Fig 7. An example plot for the subject location and air temperature on 9/11

5.3. RESULT

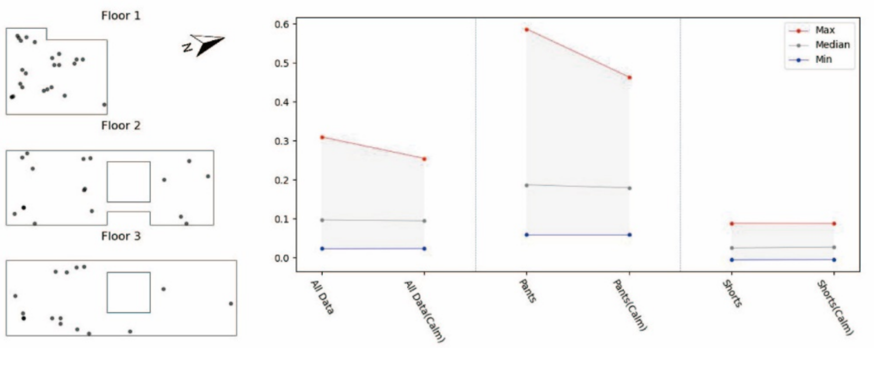


Fig 8. (Left) Subject location when the subject is predicted to feel hot
(Right) Mean of the maximum, the median, the minimum of Predicted STCV in one round

The left side of Figure 8 plots the subject's location when the Predicted STCV calculated by the ML model is higher than 0.5. The threshold of 0.5 was chosen as a level beyond which productivity in desk work may decline. According to the plot, there is always a chance that the subject feels hot anywhere in the library. It means that thermal condition changes from time to time even in the same place, and also from place to place and so does body condition, which always affects thermal comfort. Consequently, it is imperative to measure locational and chronological differences in thermal conditions and the corresponding biological reactions to estimate individual thermal comfort accurately. The invented system is useful in this way because it can frequently measure both data across diverse locations without presuming uniform thermal environments.

To investigate whether the invented system is helpful for individuals to improve their own thermal comfort by suggesting a thermally optimal place for each one, the values of Predicted STCV were compared in one round to calculate potential improvement in the subject's individual thermal comfort. To make a comparison, the mean of the maximum, the median, and the minimum of Predicted STCV in one round are calculated for each clothing level. These values are depicted on the right side of Figure 8, representing a simulated calculation of the potential improvement in thermal comfort if the subject were to relocate to a thermally optimal place. For each pair of values within the same category, the left one is calculated from all the data belonging to the category, and the right one is calculated from the data when a heart rate of the subject<80 (Calm). The plot indicates an average potential improvement in thermal comfort ranging from 0.2 to 0.3 at the most. Specifically, when wearing pants, the improvement is approximately 0.5, while for shorts, it hovers around 0.1. This result suggests that there is a high possibility that the human-based approach can help individuals improve their thermal comfort with minimum dependence on HVAC machines.

6. Conclusion and Discussion

In this study, the concept and the benefit of the human-based approach are proposed. Our fundamental studies showed that the invented mobile thermal comfort measurement system can recognize locational and chronological differences in thermal environment and human biological reactions to help estimate individual thermal comfort. Furthermore, one way to improve people's thermal comfort by the human-based approach was proposed and investigated, then it was proven highly probable from the use case that the mobile measurement system can help individuals improve their own thermal comfort with minimum cost and labor even in a space with unified air conditioning control in a situation where the human-based approach is widely accepted.

This study is still ongoing research and can be developed in the following ways.

- Important variables that affect individual thermal comfort, such as radiation and wind, are missing from this study. Mobile methods for measuring these data should be explored.
- The number of subjects is quite few even though the main purpose of this study is to propose the concept and prove an advantage of the human-based approach, not

the comparison in ways to feel thermal comfort among different people. Still, more subjects are helpful in proving its validity.

- Standardization of data collection methods and evaluation models is imperative to facilitate comparisons across different studies on individual thermal comfort.

Despite the limitations mentioned above, this study is valuable in proposing the concept of the human-based approach that enables an individually customized thermal comfort estimation model and improvements in real-life situations. The human-based approach has the potential to solve the remaining issues in previous research and lead to better people's thermal comfort due to its advantages. Furthermore, widespread adoption of the human-based approach could potentially reduce reliance on HVAC systems, leading to more energy-efficient improvements in individual thermal comfort. In a broader context, because of these advantages, the human-based approach is effective not only for thermal comfort but also for many other factors that influence the quality of the experience in architectural space. We hope that the human-based approach will be one way to improve the design quality of the increasingly dynamic relationship between buildings and humans.

References

Baig, M. M., GholamHosseini, H., & Connolly, M. (2014). Mobile healthcare applications: system design review, critical issues and challenges. *Australasian Physical & Engineering Sciences in Medicine, 38*(1), 23–38. https://doi.org/10.1007/s13246-014-0315-4

Fan, M., Fu, Z., Wang, J., Wang, Z., Suo, H., Kong, X., & Li, H. (2022). A review of different ventilation modes on thermal comfort, air quality and virus spread control. *Building and Environment, 212*, 108831. https://doi.org/10.1016/j.buildenv.2022.108831

Ghahramani, A., Castro, G., Becerik-Gerber, B., & Yu, X. (2016). Infrared thermography of human face for monitoring thermoregulation performance and estimating personal thermal comfort. *Building and Environment, 109*, 1–11. https://doi.org/10.1016/j.buildenv.2016.09.005

Kim, J., Zhou, Y., Schiavon, S., Raftery, P., & Brager, G. (2018). Personal comfort models: Predicting individuals' thermal preference using occupant heating and cooling behavior and machine learning. *Building and Environment, 129*, 96–106. https://doi.org/10.1016/j.buildenv.2017.12.011

Lu, S., Wang, S., Cochran Hameen, E., Shi, J., & Zou, Y. (2019). Comfort-Based Integrative HVAC System with Non-Intrusive Sensing in Office Buildings. *In 24th International Conference on Computer-Aided Architectural Design Research in Asia: Intelligent and Informed, CAADRIA 2019* (pp. 785-794). The Association for Computer-Aided Architectural Design Research in Asia (CAADRIA). https://doi.org/10.52842/conf.caadria.2019.1.785

Taniguchi, K., Ishiura, K., Miyata, S., & Akashi, Y. (2023). The effect of seating movement based on indoor thermal environmental distribution on overall occupant thermal satisfaction. *Nihon Kenchiku Gakkai Kankyokei Ronbunshu, 88*(808), 511–520. https://doi.org/10.3130/aije.88.511

Wang, D., Zhang, H., Arens, E., & Huizenga, C. (2007). Observations of upper-extremity skin temperature and corresponding overall-body thermal sensations and comfort. *Building and Environment, 42*(12), 3933–3943. https://doi.org/10.1016/j.buildenv.2006.06.035

Yao, Y., Lian, Z., Liu, W., Jiang, C., Liu, Y., & Lu, H. (2009). Heart rate variation and electroencephalograph - the potential physiological factors for thermal comfort study. *Indoor Air, 19*(2), 93–101. https://doi.org/10.1111/j.1600-0668.2008.00565.x

BRIDGING EXTERIOR AND INTERIOR CLIMATE

Interdisciplinary Design of a Dual-Functional Adaptive Kinetic Façade

MOON-YOUNG JEONG[1], REINER RÖSSLE[2], MICHAEL BÖHM[3], MARIA MATHEOU[4], OLIVER SAWODNY[5]
[1,4] *ILEK (Institute for Lightweight Structures and Conceptual Design), University of Stuttgart*
[2,3,5] *ISYS (Institute for System Dynamics), University of Stuttgart*
[1,4] *{moon-young.jeong\maria.matheou}@ilek.uni-stuttgart.de*
[2,3,5] *{reiner.roessle\michael.boehm\oliver.sawodny}@isys.uni-stuttgart.de, 0009-0005-1862-2646*

Abstract. Building facades serve as the interface between their interior and exterior environments. With climate change necessitating adaptive solutions, the development of adaptive facades for a sustainable operational energy cycle for buildings has been prompted. This study delves into the design and prototyping of KineticSKIN, an Adaptive Kinetic Façade (AKF) system. The system leverages kinetic mechanisms to achieve dual objectives: strategically redistributing incident solar energy and enhancing internal user comfort. This paper outlines the comprehensive schematic design phase, integrating performative architectural design through daylighting simulations and mechanical kinematic schemes with an interdisciplinary research approach bridging architecture and mechanical engineering domains. Furthermore, it provides detailed insights into the evolutionary progression of transformable joints and cable-driven actuation systems for practical solutions to the kinetic façade system.

Keywords. Adaptive Kinetic Façade, Interdisciplinary Approach, Cable-driven actuation, Dual-Functional, Performative Design

1. Introduction

Effectively harnessing solar radiation incidents on building facades is essential for a sustainable building operational energy cycle (Alva et al., 2020). This research explores the advantages gained from strategic design and deployment of façade modules in high-rise buildings. In this context, KineticSKIN, an innovative dual-functional adaptive façade system, is introduced to seamlessly adapt to both external environmental conditions (performance) and the needs of indoor occupants (user).

The energy consumption in lighting and HVAC (Heating, Ventilation and Air Conditioning) systems in buildings is significantly affected by seasonal variations.

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 435-444.

During the summer, the high solar altitude causes the UHI (Urban Heat Island) effect leading to overheating in urban areas. Conversely, in winter, the low sun altitude, while providing ample sunlight, causes glare disturbances. Beyond these external factors, user preferences are subjective depending on diverse needs and purposes. Hence, providing user control over lighting levels, views, and other parameters is essential for ensuring a comfortable user experience.

To achieve dual-functionality, three components are considered throughout the process: Morphology, Kinematics, and Construction (Figure 1). Based on a performative design methodology (Oxman, 2008; Soudian and Berardi, 2020), daylighting simulation results were integrated into the morphology design. Through these simulations, the effectiveness of seasonal countermeasures and the impact of user controllability has been assessed and verified.

Literature reviews of kinetic facades emphasize that their successful implementation in buildings depends on kinematics, reliability, and durability (Sharaidin and Salim, 2012). Thus, the challenges in prototyping lie in achieving durability, ease of installation, and maintenance without compromising architectural value. In response, an interdisciplinary approach was adopted by leveraging mechanical engineering expertise, specifically in the field of system dynamics, to ensure the reliability and controllability of the kinetic system.

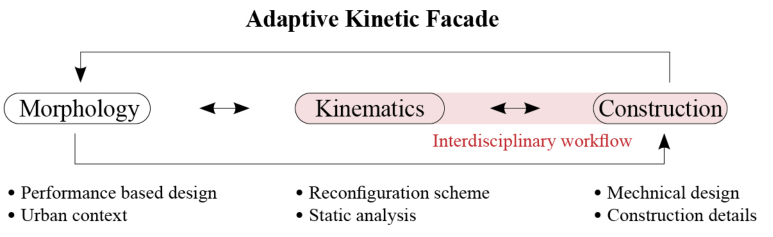


Figure 1. KineticSKIN design process with interdisciplinary workflow

This paper comprehensively details the evolution of KineticSKIN, encompassing the design, simulation, exploration of transformable joints, and actuation principles.

2. KineticSKIN - Adaptive Kinetic Façade (AKF)

In the preceding phase of research, illuminance and solar radiation simulations have been conducted with various facade typologies. The outcome highlighted the advantages of incorporating a kinetic feature that adapts to given conditions while using the same amount of material (Jeong et al., 2022). However, the optimal configuration in relation to the sun's position may not always be compatible with the user's preferences. Hence, with an emphasis on reconciling the inherent trade-offs between system performance and user experience, this section elaborates the overall design concept, simulation-based verification, and materialization scheme of KineticSKIN.

2.1. DESIGN CONCEPT

The concept started from a foldable bifurcated module, featuring upper and lower wings. The upper wing primarily contributes to the global energy cycle, particularly within urban context by redistributing the solar energy. In summer, it dynamically changes the folding angle to track the sun to mitigate solar heat gain or harvesting solar energy over time. During winter, when solar altitudes are lower, the upper wings reorient solar energy towards indoor ceilings, maximizing interior natural lighting while minimizing sun glare (Figure 2). Complementing the upper wings, the lower wings are user-controlled, allowing occupants to customize indoor conditions to meet specific requirements, considering various usage scenarios such as illuminance, view

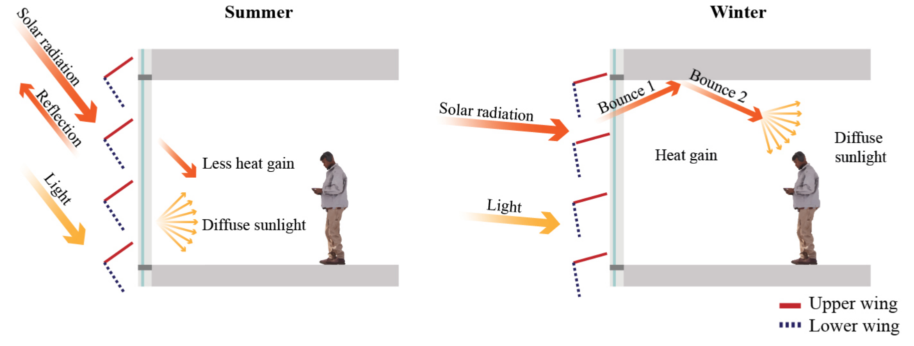


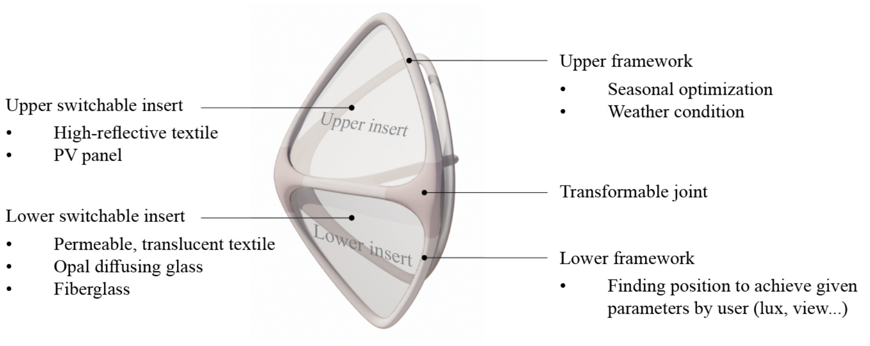
Figure 2. Seasonal countermeasure with respective role of each wing

and ventilation.

The hardware system is designed as a framework to accommodate interchangeable inserts, such as PV panels or high-reflective textiles, based on the roles of the wing. This design allows for versatility in serving a wide range of purposes (Figure 3).

Prior to investigating the specifics of the kinematic system, an assessment of the durability and folding capacity required determining the maximum and minimum folding angles of the facade modules. The findings (based on typical office hour, from 8 to 18 in Stuttgart) suggest that a maximum folding angle of 65º is sufficient for optimal performance. However, a visual study (Figure 4) revealed that pushing the

Figure 3. Framework and material concept of KineticSKIN



maximum folding angle to 85º significantly enhances flexibility, particularly for internal user views and extreme weather cases.

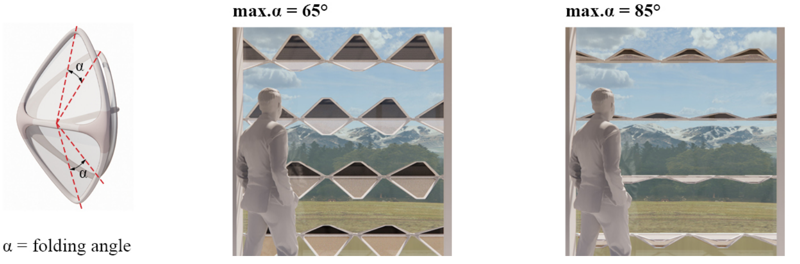


Figure 4. Visual study of maximum folding angles

2.2. ILLUMINANCE & RAY-TRACING SIMULATIONS

To assess the dual goals of user comfort and performance, ClimateStudio (Solemma LLC, 2020) and Ladybug (Roudsari and Subramaniam, 2016) were used for illuminance and ray-tracing simulations. As a parameter, the folding angles of the wings are adjusted to meet the needs of the user as well as to correspond with the position of the sun. To analyse seasonal usage, Simulations are conducted at noon on the summer solstice and winter solstice in an office environment based on Stuttgart, Germany, measuring 10m in depth and 6m wide, with south-facing windows.

To find the optimal folding angle of the upper wings, ray tracing simulations are preceded, providing a visual representation of the reflection of solar rays. Subsequently, the defined angle is used as input for illuminance simulations.

As a result, in summer, the folding angle of the upper wing is set to align its normal vector with the sun ray's altitude, reflecting direct sunlight to the external environment and facilitating effective capture for energy generation. This feature also demonstrates significant potential in alleviating Urban Heat Island (UHI) effects caused by conventional high-reflective curtain wall façades. In contrast, during winter, the folding angle is set to almost fully open angle to redirect incident solar energy toward the indoor ceiling. This promotes the deep penetration of diffuse sunlight into the interior environment (Figure 5) in an equivalent manner to light shelves.

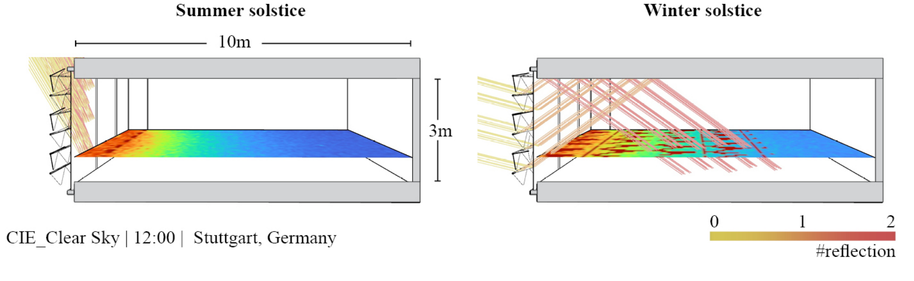


Figure 5. Ray-tracing visualization of direct sunlight

Illuminance simulations were conducted while maintaining the upper wing at angles derived from ray tracing simulations. To assess the user's control over indoor lighting levels and views, the folding angle of the lower wing was tested at both fully open and closed cases. Simulations were conducted using the glazing (U-value = 1.05)

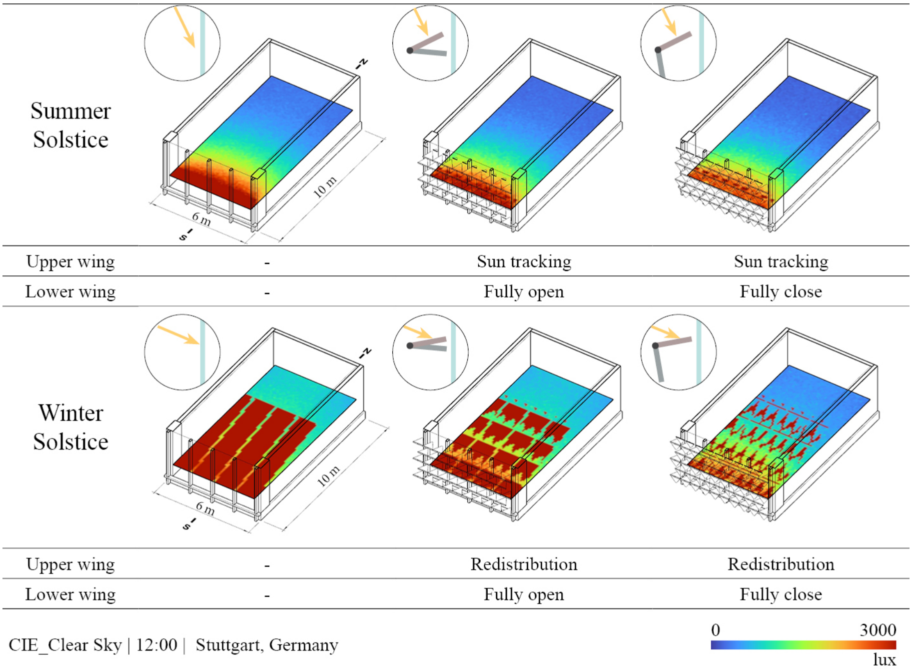


Figure 6. Illuminance simulation results based on seasonal adaptive countermeasures

and aluminium (white, high reflective) for the frames and inserts. Additionally, simulations of the glass-only case are conducted as a benchmark (Figure 6).

3. Development of Kinetic Mechanism

With the aim of installing a full-scale prototype in an outdoor setting, an interdisciplinary team developed the kinetic mechanism design. This section offers a comprehensive overview of the development process, with a focus on enhancing durability of folding frameworks, incorporating wind load factors, and minimizing the number of actuators.

3.1. HINGE DEVELOPMENT

Iterative development considered various materials, actuation principles, and hinge mechanisms, striking a balance between manufacturing ease, maintenance, functional requirements, and aesthetics with lightweight construction principles. Four sequential alternatives are elaborated in this section.

3.1.1. *Compliant Mechanism Hinge*

As the starting point, the first prototype (Figure 7) featured a monolithic structure with elastic TPU elements —acting as springs and hinges—and backside actuation cables. TPU elements compensate for wind loads through their elasticity offering following advantages: Reduced mechanical components (joints, springs, etc.), minimized friction, and no need for lubrication. Despite its advantages, the prototype proved unsuitable for the final solution due to factors like scalability and wind resistance requirements at higher elevations.

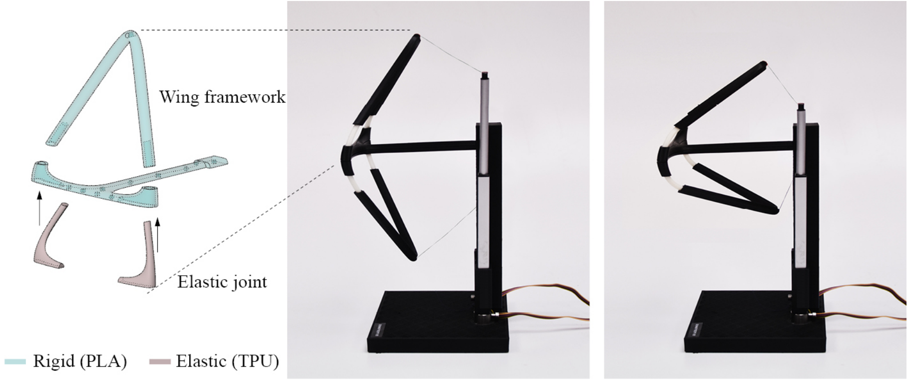


Figure 7. Compliant hinge joint with cable mechanism and 3D printed prototype in 1:5 scale

3.1.2. *Elastomeric Layered Panel*

Alternatively, elastomeric layered panel by employing stainless steel spring sheet has been explored. It is easy to assemble, ideal for integration of various insert panels including PV panel and provides the sufficient stiffness against wind loads. However, due to the required length of the spring sheet, limitations in the folding angle (as discussed in Section 2.1) and effective surface have been arose which prevented it from further development.

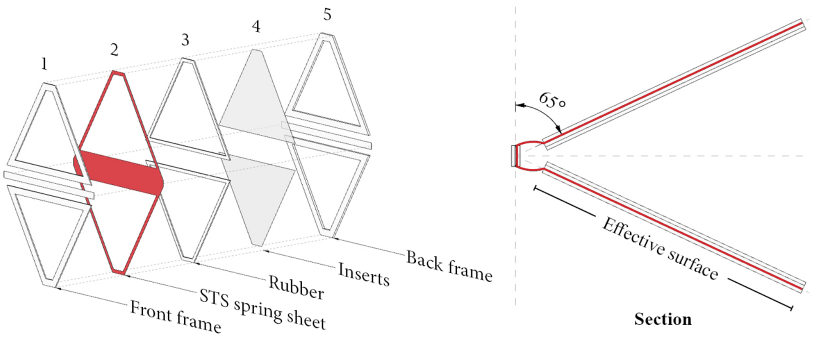


Figure 8. Elastomeric layered panel with Spring metal sheet exploded drawing & section

3.1.3. Torsion Spring Hinge

Further considerations prompted the use of torsion springs to bear wind loads (Figure 9). By integrating a rotating shaft and suspension subframe, the joint part was minimized, ensuring maximized effective surface area. However, scaling challenges have emerged, particularly at higher elevations, resulting in the requirement of larger springs, which is inconsistent with the principle of minimum joint parts.

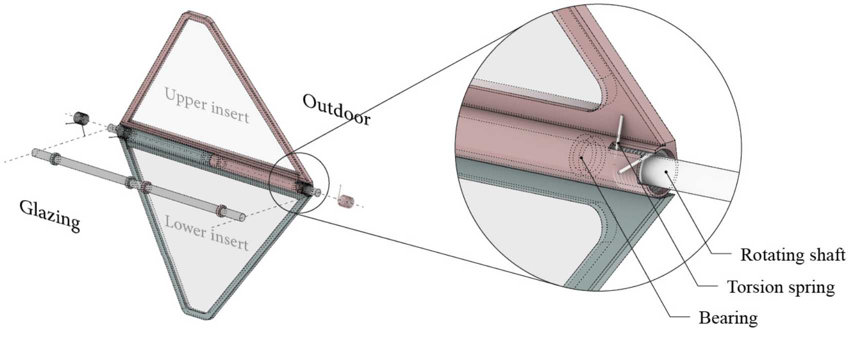


Figure 9. Torsion spring hinge concept

3.1.4. Dual-cable driven Mechanism

The spring-loaded system transitioned to a dual-cable system, featuring cables on both the back (inside) and front (outside), see Figure 10. The constant opposing tensions in these cables ensure stability during dynamic movements and sustain a static position for certain periods. These tension forces maintain equilibrium, facilitating controlled adjustments and positional stability. These features make the dual-cable driven mechanism suitable to the façade system that requires precision and durability.

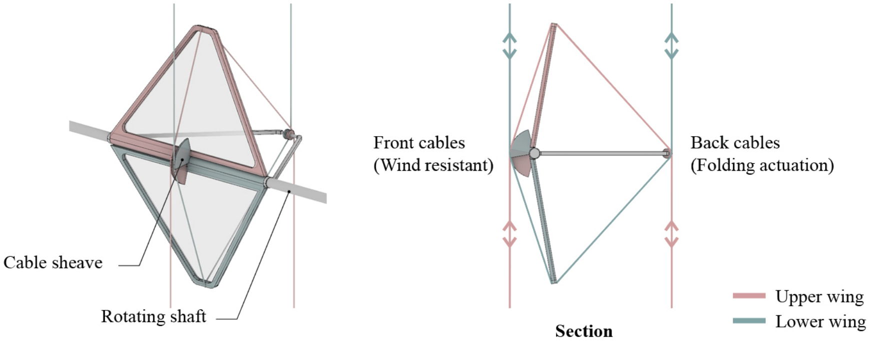


Figure 10. Dual cable system concept

In the meantime, kinematic design iterations were reflected in morphology design. Besides the integrated hinge from Section 3.1.3, an arc-component is added to the surface of each wing to guide the cables on the outside. The arc-components allow only one rotary motor for both cables by making similar dynamic variations in their lengths.

3.2. MOTORIZED CLUSTER

The primary advantage of cable-driven actuation lies in the ability to connect numerous modules to a single actuator, achieving mechanical linkage with minimal visual impact and thereby reducing the overall number of actuators. This enables the grouping of modules, where multiple units share a motor, facilitating efficient cable actuation.

Each upper and lower wing is linked to pulleys on the rotating shaft of an electric worm gear motor via steel wires. The folding motion is controlled by winding the front and back cables in different directions on separate pulleys, allowing for controlled retraction and extension based on the rotary direction of the motor. For a single mullion span about 2m, a cluster of 12 modules is managed by two motors—one for the upper wings and another for the lower wings. The rotary motor B, situated on the floor, operates the 12 upper wings, while the rotary motor A, mounted on the ceiling, drives the lower wings (Figure 11).

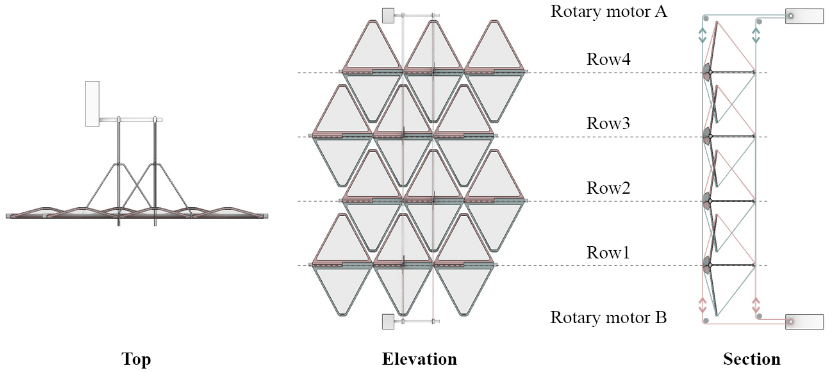


Figure 11. Motorized cluster and layout

4. Result and Discussion

Four sequential developments of transformable joints are presented, with a focus on their mechanical properties and flexibility. This study provided insight into how these joints can be reconfigured for facade applications that require a minimum number of actuators, durability, precision, and flexibility.

The simulation results with the façade system revealed significantly improved illumination levels during both summer and winter solstices compared to glazing alone. In summer, the concentrated incident near the glazing led to extremely high illuminance level, which is mitigated by 30-40% depending on the lower wing's position. During winter, the upper wing effectively redirected low-altitude sunlight to the interior ceiling. As shown in Figure 6, the reflected diffuse light illuminated the interior, providing adequate brightness into the deeper area. The role of the lower wing in both seasons is encouraging, showing that the lower wing alone can provide 10% of indoor illuminance control in summer and 30% in winter. The difference between mean and median lux values is notably reduced by approximately 10% compared to without façade module (Figure 12). This improvement indicates enhanced uniformity in daylighting, particularly attributed to the diffuse light.

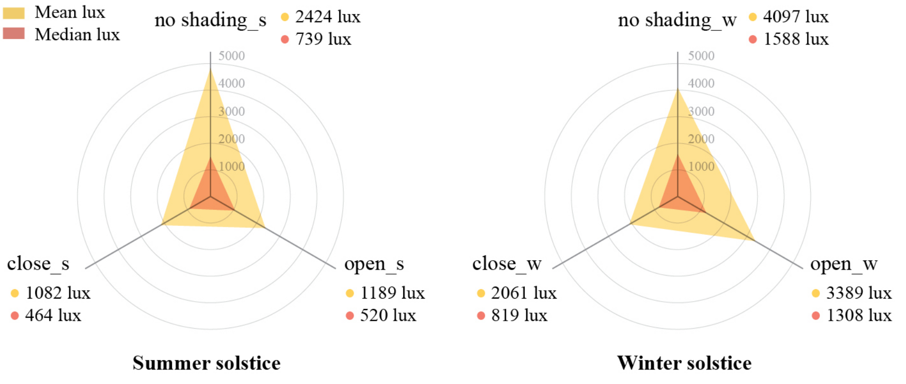


Figure 12. Comparison between mean and median lux level

In the near future, KineticSKIN will be showcased on a 6m wide and 3m in height glazed façade—a full-scale prototype (Figure 13). There will be three clusters of 12 modules each, covering a span of 2m with only two actuators. Next steps involve verifying the kinematic design through a 1:2 prototype and detailed manufacturing of the façade modules. This phase will validate kinematics at a larger scale and explore the feasibility of interchanging insert materials.

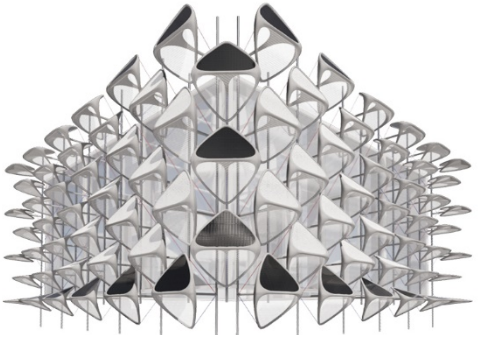


Figure 13. Vision of KineticSKIN

5. Conclusion

AKFs have been studied extensively for their potential in sustainable building energy cycles (Barozzi et al., 2016). Despite their recognized benefits, practical challenges, including installation and maintenance complexities, limit their widespread use. This paper addresses these challenges by presenting an integrated design and development process for an AKF that not only preserves architectural value and functionality but is also practical for implementation.

The outcome is a versatile façade system featuring adaptive design elements, informed by seasonal daylighting simulation results. Employing an interdisciplinary approach, various kinematic design options were rigorously analysed and developed. The result is a cable-driven actuation system that enables folding motions with high resolution while ensuring wind stability through cable interconnection, minimizing the number of required actuators.

This research distinguishes itself by emphasizing a harmonious balance between technical and architectural considerations. By adopting a continuous feedback loop and iterative processes within an interdisciplinary workflow, this integrated approach not only optimizes the entire system for functionality but also streamlines its implementation.

Acknowledgements

This research is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project- ID 279064222 – SFB 1244.

References

Alva, M., Vlachokostas, A., & Madamopoulos, N. (2020). Experimental demonstration and performance evaluation of a complex fenestration system for daylighting and thermal harvesting. *Solar Energy, 197,* 385-395. doi:10.1016/j.solener.2020.01.012

Barozzi , M., Lienhard, J., Zanelli, A., & Monticelli, C. (2016). The Sustainability of Adaptive Envelopes: Developments of Kinetic Architecture. *Procedia Engineering*, 275-284. doi:10.1016/j.proeng.2016.08.029

Jeong, M. Y., Matheou, M., & Blandini, L. (2022). *Materials Ecology 2— Optimization of Daylighting Performance and Solar Heat Gain.* (M. Kanaani, Ed.) New York: Routledge. doi:10.4324/9781003183181-23

Oxman, R. (2008). Performance-Based Design: Current Practices and Research Issues. *International Journal of Architectural Computing*, 1-17. doi:10.1260/147807708784640090

Sharaidin, K., & Salim, F. (2012). Design Considerations for Adopting Kinetic Facades in Building Practice. *eCAADe* (pp. 619-628). Prague: Czech Technical University in Prague, Faculty of Architecture. doi:10.52842/conf.ecaade.2012.2.629

Soudian, S., & Berardi, U. (2020). Development of a performance-based design framework for multifunctional climate-responsive façades. *Energy and Buildings*. doi:10.1016/j.enbuild.2020.110589

Solemma LLC. (2020). ClimateStudio Documentation. Retrieved from https://climatestudiodocs.com/

Sadeghipour Roudsari, M., & Subramaniam, S. (2016, 8 30). Automating Radiance workflows with Python. Retrieved from The 15th Annual Radiance Workshop. Padua, Italy: https://www.radiance-online.org/community/workshops/2016-padua/presentations/213-SadeghipourSubramaniam-AutomatingWorkflows.pdf

COASTAL INFRASTRUCTURE DESIGN: RESEARCHING SEA-WAVES AND TEXTURED SURFACES INTERACTION USING PHYSICAL AND VIRTUAL WAVE FLUMES

ROY KOZLOVSKY [1], YASHA GROBMAN [2] and HANNA LEVY[3]
[1]Tel Aviv University, Tel Aviv, Israel; [2]Technion Israel Institute of Technology, Haifa, Israel; [3]independent scholar.
[1] rkozlov@tauex.tau.ac.il, http://orcid.org/0000-0003-4079-4637
[2] yasha@technion.ac.il, http://orcid.org/0000-0003-4683-4601
[3] hanny_levy@hotmail.com, https://orcid.org/0009-0001-6030-0588

Abstract. Projected global rise in sea level and intensification of storms place the shoreline at risk, requiring extensive investment in coastal defence infrastructure. These structures are designed to efficiently dissipate wave energy at the expense of ecological and landscape values. The aim of the research is to establish a multifunctional approach to coastal infrastructure. Within this framework, it proposes a method for utilising simulation tools to creatively shape the interaction of sea waves with coastal structures for scenic and ecological benefits. It sets two primary goals: to establish that computational fluid dynamics tools can be used by architects to design the interaction of sea-waves with solid surfaces. This goal is explored by creating a digital simulation of a physical wave flume facility, and running physical experiments to calibrate the virtual simulation tool. Secondly, it uses these tools to systematically explore the range of possibilities latent in wave-structure interaction by initiating basic research into the flow properties of different types of textured surfaces to improve the aesthetic and ecological performance of such structures.

Keywords. Computational Fluid Dynamics, Coastal Infrastructure, Ecological Enhancement, Textured Surfaces, Physical and Virtual Simulations, Computational Design.

1. Introduction

Current forecasts of rising global sea levels and intensification of storms place the shoreline at risk, requiring extensive investment in new coastal defence systems and modifying existing facilities to protect urban development and critical waterfront infrastructure such as ports and power plants. The massive construction of seawalls and other types of protective coastal infrastructure may conflict with ecological and urban values, by employing materials and techniques that contribute to climate change, reducing public access to the shoreline and damaging sensitive coastal ecosystems (Hosseinzadeh et al. 2022). Mitigating these conflicts requires a multi-functional and

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 445-454.

multispecies approaches in which defensive structures assume additional functions such as providing ecological services, parks and seaside promenades (Salauddin et al. 2023; Grobman et al. 2023). These approaches provide the opportunity to rethink the design of these structures, and especially how they interact with incoming waves.

The architectural, creative approach to designing these structures conceives sea waves differently than the defensive approach of coastal engineers. First, it would plan for moderate and calm sea states, which coastal engineers discount because their main design parameter are peak waves. Second, it would attend to the ways in which breaking waves flow back to the sea. This type of flow is of lesser interest to engineers, as it poses little risk to the stability of the structure or its resistance to overtopping. Third, it would attend to the texture scale of these structure, which is not an important factor for defensive design of marine infrastructure. To address these gaps in knowledge, the proposed research explores the ecological and experiential potential of coastal infrastructure by focusing on everyday sea conditions and the scale of texture.

The interaction of the sea with the built shoreline for both run-up and run-down flows is difficult to model because of the non-linear and turbulent behaviour of sea waves as they encounter a solid boundary and break. This complexity presents technical difficulties for developing methods for creatively shaping them through morphological and material interventions.

Until recently, the main method for studying the real-life interaction of sea waves with a solid structure without building it first has been to test a physical model in a wave flume facility. A major impediment to integrating wave flumes into architectural research is the high costs and time-consuming process of preparing and testing models in this highly specialized and expensive laboratory equipment.

Recent developments in digital computation and simulation technologies may facilitate design-oriented research into sea-wave flows (Chronis et al. 2017). Computational Fluid Dynamics (CFD) are being used by architects and designers to model and visualize airflows to design, healthy and energy-efficient buildings, as well as predict structural wind loads and environmental comfort around them. The aim of this paper is to expand the application of these tools to the design of coastal structures. This is possible because specialized CFD tools for simulating the complex interaction of sea waves with solid surfaces have advanced past the threshold of accuracy that enables architects to use them as substitutes to physical wave flumes, therefore removing a high entry barrier to this emerging field of design.

Previous research into multi-functional coastal infrastructure design utilized CFD simulation tools to explore seawall morphology as a means for creating different types of flow effects (Kozlovsky & Grobman, 2016; Grobman et al., 2017). Its contribution remained speculative since these tools were not empirically validated. This paper presents the next stage in the research program that addresses this gap. It sets two main goals: Validating the claim that currently available computational fluid dynamic tools can be reliably used to research the interaction between solid surfaces and sea-waves by comparing the virtual simulations with physical tests in a wave flume. Secondly, using these tools to systematically explore the range of possibilities latent in wave-structure interaction by initiating basic research into the flow properties of different types of textured surfaces. This stage of research contributes to CFD aided design the relatively unexplored scale of texture.

2. Methodology

The research combines analytical and experimental methods. It simplifies the infinite variety of textures and sea-wave conditions into basic parameters, and tests their interaction in a physical wave flume. The results are used to calibrate the virtual wave flume, and create systematic knowledge on texture-induced fluid motion.

The experiment is divided into three stages. In the first, it establishes simplified parameters for designing and fabricating textured surfaces such as geometry, density and extrusion, and standardized sea-wave flow variables for testing their performance in a physical wave flume, such as surface inclination and wave height. In parallel, the researchers utilize computational tools to construct a virtual wave flume based on the properties of the physical wave flume facility available to them.

In the second stage, the parametrically produced textures are tested in a physical wave flume. The results of the experiments are used to calibrate the computational tool and evaluate its capacity to achieve a reasonable degree of similitude.

In the third stage, the experimenters develop a set of criteria for measuring and assessing the flow characteristics of each texture, and used the comparative method to generate knowledge on the impact of geometry, density and extrusion on run-up and run-down wave motion. Based on the initial empirical findings, it proposes a method for systematically exploring the range of possibilities latent in wave-structure interaction and integrating this knowledge into the early stages of the design process.

3. Texture parameters

In architectural discourse, texture is defined as a formal category that accounts for the physical and visual qualities of surfaces. Modern architecture defined texture as subservient to primary categories of space, structure or form. Considered superficial, it is permitted if it adheres to the principle of material honesty, that is, exposing the material's inner essence or the workmanship and tools applied to it. As an operative concept of design thinking, texture is considered a secondary artistic means for unifying or varying planes that enclose space, expressing the weight of a form, reflecting or absorbing light, and other compositional aims (Ching, 2007). In phenomenologically informed practice, texture's association with the sense of touch and tactile modes of perception informs its revaluation as a significant component of architectural experience (Pallasmaa, 2012). Digital technology and culture are reframing texture as an artifice of a new aesthetics that registers themes such as emergence, fluidity, animation, interactivity and differentiation (Schumacher, 2009).

Texture in coastal infrastructure has a functional rather than visual or experiential significance. It is assessed in terms of its performance, be it in dissipating wave energy or increasing biodiversity, suggesting that surface texture is not as a fixed, autonomous property, but a process that shapes and is shaped by external forces and programmatic requirements.

Hydraulic engineers consider texture in terms of its energy dampening effect on wave run-up. Increase in surface roughness is correlated with higher energy dissipation through friction and turbulence. This relationship, expressed in empirical equations, is used to determine the size of elements for building riprap revetments (van der Meer, 2018; Kreyenschulte et. al 2020).

The dimension of texture at its different scales has been explored by marine biologists, since it has a quantifiable impact on the ecological performance of coastal structures (Sella et al., 2022). Smooth surfaces that characterize modern concrete coastal structures have an adverse impact on biodiversity, since many species are adapted to naturally occurring textures. Increase in texture density enables more species to attach themselves to the surface or find shelter from predation, resulting in more abundant and diverse intertidal ecosystems. The practical application of biological research into the habitat properties of texture is the bioengineering of coastal and marine structures to recruit and sustain complex ecosystems. The texture may have aesthetic value, but it is incidental to its calculated effect on biodiversity.

The engineering and biological understanding of texture as a function of hydraulic or biological performance, can be transferred back to architectural thought, leading to the concept of the performative surface. The shape of the solid texture is a function of its effect on water flows. Since water is transparent, one can simultaneously observe both the texture of the solid surface in itself, and the emerging pattern of water flows during run-up and run-down. To begin exploring this unique condition, while insuring that the research into texture is not biased towards functional goals such as optimizing the structure's engineering and ecological performance, or humanizing texture by mimicking pre-existing culturally meaningful patterns, the research develops an analytical approach that sorts texture into basic morphological parameters.

3.1. BASIC TEXTURE PARAMETERS AND MANUFACTURING

In order to establish a basic understanding of how texture could be purposely applied to shape flows for practical and creative aims, the research simplifies texture into basic types of point-based and line based textures. These groups are tested for three elementary geometric forms, the circle, the triangle and the square.

The textured plates are of standard size that fits the width of the physical wave flume. Individual plates are designed to analyse one morphology. Using a parametric model in Grasshopper plugin for Rhinoceros design software, each plate includes a distribution of texture scale and roughness. Density of texture varies from highest to lowest value from right to left. Extrusion is distributed from lowest at the top to highest at the bottom according to a Voronoi grid which changes density of cells according to the texture density distribution described above (Figure 1). This setup is designed to study the minute variations in flow regimes and identify emerging flow patterns that are scale sensitive. Altogether seven texture pattern plates were constructed from birch wood plates using a CNC machine.

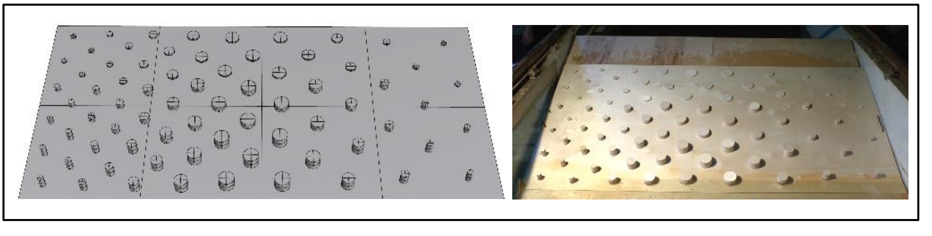


Figure 1: Point-based circular texture plate.

3.2 PLATE INCLINATION

The inclination of a textured surface is an important parameter for wave performance, both for run-up and run down flows. Due to gravity, run-up velocity is reduced by steep inclination, while run-down velocity is accelerated. Each inclination would produce a different flow regime for the same pattern. Minute differentiations in inclination angle may produce a qualitative change in flow regime. While digital simulation tools allow to test any inclination with a simple command, the practical constraints of time and costs posed by experiments in physical wave flumes limited the inclination parameter to two angles. The inclinations of 20° and 40° were chosen in accordance with the optimal inclination of maritime structures such as revetments and seawalls. Two demountable substructures with the mentioned above inclinations were built from wood to support the removable textured plates during the physical trial.

4. Wave flume experiment

The laboratory experiments were carried out in the Coastal and Marine Engineering Research Institute (CAMERI) at the Technion, Haifa. The facility's 45×2.45×1.5 meter wave flume is equipped with a wave generator, computerized control for real sea simulation, wave gages and high sampling rate pressure gages.

The water depth was set to 1 meter while the wave piston was set to move 10 cm back and forth, generating sinusoidal waves with height of 0.188 m and amplitude of 0.09. This type of waves imitates weak sea state level 2 according to the World Meteorological Organization scale. The simulations were recorded from three points of view simultaneously: from a top view, from a side view through a glass window to the side of the plate and from the back (Figure 2).

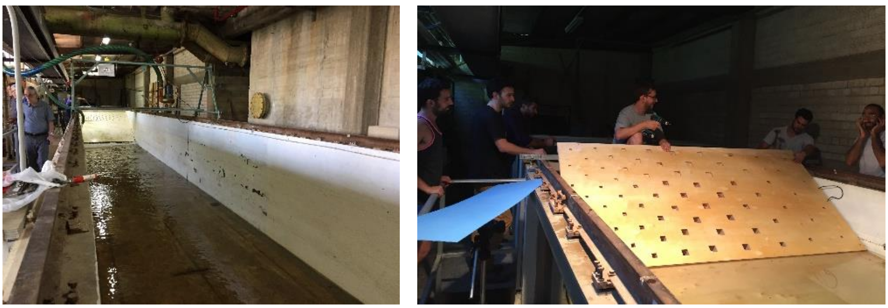


Figure 2: CAMERI wave flume with demountable texture plate.

5. Experimental Results

Initial examination of the results of the physical simulation showed that texture has a significant qualitative and quantifiable effect on water flow regimes, and that this effect can be analysed to establish a basic understanding of how texture could be purposely applied to shape flows for both practical and creative aims.

After observing the visual properties of run-down and run-up of flows on different texture types, it was decided to generate quantitative data on a series of parameters considered significant for understanding the relationship between texture and flow:

- Run-up and run-down flow times: the time it takes the wave to climb the textured plate, and the duration of the run-down flow.
- Run-up distance: the maximal distance reached by the wave during run-up motion.
- Flow wake width above and below: the width of the trail formed above and below the obstacle during run-up and run-down.
- Flow splash height: the height of the splash created by the impacting wave.
- Returning flow water jump height: the height of the jump of returning flow over textural elements.

Measurement were taken for three different waves. The outcomes shown below are the result of a mathematical average of the three sensitivity checks.

Parallel to the quantitative analysis, the observations identified qualitative flow patterns generated by specific textures, and analysed the parameters for their emergence.

5.1. QUANTITATIVE ANALYSIS OF TEXTURED FLOW

This section presents the most significant findings of the quantitative analysis.

5.1.1. Run-up and Run-down Time

The run-up and run-down times for each dot-based texture placed with slopes of 20 and 40 degrees is presented in table 1 as percentage value to maximal time.

One first observation is the likeness of run-up time records of the plates with cavities and textures both with an inclination of 20, which variate only by 7%. The second observation is the likeliness between the run-down time of the texture plates with slope of 40° and that of the cavities plates with a 20° inclination.

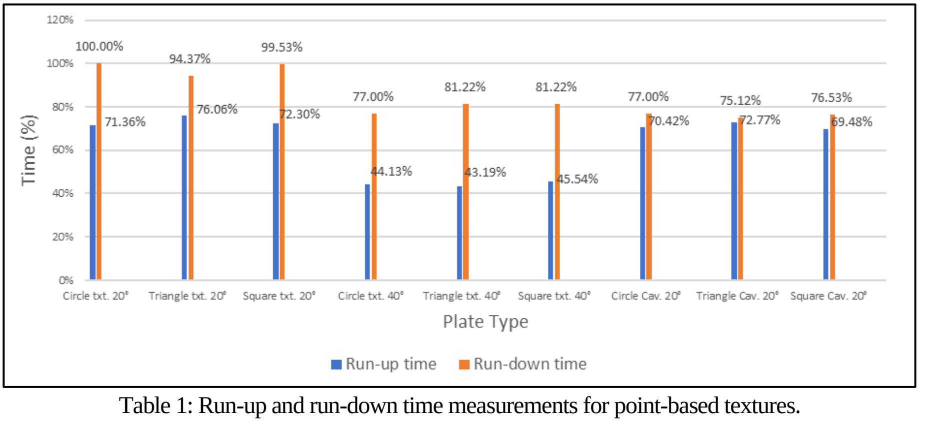


Table 1: Run-up and run-down time measurements for point-based textures.

The conclusion from the table is that the run-up time is almost the same for the circular texture and cavities plate (difference of 0.94%) while the difference slightly increases for triangle (3.29%) and square (2.82%) texture geometries. This happens because circles have a hydrodynamic shape in both directions of flow. In addition, the change in slope angle from 20° to 40° produces a stronger effect on run-down time, due to the effect of gravity.

5.1.2. Water Wake Width

Figure 3: wakes formed during run-down in circular texture at 20° slope.

This set of measurements investigate the flow created after the run-down flows pass the textured field (Figure 3). The wake width measurement in Table 2 was taken at a distance of 80 mm below the obstacle. It presents an average of the three measurements. From the resulting trend line it is possible to conclude that, by average, the width of the wake on the y-axis is about 2.4 times bigger than the diameter, shown on the x-axis. This value can be used by designers to create a trail that has a desired width by playing with the diameter of the circular textures.

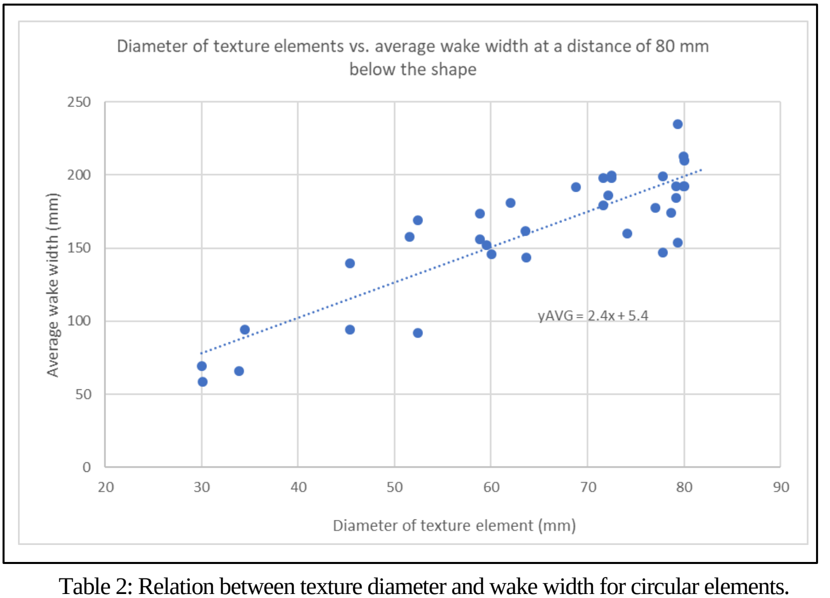


Table 2: Relation between texture diameter and wake width for circular elements.

5.2. QUALITATIVE ANALYSIS OF TEXTURE FLOWS

Concluding the catalogue of visual effects of texture are unexpected phenomena that may have aesthetic or educational value. These effects may be specific to a given texture geometry or flow condition such as the first, weakest run-up flow of a series.

One of these visual effects (figure 4) was a water jump created by the returning flow. The height of the jump could be calculated in relation to the texture extrusion height, surface inclination, and the position of the protrusion on the plate.

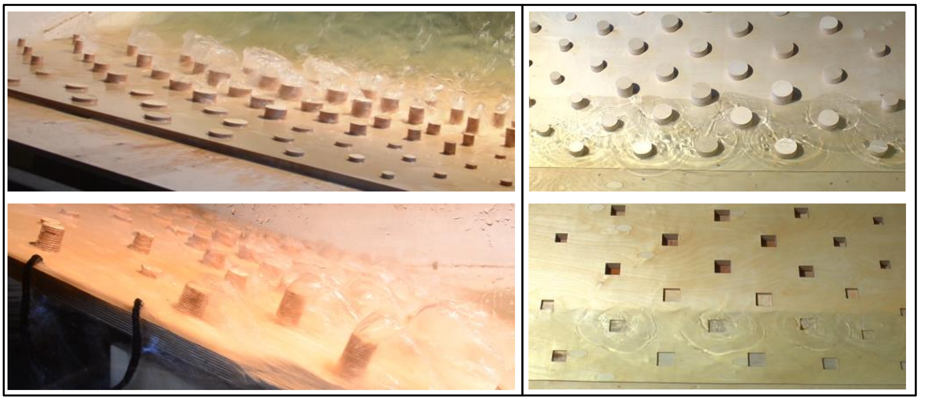


Figure 3: quantitative effects: wake jump. Traces produced by cavities.

Additional qualitative visual effects can be classified as echoes or traces left by the texture on the water surface (Figure 5). This delicate wave event was observed in the first, weakest wave of the series, for both protruding and cavity based textures. The visual interest in this wave taxonomy can be attributed to the tension between the texture's geometry and the circular movement of waves. As discussed in the introduction, this effect is limited to calm sea state, a relatively under-studied condition.

6. Calibration of Virtual Simulation Tools

The final step of the research dealt with the calibration of the virtual simulation. A digital version of the CAMERI wave flume was prepared according to its true measure, including the properties of its wave producing mechanism. The digital drawings of the textured plates were inserted into the virtual flume to run the simulation. The researchers used visual records and measurements from the physical wave flume to calibrate the digital tool by adjusting flow parameters such as viscosity, particle separation, and surface tension, as well as the movement of the virtual wave generator to reproduce the specified wave height of the physical experiment.

The initial computational set-up for the virtual simulation used Blender and its recent fluid simulator Mantaflow. Due to inconsistencies, it was decided to compute the simulation step with Houdini and then import it to Blender for final rendering.

The digital model of the wave flume and textured plates was imported into Houdini. Using the simulation of the triangular texture at 20° as a reference model, the flow simulation parameters of Houdini were adjusted in an iterative process until the final visual effect was close enough to the behaviour recorded in the physical wave flume experiment. To make the water more realistic, the element was composed of the actual fluid (made of FLIP particles) and white-water particles to recreate drops, bubbles and foam. Once the flip particles were simulated, little meshes were generated and located on each white-water particle to transfer their location to Blender.

As to the accuracy of the process, the calibration exercise shown in Figure 6 shows inconsistencies between the virtual version and the one obtained in the laboratory. However, there is a clear similarity between the run-up distance profile reached by the virtual and the laboratory wave. In fact, as in the real experiment, in the virtual version the wave reaches the maximal distance on the right side of the plate while it is minimal on the left side of the plate. The virtual tool therefore replicates the findings that higher texture density produces shorter run-up distance.

Our conclusion from the calibration process is that the tool can be used by designers to create textures and other features of coastal infrastructure, but it cannot at this stage of technical development substitute physical wave flumes for engineering and other purposes that require accurate simulation of fluid motion.

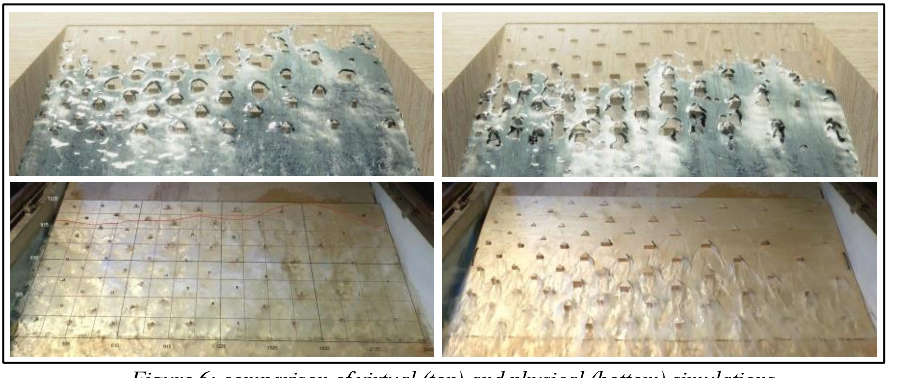


Figure 6: comparison of virtual (top) and physical (bottom) simulations.

7. Discussion and conclusions

The main aim of the research was to demonstrate the potential of CFD simulation tools to advance a multifunctional approach to coastal infrastructure. It initiated basic research into the capacity of surface texture to shape the ways in which waves break on the built shoreline for programmatic, ecological and experiential benefits. This was pursued by testing basic geometric textures and developing protocols for measuring and assessing flow regimes, as a first step in closing the gap in knowledge of textured surfaces-sea wave interaction.

The main findings of the experimental stage were the quantifiable impact of different texture morphologies on wave behaviour. Run-up distance is a function of texture density and geometry, while run-down time and wake length are closely correlated with the hydrodynamic characteristics of texture morphology, scale and density. Further investigations are required to assess the effect of texture on the ecological and engineering performance of waterfront structures.

The qualitative assessment of flow characteristics of the textures observed in the physical experiment foregrounded flow effects that have aesthetic and experiential value, mainly for run-down flows. This suggests that the otherwise overlooked phenomenon of wave run-down could be further developed for designing the experiential component of waterfronts and coastal infrastructure.

The research engaged with architectural issues beyond the expert domain of coastal infrastructure. The experimental use of digital simulation tools combined with digital fabrication to produce the textured plates may also open new ways of conceptualizing solid surfaces. Simulation technologies bring to awareness the dynamics of change, transformation and flux. This focus informs an epistemological shift from the object and its autonomous formal or structural properties to the complex, emergent interaction between fixed and fluid elements. Since water is transparent, the observer can perceive the material and fluid textures simultaneously as they enfold in time. Likewise, the process of designing such surfaces entails an iterative process of attenuation of the texture in relation to the feedback of the simulation. This dynamics expands the concept of parametric surfaces to include the dimension of performance, be it ecological, hydraulic or experiential.

References

Ching, F. (2007). *Architecture: Form, Space & Order* 3rd ed. Wiley.

Chronis A., Dubor, A., Cabay, E. & Roudsari, M. Sadeghipour (2017). Integration of CFD in Computational Design - An evaluation of the current state of the art. *Proceedings of the 35th eCAADe Conference* (pp. 601-610). Education and research in Computer Aided Architectural Design in Europe (eCAADe).

Grobman, Y. Kozlovsky, R. & Levy, H. (2017). A Multifunctional Computational Approach to Waterfront Design. *Architectural Science Review*, 60(6), 446-459.

Grobman, Y., Weisser, W, Shwartz, A., Ludwig, F., Kozlovsky, R., Ferdman, A., Perini, K., Hauck, T. E., Selvan, S. U., Saroglou, S. (2023). Architectural Multispecies Building Design: Concepts, Challenges, and Design Process. *Sustainability*, 15(21).

Hosseinzadeh, N. Ghiasian, M., Andiroglu, E., Lamere, J., Rhode-Barbarigos, L. Sobczak, J., Sealey, K. S. & Suraneni, P. (2022). Concrete seawalls: A review of load considerations, ecological performance, durability, and recent innovations. *Ecological Engineering*, 178(106573).

Kozlovsky, R., & Grobman, Y. (2016). The Blue Garden: Coastal Infrastructure as Ecologically Enhanced Wave-scapes. *Landscape Research Journal*, 42(5), 439-454.

Kreyenschulte M, Schürenkamp D., Bratz B., Schüttrumpf, H & Goseberg, N. (2020). Wave Run-Up on Mortar-Grouted Riprap Revetments. *Water*, 12(12), 3396.

Pallasmaa, J. 2012. *The Eyes of the Skin; Architecture and the Senses*. Wiley.

Salauddin M, O'Sullivan JJ, Abolfathi S and Pearson JM (2021). Eco-Engineering of Seawalls—An Opportunity for Enhanced Climate Resilience From Increased Topographic Complexity. *Frontiers in Marine Science*, 8(674630).

Sella, I., Hadary, T., Rella, A. J., Riegl, B., Swack, D., & Perkol-Finkel, S. (2022). Design, production, and validation of the biological and structural performance of an ecologically engineered concrete block mattress: A Nature-Inclusive Design for shoreline and offshore construction. *Integrated Environmental Assessment and Management*, 18(1), 148–162.

Schumacher, P. (2009). Parametric Patterns. *Architectural Design*, 79(6), 28–41.

van der Meer, J. W., Allsop, N., Bruce, T., De Rouck, J., Kortenhaus, A., Pullen, T., Zanuttigh, B. (2018). EurOtop: Manual on wave overtopping of sea defences and related structures. www.overtopping-manual.com

D-PREDICT: INTEGRATING GENERATIVE DESIGN AND SURROGATE MODELLING WITH DESIGN ANALYTICS

ESMAEIL MOTTAGHI[1], AHMED M. ABUZURAIQ[2]
AND HALIL ERHAN[3]
[1,2 3] Simon Fraser University, Canada.
[1]ema63@sfu.ca, 0009-0005-6082-9966
[2]aabuzura@sfu.ca, 0000-0002-3604-7623
[3]herhan@sfu.ca, 0000-0002-7630-6351

Abstract. The increasing importance of performance prediction in architecture has driven designers to incorporate computational tools like generative design and building simulations to widen and guide their exploration. However, these tools pose their own challenges; specifically, simulations can be computationally demanding and generative design leads to large design spaces that are hard to navigate. To address those challenges, this paper explores integrating machine learning-based surrogate modelling, interactive data visualisations, and generative design. D-Predict, a prototype, features the generation, management and comparison of design alternatives aided with surrogate models of daylighting and energy.

Keywords. Generative design, building performance assessment, surrogate modelling, machine learning, design analytics.

1. Introduction

Architects have increasingly turned to computational methods, including Generative Design (GD) and building performance simulations, as to develop sustainable built environments. Integrating performance analysis into the design exploration aims to proactively identify and address potential issues as early as possible, ensure sustainability (Iyengar, 2015), align project preferences and constraints with performance objectives (Bernal et al. 2019), and reduce environmental impacts, ultimately contributing to the resilience of built environments. Such integration is more pronounced in GD, where large sets of design alternatives can be rapidly generated.

However, both performance simulations and GD methods pose their own challenges to designers. First, performance simulations, such as for daylight use and energy efficiency, are computationally intensive, mainly when conducted for multiple design alternatives such as those generated through GD. A promising solution is adopting machine learning-based performance assessment methods, including surrogate models. These methods can replace laborious simulations with fast performance predictions, accelerating the exploration of alternatives (Westermann et al., 2019; Yousif and Bolojan, 2022). These methods also enrich design alternatives by

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 455-464. © 2024 and published by the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA), Hong Kong.

offering additional layers of building information.

Secondly, performance simulations and GD methods produce large and diverse volumes of data that must be managed and presented well to designers. Innovative interfaces that facilitate practical design data analysis are required to achieve that. Furthermore, exploring design alternatives can be overwhelming, particularly when confronted with a vast design space, leading to the choice overload phenomenon (Erhan et al., 2017). The situation is further complicated when dealing with multiple objectives and conflicting performance goals. To address this issue, architects and designers can rely on data analytics systems (e.g., Chaszar et al., 2016; Garg et al., 2019; DesignExplorer 2, 2023), such as interactive data visualisation, to navigate and assess the data-rich design space, enabling informed decisions while working on multiple alternatives.

In this study, we employ a design study methodology (Sedlmair et al., 2012), which builds on the literature review above and involves architectural design practitioners as collaborators in the research process. The result is D-Predict, a Design Analytics system prototype incorporating GD and machine learning-driven performance prediction and enhances design decision-making through interactive data visualisation. D-Predict focuses on two important but conflicting performance concerns instrumental for sustainable architecture: daylight and energy use. It combines interactive data visualisation, surrogate models, and GD to enhance design decision-making considering performance metrics.

2. Background

Numerous studies have been conducted to explore methods to enhance architects' design decision-making in terms of building performance. These studies have been categorised into three main sections. The first explores integrating the GD and building performance assessment into the design process (Bernal et al., 2019; Anton and Tănase, 2016). Architects used experience, while researchers applied GD principles to optimise solutions for conflicting goals like energy and daylight optimisation. The second category is more relevant to our research investigating how machine learning and surrogate models can enhance building performance prediction. Below we briefly discussed the relevant research on this topic following with research on improving design decision-making when confronted with a large set of design alternatives by employing data analytics interfaces.

2.1. COMPUTATIONAL BOTTLENECKS AND SURROGATE MODELS

Using simulation processes for building performance assessment is computationally intensive, requiring complete models, particularly when considering multiple performance criteria assessed together. However, a complete model means that designers must prematurely commit some decisions, removing the exploratory opportunities of design from the process. Architects have recently started using surrogate modelling with deep-learning methods to replace simulation methods for reducing the time and required labour when predicting performances in the early stages of design (Westermann et al., 2019; Yousif and Bolojan, 2022; Zorn et al., 2022).

Yousif et al. (2022) proposed an automated performance-driven GD approach aided with surrogate daylighting models to enable the creative and informed exploration of complex floor layouts. In a different study, Zorn et al. (2022) demonstrated the potential of surrogate models (SM) as a novel approach for early-stage design evaluation. Additionally, the authors integrated the SMs into a dashboard presenting several performance indicators, which update in real-time. Westermann et al. (2020) propose a platform (Net-Zero Navigator) for exploring building performance design at the conceptual stage, powered by highly accurate surrogate models for energy that are trained on a dataset of 16 building types over 30 parameters and 20 different climates. The system is available online and features data visualisations, such as parallel coordinates plots, that can filter amongst existing designs or quickly sample designs from select parametric ranges.

The approaches mentioned above focus on surrogate models for either daylighting or energy; furthermore they do not present an interface for interactive design exploration (Yousif and Bolojan, 2022), present a data-only dashboard that does not show design forms (Zorn et al., 2022, Westermann et al., 2020), or an interactive design exploration interface that lacks features of creativity-support (Westermann et al., 2020), as outlined by Shneiderman (2007). Our work is distinguished by attention to daylighting and energy usage as interrelated performance factors. Furthermore, we present a feature-rich prototype that enables generating, evaluating and comparing designs quickly with the support of surrogate models and an emphasis on visualising both design forms and their performance to designers to allow holistic assessment.

2.2. ENHANCING DESIGN BY EMPLOYING DESIGN ANALYTICS

Integrating GD and BPA poses challenges in data management and design decision-making. With the GD approach, architects produce numerous design solutions, and BPA adds layers of information. Consequently, interpreting extensive data to identify the most suitable design alternatives becomes challenging, often leading to an overload of choices (Erhan et al., 2017). In contrast, evidence shows that, in conventional design processes, designers can effectively manage multiple alternatives in large office walls (Woodbury 2010).

This issue requires practical data analysis and decision-making strategies. Architects can use data analysis techniques and interfaces to navigate the design space effectively. These systems facilitate the exploration, analysis, and interpretation of large data sets through data visualisation and analysis, empowering architects to make informed decisions. It’s common to include interactive data visualisations such as parallel coordinates plots (e.g., in Design Explorer 2, 2023 and D.Star), bar charts (e.g., in DANZ by (Garg 2019), and scatterplots (e.g., DreamLens, 2018), which can be used to filter through design alternatives. Accompanying data visualisations are often graphical representations of designs (e.g., a single 2D image, a 3D view, or interactive image galleries), which, combined with visualisations, can help designers judge designs qualitatively and quantitatively.

Supportive techniques like rating, clustering, grouping and annotating designs have also been incorporated into design navigation interfaces. Including those features shows support from studies on designers' behaviour when exploring design spaces under cognitively overloading conditions (Shireen et al., 2017). A longer survey and analysis of design space navigation interfaces can be found in other sources (Abu

Zuraiq, 2020). In addition to the above features, sensitivity analysis (Østergård et al., 2017), clustering and Pareto Frontiers (Brown and Mueller, 2017) can be integrated into design space exploration. Visualising already generated designs helps understand relationships in the design space and filter designs to a few choices. But often, it's also necessary to generate new designs as designers gain new insights about the design problem and their criteria mature. Systems like D-Star (Mohiuddin et al., 2018), the Navigator (Garcia and Leitão, 2022), and the Net-Zero Navigator (Westermann et al., 2020) enable their users to create new alternatives on demand. Finally, each new design project may require bespoke needs in terms of design models or performance analysis. Ritter et al. (2015) introduce a system that connects a parametric modeller like Dynamo to building simulations so that designers can flexibly adapt the parametric model based on their unique needs for each new design task.

3. Developing D-PREDICT: Methods

We employed the design study approach as a problem-oriented research methodology. It involves "analysis of a specific real-world problem faced by domain experts, designing a visualisation system that supports solving this problem, validating the design, and reflecting about lessons learned to refine visualisation design guidelines." (Sedlmair et al. 2012). Below, we summarised the initial high-level requirements for a design analytics system combining performance evaluation with GD, which are derived from a literature review:

R1. Support exploration generated design alternatives (Shneiderman 2007).

R2. Facilitate user-friendly, engaging interactions with design models and surrogate models for performance analysis, catering for designers with varying levels of expertise (Shneiderman 2007).

R3. Guide selecting parameters and make their sensitivity on design generation transparent (Hamby 1994; Bernal et al. 2020).

R4. Support collecting and retrieving design (Shneiderman 2007).

R5. Keep a history of choices within the GD (Shneiderman 2007).

R6. Support the different design exploration methods and styles ("wide walls") while providing advanced data visualisation techniques for those who can use them (Shneiderman 2007; Abu Zuraiq 2020).

We present D-Predict as a low-fidelity design analytics tool for design decision-making. D-Predict can be classified as a Creative Support Tool (CST) (Shniderman, 2007), offering design generation, exploration, and comparison capabilities. D-Predict is developed as a prototype to provide insights into the GD, analysis of alternatives using surrogate models, and their comparative evaluation. The prototype serves as an integral part of the design exploration process. D-Predict can connect with surrogate model-driven performance prediction systems. Its prototype includes functionalities that demonstrate possible solutions rather than being a usable system.

4. D-Predict System Design

D-Predict adopts a specify-generate-evaluate cycle in GD (Figure 1). Central to the workflow, design analytics interfaces combine GD and performance prediction using data visualisations. The building performance prediction layer generates data by selecting surrogate models for performance concerns. After merging data from GD and performance prediction, designers explore the potential of each solution or set of solutions using coordinated views.

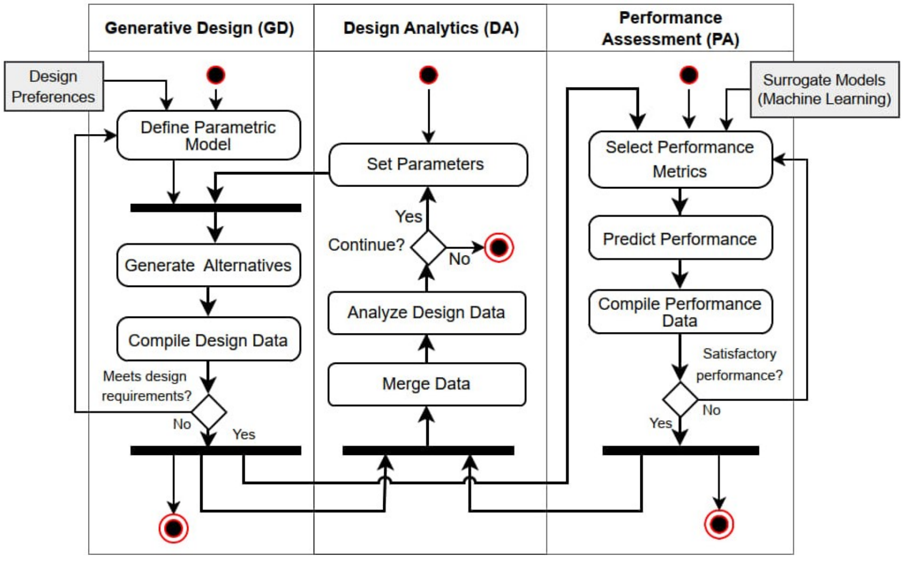


Figure 1. The proposed workflow integrates Design Analytics interfaces as a central data flow control between GD and performance prediction. The design alternatives are evaluated through such interfaces for their potential; then, they can be used for performance prediction. A design alternative can be directly used for performance prediction and evaluation. Therefore, the flow makes the design decision-making as flexible as possible.

4.1. D-PREDICT INTERFACES

D-Predict interfaces comprise three primary views: setup, data generation, and comparison views (Figure 2). In the Setup View, architects can establish a modelling environment in D-Predict by connecting it to a parametric model in a building modelling system, such as Grasshopper. This feature enables architects to leverage their parameters, previously crafted in the source software, and adjust parameters.

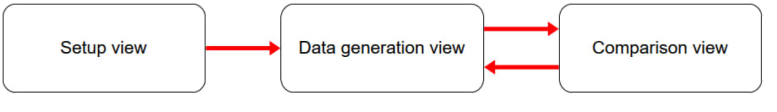


Figure 2. Three main views of D-Predict: The setup view for linking parametric from a model to initiate the interaction in the generate-predict-evaluate cycle.

4.2. DESIGN GENERATION VIEW

The design generation view is composed of two dynamic sidebars. The left sidebar is dedicated to architectural modelling, allowing designers to adjust parameters (R1) (Figure 3A). On the right one, the surrogate models are selected for performance prediction (Figure 3B). In the same panel, the repository panel stores visited alternatives (Figure 3E). Two vertical panels in the centre feature a parallel coordinate chart (Figure 3C) displaying the parameters and performance data. Simultaneously, the lower panel includes the 3D model, akin to the views in 3D software (Figure 3D).

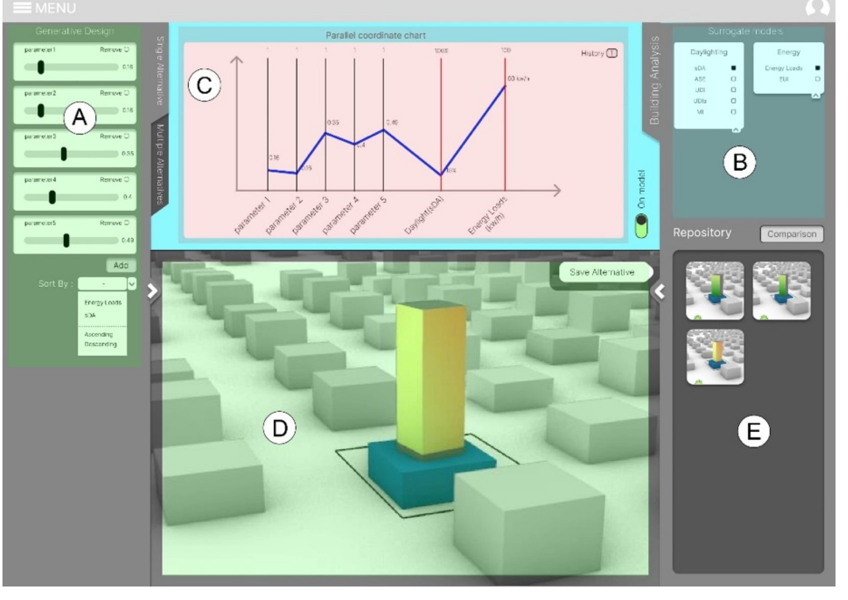


Figure 3. The generation comprises five views: (A) GD view to create multiple alternatives; (B) Surrogate Models to define performance metrics and select surrogate models; (C) Parallel coordinates displaying input and output parameters; (D) 3D View providing a visual representation of the form generated and used for performance analysis; (E) Repository view serving as a visualisation of visited alternatives and allowing persistent storage and retrieval on-demand.

The parameters from the CAD modeller are viewed to enable the generation of alternatives (R1). The users can add or remove parameters and modify the building configuration and material by adjusting the parameters. Once the form is defined, the users can choose one or multiple surrogate models to perform performance prediction, in which the results are viewed on the 2D parallel coordinate and the 3D model view (R2). The surrogate models' panel is designed explicitly for daylighting and energy load metrics, encompassing sDA (spatial daylight autonomy), ASE (annual sunlight exposure), UDI (useful daylight illuminance), and MI (mean illuminance) for daylight, as well as EUI (energy use intensity) and energy load for energy analysis. The repository enables storing preferred designs based on architectural concerns.

4.2.1. History tracking

One of the main features of D-Predict is history tracking (R5), allowing designers to monitor their past decisions. On the top-right corner of the parallel coordinate chart is a space for specifying the number of previous actions the designer wishes to revisit. By defining this number, the designer can see and retrieve the previous values for each parameter either on the parameter slider or the chart. (Figure 4).

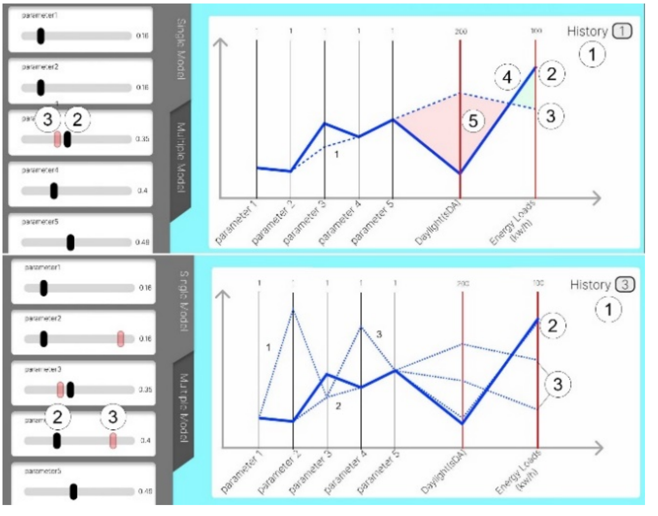


Figure 4. History tracking functionality allows monitoring and retrieving previous actions. (1) Defines the number of previous actions. (2) Displays the current state. (3) Displays the previous states. (4) The green area depicts an increase in value compared to the previous alternative. (5) The red area depicts a decrease in value compared to the previous alternative.

4.2.2. Alternative generation options

To generate alternatives, we have proposed three methods (Figure 5): single alternative generation by adjusting parameter values individually, multiple alternative generation by setting multiple discrete values for each parameter, and multiple alternative generation by value ranges. Subsequently, design forms are generated and used for performance prediction. The potential solutions can be stored in the repository for further investigation or revisiting later.

4.2.3. Sensitivity analysis

The impact of different parameters on each performance metric varies. Sensitivity analysis helps designers predict which parameters enhance or decrease performance more significantly. D-Predict offers a feature to view the impact of each parameter in two ways (R3). First, designers can rank the parameters based on their influence on a specific metric. This shows the relative importance of each parameter in achieving an expected outcome. The second visualises the results of a specific building performance metric, considering different values of one parameter while keeping the values of other parameters unchanged (Figure 6).

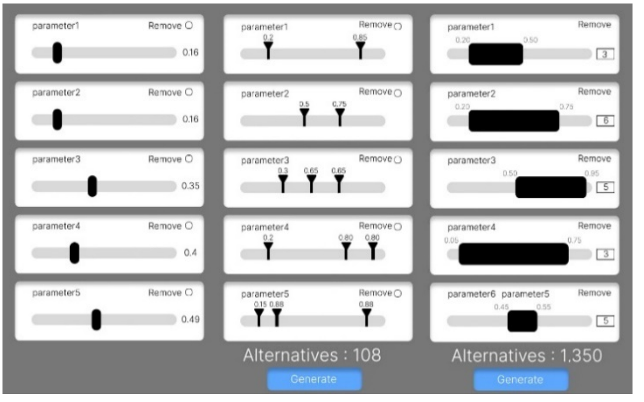


Figure 5. Three different methods to generate design alternatives: (left) single alternative by values, (middle) multiple alternatives by discrete values, and (right) multiple alternatives by value range.

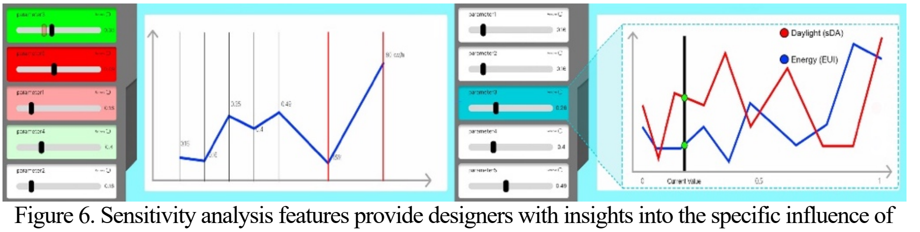


Figure 6. Sensitivity analysis features provide designers with insights into the specific influence of design parameters on performance metrics in two ways: (left) sorting the parameters based on their impacts and (right) displaying the effect of each parameter on overall performance.

4.2.4. Comparison View

The repository allows reexploring visited alternatives. By applying specific filtering techniques (Figure 7), designers can navigate a large set of alternatives and sort them to find potential designs by narrowing the search space. In this view, designers can create multiple subsets within the repository.

5. Conclusion

D-Predict represents an example towards making progress in sustainable architectural design by integrating GD, performance prediction, and interactive data visualisation. The discussion concerns overcoming computational challenges by adopting surrogate models, providing a quick alternative to laborious performance simulations. This approach aims to accelerate the design space exploration process and enrich potential by introducing additional layers of building information. D-Predict's interfaces, particularly the Comparison view, address the complexity of handling extensive data, offering architects mechanisms to navigate the design space and make data-informed decisions. The design study methodology ensures building a practical tool shaped by the insights and needs of design practitioners. As a part of this study, we identified high-level requirements for such systems.

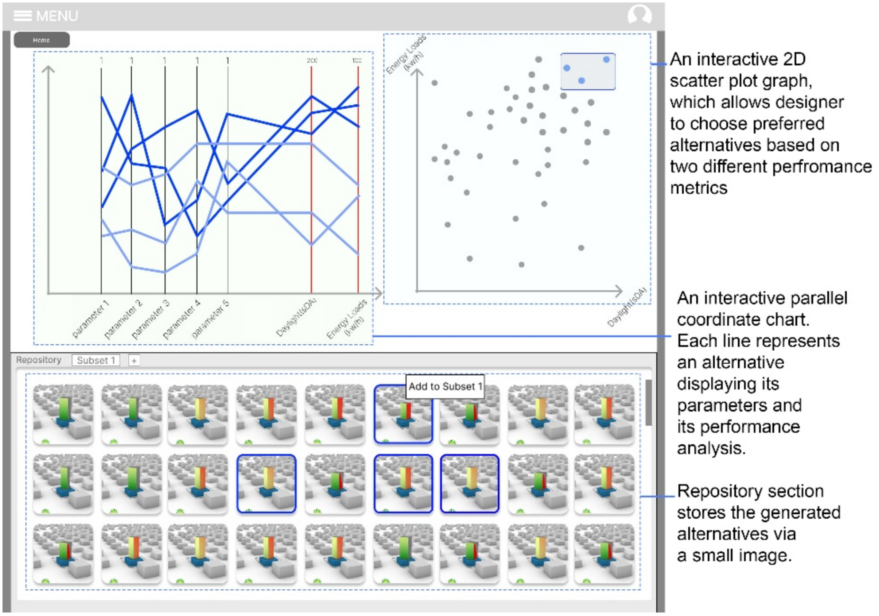


Figure 7. The comparison view enables evaluating sets of the generated design alternatives.

Based on the first version of the prototype, we developed a partial implementation of D-Predict as a Rhino plug-in. It integrates parametric design with machine learning-based analysis of daylight and energy load assessment (Figure 8). Our ongoing efforts involve conducting a user study with domain experts to evaluate the utility and adaptability of the system in practice. The Initial feedback has been positive and motivating. We will publish the updated system and the user study at a future venue.

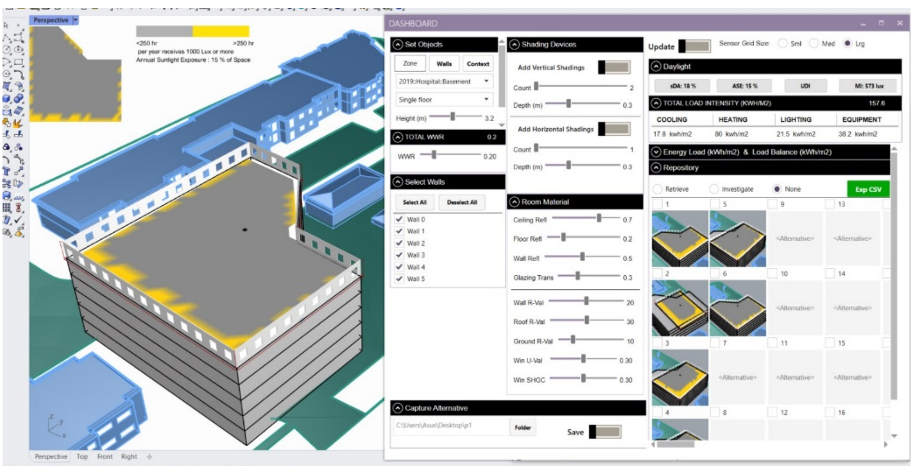


Figure 8. D-Predict's partial implementation on Rhino.

References

Abu Zuraiq, A. (2020). DesignSense: A Visual Analytics Interface for Navigating Generated Design Spaces. Retrieved December 13, 2023, from https://summit.sfu.ca/item/34308

Anton, I., & Tănase, D. (2016). Informed Geometries: Parametric Modelling and Energy Analysis in Early Stages of Design. Energy Procedia, 85.

Bernal, M., Marshall, T., Okhoya, V., Chen, C., & Haymaker, J. (2019). Parametric Analysis versus Intuition - Assessment of the Effectiveness of Design Expertise.

Brown, N., & Mueller, C. (2017). Designing with Data: Moving beyond the design space catalog. Proceedings Catalog of the 37th Annual ACADIA 2017 (pp. 154–163).

Chaszar, A., von Buelow, P., & Turrin, M. (2016). Multivariate interactive visualization of data in generative design. Proc. of SimAUD, 16.

Erhan, H., Abuzuraiq, A., Zarei, M., AlSalman, O., Woodbury, R., & Dill, J. (2020). What Do Design Data Say About Your Model? - A Case Study on Reliability and Validity. Proceedings of the 25th CAADRIA Conference. pp. 557-567

Erhan, H., Chan, J., Fung, G., Shireen, N., & Wang, I. (2017). Understanding Cognitive Overload in Generative Design - An Epistemic Action Analysis. Protocols, flows and glitches. p 127Protocols, Flows, and Glitches, CAADRIA 2017.

Garcia, S., & Leitão, A. (2022). Navigating Design Spaces: Finding Designs, Design Collections, and Design Subspaces. Int. J. of Architectural Computing, 20(2), 176–195.

Garg, A., & Erhan, H. (2019). Use of Data in Design Exploration: Design Analyzer. In J.-H. Lee (Ed.), Hello, Culture! (pp. 399–411). CUMINCAD.

Hamby, D. M. (1994). A Review of Techniques for Parameter Sensitivity Analysis of Environmental Models. Environmental Monitoring and Assessment, 32(2), 135–154.

Iyengar, K. (2015). Sustainable Architectural Design: An Overview.

Matejka, J., Glueck, M., Bradner, E., Hashemi, A., Grossman, T., & Fitzmaurice, G. (2018). Dream Lens: Exploration and Visualization of Large-Scale Generative Design Datasets. In Proceedings of the Conf. on Human Factors in Comp. Sys. (pp. 1–12). ACM.

Mohiuddin, A., Ashtari, N., & Woodbury, R. (2018). Directly Interactive Design Gallery Systems: Interaction Terms and Concepts. Comp. on Cultural V. and Heredity (pp. 15–32)

Østergård, T., Jensen, R. L., & Maagaard, S. E. (2017). Early Building Design: Informed Decision-Making by Exploring Multidimensional Design Space Using Sensitivity Analysis. Energy and Buildings, 142, 8–22.

Ritter, F., Geyer, P., & Borrmann, A. (2015). Simulation-based Decision-making in Early Design Stages. In Proceedings of the 32nd International Conference of CIB W78, Eindhoven, The Netherlands, 27-29 October (pp. 657-666).

Sedlmair, M., Meyer, M., & Munzner, T. (2012). Design Study Methodology: Reflections from the Trenches and the Stacks. IEEE T. on Vis. and Comp. Grph., 18(12), 2431–2440.

Shireen, N., Erhan, H., Woodbury, R., & Wang, I. (2017). Making Sense of Design Space - What Designers Do with Large Numbers of Alternatives?, 17th Int. CAAD Futures 2017.

Shneiderman, B. (2007). Creativity Support Tools - Accelerating Discovery and Innovation. Communications of the ACM, 50, 20–32.

CORE Studio. (2023). Design Explorer 2. https://tt-acm.github.io/DesignExplorer/

Westermann, P., & Evins, R. (2019). Surrogate Modelling for Sustainable Building Design – A Review. Energy and Buildings, 198, 170–186.

Westermann, P., Rulff, D., Cant, K., Faure, G., & Evins, R. (2020). Net-Zero Navigator: A Platform for Interactive Net-Zero Building Design Using Surrogate Modelling.

Woodbury, R. (2010). Elements of Parametric Design. London; New York, NY: Routledge.

Yousif, S., & Bolojan, D. (2022). Deep Learning-Based Surrogate Modeling for Performance-Driven Generative Design Systems. In Proc of the 27th CAADRIA (pp. 363–372).

Zorn, M., et al. (2022). Replacing energy simulations with surrogate models for design space exploration. Bauphysik, 44(6), 311–316.

DESIGN FACTOR-ORIENTED LIFE-CYCLE ENERGY OPTIMISATION FOR EARLY-STAGE BUILDING DESIGN

YANG YANG[1], LIKAI WANG[2] and MARCO CIMILLO[3]
[1,2,3]*Xi'an Jiaotong-Liverpool University.*
[1]*yang.yang2104@student.xjtlu.edu.cn, 0009-0007-8659-1831*
[2]*wang.likai@outlook.com, 0000-0003-4054-649X*
[3]*marco.cimillo@xjtlu.edu.cn, 0000-0002-6260-4470*

Abstract. The need for life-cycle energy (LCE) optimisation is imperative in the building sector. Computational optimisation has been adopted to evolve design populations search for desirable solutions with competitive LCE performance. However, prior studies typically focused on a predefined building form or typologies, which is less effective in assisting designers in space design exploration and informed decision-making. To remedy this gap, this study introduces a design factor-oriented LCE optimisation workflow integrating EvoMass and ClimateStudio within the Rhino-Grasshopper environment, with further steps of taking Window-to-Wall Ratios (WWRs) and multiple thermal zones into consideration. Through a case study, the results of the optimisation demonstrate the efficacy of the approach, revealing a considerable reduction of total energy needs. This study underscores the potential for energy savings through careful consideration of building massing, WWR, and multi-thermal zones in the LCE optimisation process when using a computational process, which also provides useful information for designers' decision-making at the early design stage.

Keywords. Building Massing, Multi-Thermal Zones, Window-to-Wall Ratio, Design Exploration, Life-Cycle Energy, Design Optimisation, Life-Cycle Analysis

1. Introduction

The construction and operation of buildings are significant contributors to global carbon emissions, accounting for approximately 37% of the total (United Nations Environment Programme, 2021). To address this pressing environmental concern, life-cycle energy (LCE) has been widely accepted as an effective approach to formulating strategies to reduce primary energy needs in buildings (Ramesh et al., 2010). LCE accounts for all energy inputs to a building in its life cycle, including operational energy (OE) and embodied energy (EE).

Energy efficiency measures (EEMs) are intended to improve the building's performance by changing critical building design factors (building forms, orientation

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 465-474.

and window-to-wall ratios/WWRs), material types, material quantities and so on (Shadram & Mukkavaara, 2019). As substantial design modifications are more viable and impactful for LCE at early design stages (Yu et al., 2022), researchers are focusing more and more on optimising the impact of EEMs in LCE at early design stages. For instance, Vollmer et al. (2023) conducted a lifecycle-based parametric optimisation, considering building materials, energy systems and renewable energy systems as EEMs. Additionally, Shadram & Mukkavaara (2018) studied an integrated BIM-based framework for the optimisation of the trade-off between EE and OE for early design, considering material types and quantities.

Design factors should not be neglected as the first set of EEMs that can be implemented or changed early in the design phase (Shadram & Mukkavaara, 2019), but they were typically not considered in aforementioned studies. Thus, Shadram & Mukkavaara further studied building forms based on six different extrusions, revealing significant variations influenced by building forms in LCE pre-optimisation but modest differences in post-optimisation (Shadram & Mukkavaara, 2019). Other studies also focused on the life-cycle analysis (LCA) concerning the design factors of buildings. For instance, Jusselme et al. (2018) conducted a sensitivity analysis of LCA-based optimisation, which includes one archetypal geometry. Harter et al. (2020) investigated uncertainty analysis of LCE in early design, incorporating design factors (seven fixed extrusions and their orientation). In Yu et al.'s study (2022), the geometric variability is considered a critical point for LCE optimisation. The study tested an extrusion with partial variability on one side of the building using Design Explorer, conducting a case study on buildings with a single thermal zone. The aforementioned studies indicated that the design factor is one of the unignorable EEMs for LCE or LCA optimisation.

However, prior studies on design optimisation focusing on LCE typically adopted simple geometrical operations, such as extrusion, instead of more articulated variations of the spatial configuration of the building massing (Figure 1). At the same time, WWRs with single thermal zone were often widely used to simplify the optimisation task, which leads to inaccurate predictions of thermal loads. These limitations hinder architecture designers from fully exploring building typologies, implementing more accurate WWRs design optimisation based on multiple thermal zones.

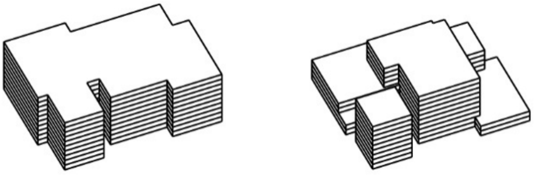


Figure 1: a) Building extrusion b) The spatial configuration of the building massing

To address these limitations, this paper proposes an LCE optimisation workflow focusing on early-stage design exploration by combining two tools (EvoMass and ClimateStudio) and taking WWRs for multi-thermal zones into consideration. The workflow includes the design factors of building massing, WWRs, and multi-thermal zones into the optimisation in order to enlarge the design space of the optimisation search so as to avoid the bias caused by the predetermined building forms and WWRs.

Thus, the study investigates the feasibility of design factors-oriented parametric optimisation and assesses its potential in terms of embodied, operational and total life-cycle energy through a case study. The case study highlights how the proposed design optimisation workflow enables architects to identify promising design directions and obtain information about the trade-offs and correlation between building massing forms, WWR and LCE needs.

2. Method

The study proposes a design optimisation workflow that combines EvoMass and ClimateStudio to achieve an automatic design generation, evaluation, and optimisation process. There are four steps as depicted in Figure 2: 1) Setting up the massing generation in EvoMass; 2) Defining thermal zones and corresponding WWRs based on the building massing generated by EvoMass; 3) LCE calculation and design evaluation using ClimateStudio; 4) Evolutionary optimisation based on Steady-stage Island Evolutionary Algorithm (SSIEA) in EvoMass.

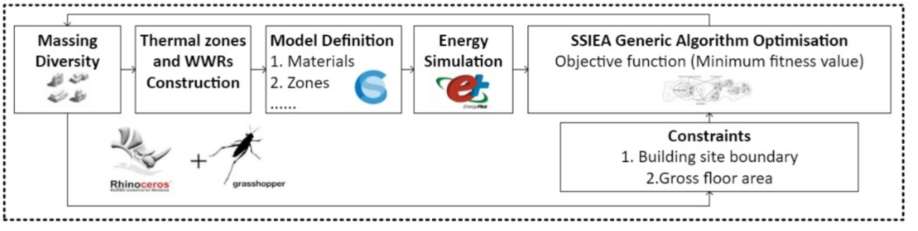


Figure 2 Workflow of the design optimisation.

In the first step, the massing generation is configured in EvoMass, an integrated evolutionary building massing design tool facilitating the automatic generation of optimised design iterations for architecture designers (Wang, 2022).

Subsequently, the generated massing is automatically segmented into multiple thermal zones using the Boolean function within a cutter massing approach. At the same time, the orientations of windows are specified and categorised within individual thermal zones. This step is aimed at avoiding the issue caused by single-zoning approaches, such as unbalanced WWRs on different building façade surfaces and inaccurate predictions of thermal loads since a surplus of solar gains in one zone may be credited to the heating or cooling required in another.

In the third step, the energy model is built using ClimateStudio (ClimateStudio, 2023), a Grasshopper plug-in for building EE and OE models. ClimateStudio is built upon validated simulation engines, EnergyPlus (Testing and Validation, 2014) and Radiance (Gregory, 2019). The thermal zones, construction elements, orientation, WWR values and other design parameters subject to LCE optimisation are defined in ClimateStudio for the initial designs of the buildings. The output of EnergyPlus on the annual energy use is used to obtain the OE.

In the final step, SSIEA is used to evolve the design population. SSIEA adopts a multi-island approach, subdividing the design population into several subpopulations. This approach guides each subpopulation to focus on different regions in the design space, preventing optimisation from being confined to a single region (Wang et al., 2020). Thus, SSIEA is particularly adept at providing diverse solutions for designers at the initial design stage.

3. Case study and results

3.1. DESIGN SETTING AND CONDITION

The selected case for this study is a five-story educational building within Xi'an Jiaotong-Liverpool University in Suzhou, China, which contains offices, lecture halls, studios, and fab labs. The gross floor area of this building is 15000m². Figure 3 provides detailed information about the building. Suzhou is characterised as a hot-summer-cold-winter (HSCW) climate whose data are obtained from standard EnergyPlus Weather files (.EPW) via the Ladybug website (EPW Map, 2023).

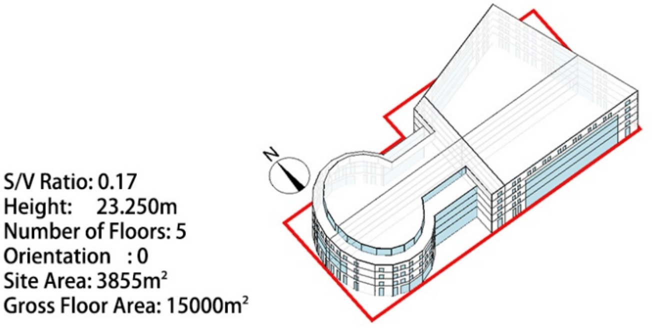


Figure 3 ClimateStudio energy model of the actual building

3.2. SCENARIO SETTING AND BENCHMARKS

To demonstrate the effectiveness of the proposed workflow, the automatic generative design is compared with following scenario-based benchmarks: the actual building (AB) and best international practice (BI).

The AB was defined according to the specifications outlined in the section of the HSCW zone in GB50189-2015 (MOHURD & AQSIQ, 2015), except WWRs. WWRs are defined as the real situation. The standard stipulates the baseline requirements for all commercial buildings in China (Table 1).

The high-performance benchmark was created to represent the Best International Practice (BI) based on the BREEAM In-Use International Technical Manual: Commercial SD243-V6.0.0 (BRE group, 2020). The Building Research Establishment's Environmental Assessment Method, or BREEAM, is one of the most comprehensive and widely used environmental assessment tools worldwide.

The scenarios of benchmarks AB and BI are summarised in Table 2. The values that meet benchmarks in the ClimateStudio database are written in parenthesis and applied in the simulation. Window distribution in AB and BI has been modified according to relevant standards, expressed as WWRs by GB50189-2015 and as window-to-floor ratio (WFR) by BREAAM. Glazing components have been selected to meet both U-value and SHGC requirements.

The building's other construction elements were chosen by considering designs suitable for the HSCW zone, following "NECB-2020 Non-residential School/university" (Table 3). While the same materials are applied across all scenarios, different insulation thicknesses are defined to alter the U-value.

Table 1 Parameters applied in ClimateStudio

Parameters	Values	Parameters	Values
Buildings’ lifespan	50 years	WWRs	0.4-0.7
Heating set point	20 °C	Mechanical ventilation	10 L/s/p
Cooling set point	26 °C	People	6 m² per person
Opening time	7:00~18:00	Infiltration	0.15 ACH

Table 2 Scenarios of benchmarks

Parameter	AB	BI
U-value of external wall	1 W/(m²·k)	0.625 (0.616) W/(m²·k)
EE of external wall	390.5 MJ/ m²	506 MJ/ m²
U-value of roof	0.7 W/(m²·k)	0.5 (0.492) W/(m²·k)
EE of roof	481 MJ/ m²	506 MJ/ m²
U-value of glazing	3.0 (2.69) W/(m²·k)	2.32 (1.8) W/(m²·k)
SHGC	0.44 (0.358)	0.5 (0.296)
Total WWRs	0.49	0.63
EE of glazing	431 MJ/ m²	427 MJ/ m²
Total WFRs	0.2	0.257

Table 3 Information of other construction sets

Construction	U-Value	Thermal Capacitance	EE
Partition	2.422 W/(m²·k)	118.6 kJ/k/m²	26.6 MJ/m²
Slab	2.273 W/(m²·k)	456 kJ/k/m²	360 MJ/m²
External floor	0.386 W/(m²·k)	293.566 kJ/k/m²	183.6 MJ/m²
Ground slab	0.703 W/(m²·k)	472.001 kJ/k/m²	441.095 MJ/m²
Ground wall	0.955 W/(m²·k)	464.131 kJ/k/m²	446.025 MJ/m²

3.3. OPTIMISATION SETTING

The following values are defined in EvoMass based on the site characteristics: 1) Five floors; 2) Column spacing of X=5.265, Y=5.15; 3) Unit mass of vertical size from 1 to 5; 4) Unit mass of horizontal size from1 to 18; 5) 15000 m² of target gross floor area (TGFA). The 'additive form generation' in EvoMass is tested for geometric exploration. Figure 4 illustrates four samples of the EvoMass generated massing and WWRs with four thermal zones aligned with the settings in this paper.

The generated massing design is segmented into four thermal zones—north-east, north-west, south-east, and west-east—via a cutting volume created by connecting centre points of the site boundary perpendicularly. Individual WWRs are defined for each façade surface based on its orientation, i.e. the north, south, west, and east.

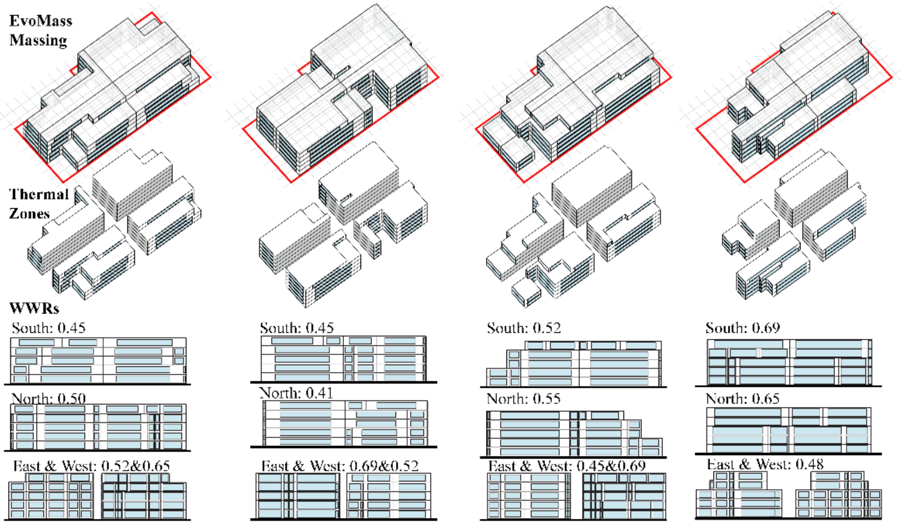


Figure 4: EvoMass generated massing and WWRs with thermal zones aligned with the setting

In order to mitigate the unbalance between lower EE and higher OE in total energy (TE) and to equally weigh the EE and OE, the fitness is calculated based on the multiplication of EE/m² and OE/m². The optimisation objective is to minimise the fitness value, with a TGFA serving as a penalty function for excluding designs failing to satisfy the GFA requirements. The fitness function is defined as follows:

$$fitness = \left(1 + \left|\frac{GFA - TGFA}{TGFA}\right|\right) \times \frac{\mathrm{EE/m^2 \times OE/m^2}}{1000} \tag{1}$$

In Equation 1:

EE/m² (MJ) is the embodied energy per square meter.

OE/m² (MJ) is the operational energy per square meter.

GFA (m²) is the gross floor area.

TGFA (m²) is the target gross floor area.

OE and EE are equalised to reduce the bias toward OE in this particular case study, which does not necessarily provide a representative ratio between the two values. A number of simplifications in the model are, in fact, likely to produce an overestimation of the OE and an underestimation of the EE. These include the removal of the internal components and the adoption of uniform occupancy and ventilation rates for the whole building.

3.4. RESULTS AND COMPARISON STUDY

Simulations and optimisations are conducted on a computer with a 2.7 GHz Intel Core CPU and 8GB RAM, running Microsoft Windows 10 Enterprise LTSC as the operating system. Each simulation run takes 1 minute on average. In order to eradicate the influence of GFA on energy needs, the results with GFA (15000±750 m²) are selected in the following sections. There are 797 satisfying designs out of 1200 generated designs during the whole optimisation process (Figure 5).

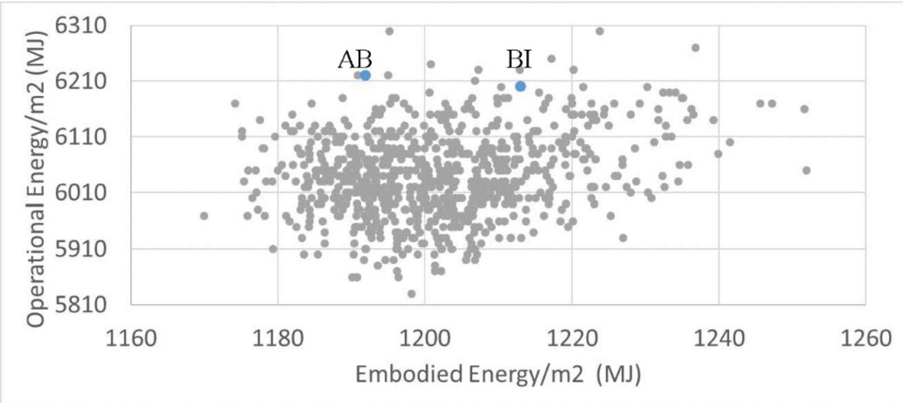


Figure 5 The results of optimised solutions and scenarios

3.4.1. General energy saving

The primary energy of EE/m² and OE/m² for AB and BI is simulated, which reveals that BI and AB consume similar TE/m². The optimisation is conducted based on the BI scenario setting in Section 3.2. The comparison in Figure 6 shows a potential for LCE saving by design factor-oriented optimisation, particularly for OE. The optimal solution accounts for 7.4% in OE/m² and 6.6% in TE/m², when it is compared to the highest energy-need solution. Compared to BI, the energy use decreases by up to around 5.5% for TE/m² and 6% for OE/m².

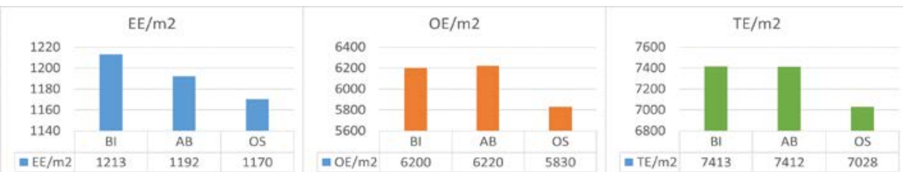


Figure 6 Comparison of best-optimised solutions and benchmarks

3.4.2. Building massing characteristics

The fifteen highest-ranking elites are selected from optimal solutions and grouped into three categories (C1-Maximum Fitness Percentage, C2-Minimum OE, C3-Minimum Surface/Volume Ratio). Figure 7 summarises their performance indicators, while Figure 8 provides the average of each indicator of the elites from C1, C2, and C3.

C1 elites achieve the best average fitness among these three optimisation processes. In terms of massing, the C1 elites showcase a tendency toward a rectangular-extrusion volume elongating along the east-west direction, favouring the reduction of OE/m². In addition, as OE/m² accounts for major energy needs in LCE, reduction of OE can achieve fewer energy needs, even taking the risk of the trade-off with EE/m².

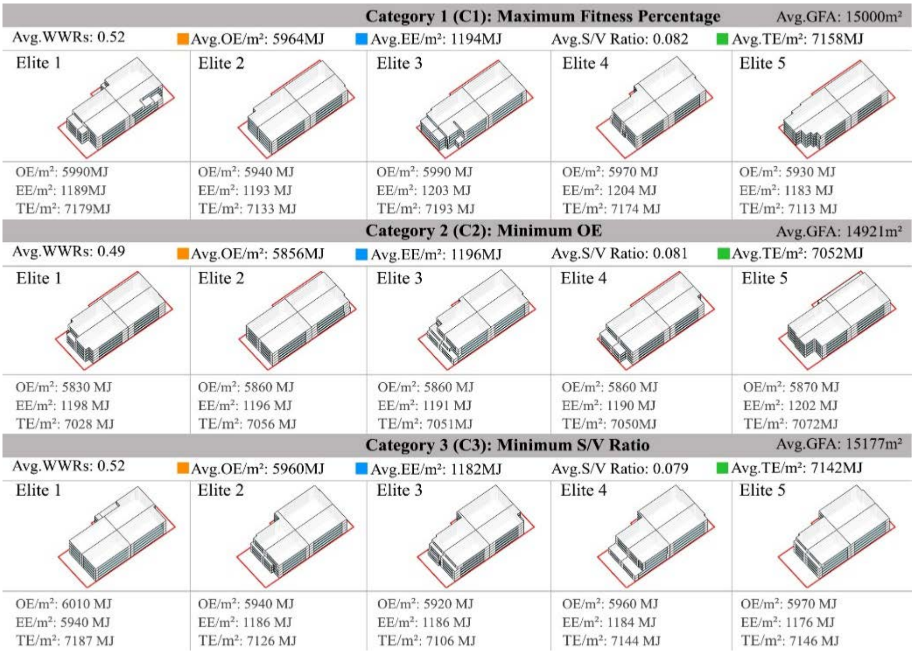


Figure 7: Fifteen highest-ranking design elites from the results of the case study in three categories

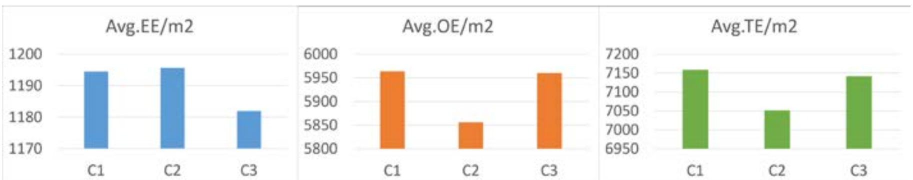


Figure 8: Average EE, OE, WWRs, TE/m2 of elite results of the case study in three categories

C2 elites share a similar north-south-side longer rectangular profile as C1 but with fewer step-backs or setbacks, achieving the lowest OE/m², highest EE/m², and least TE/m². As mentioned in C1 elites, the results of C2 elites reveal that the longer rectangular extrusion is beneficial for reducing OE/m². Moreover, it allows C2 elites to achieve a greater reduction of OE/m² by taking advantage of less S/V ratio and GFA.

On the other hand, the elites in C3, with the lowest S/V ratio, consume more energy than C1 and C2. Despite common recommendations for lower S/V ratios to reduce energy needs, C3 elites highlight the importance of considering the integral impact of various design factors (like WWRs), as solely focusing on minimising the S/V ratio may not lead to reduced energy needs. Thus, the C3 elites demonstrate the importance of considering multiple design factors for effective LCE optimisation.

3.4.3. Correlation Analysis

The impact of variables is illustrated by the correlation coefficients listed in Table 5.

The correlation has been calculated for EE/m², OE/m², and TE/m², assigning numerical values to all variables. The correlation analysis shows that the S/V ratio significantly influences TE/m², with a notable impact from north, south and west WWRs.

Table 5 Correlation analysis

	North WWR	South WWR	East WWR	West WWR	S/V Ratio
EE/m2	-0.09282	0.01233	-0.10603	0.019703	0.789915
OE/m2	0.479456	0.33908	0.329014	0.345862	0.525172
TE/m2	0.438372	0.323147	0.293695	0.33081	0.630219

4. Discussion and Conclusions

The outcomes of the case study underscore the considerable impact of different combinations of design factors (specifically, in this study, building massing and WWRs across multiple thermal zones) on LCE performance. In addition, the comparison study also highlights the necessity of the proposed automatic generative design optimisation workflow. The proposed design workflow should be considerably more accurate than conventional scenario workflows with a limited number of combinations modelled and iterated manually. Moreover, the workflow is able to help designers understand the design trade-offs between OE and EE, and the extracted design implications related to these two aspects can also allow designers to synthesise this information in the subsequent design ideation and development process.

While the preliminary results of optimisation in this study are promising, further research is needed, particularly in devising segmentation solutions for more intricate thermal zones. For instance, subdividing a large thermal zone into several logically generated parts based on the building's configuration could enhance the precision of the optimisation process. Moreover, the integration of additional factors into the optimisation workflow is imperative. Consideration of material properties, for instance, could offer a more holistic approach to optimisation. Additionally, the incorporation of constraints on objective functions, such as comfort hours and solar irradiation, is essential for a comprehensive evaluation. Although the single-objective optimisation in this study has yielded valuable results, delving into multi-objective optimisation (MOO) could be the subsequent step. This would enable a more thorough exploration of the impact of EEMs implemented during the design phase on the EE and OE trade-off, ultimately striving to minimise the building's LCE. Finally, due to the development of the study on net-zero buildings, the OE tends to account for less in LCE at some point when buildings are renovated into net-zero ones. Thus, the scope of OE before renovation and how EE makes up for it should be studied in future.

To conclude, this study proposes a design factor-oriented LCE optimisation workflow by integrating EvoMass and ClimateStudio to take into consideration building massing variations and WWRs for multi-thermal zones. The paper provides a better understanding of design factors' optimisation that may contribute to the reduction of energy needs and provide valuable information to support sustainable building design. Further research should consider more design factors, constraints, MOO and net-zero renovation to minimise a building's LCE holistically.

Acknowledgements

This study is supported by Xi'an Jiaotong - Liverpool University Research Development Fund (RDF-23-01-107).

References

BRE group. (2020). SD6063 BREEAM In-Use International Commercial Technical Manual.

ClimateStudio. (2023). Solemma. Retrieved November 13, 2023, from https://www.solemma.com/climatestudio/accuracy

EPW Map. (2023). Retrieved November 18, 2023, from https://www.ladybug.tools/epwmap/

Gregory, W. (2019, November 9). *The RADIANCE Lighting Simulation and Rendering System.* Radsite. https://www.radiance-online.org/archived/radsite/radiance/papers/sg94.1/accurate.html

Harter, H., Singh, M. M., Schneider-Marin, P., Lang, W., & Geyer, P. (2020). Uncertainty Analysis of Life Cycle Energy Assessment in Early Stages of Design. *Energy and Buildings*, 208, 109635. https://doi.org/10.1016/j.enbuild.2019.109635

Jusselme, T., Rey, E., & Andersen, M. (2018). An integrative approach for embodied energy: Towards an LCA -based data-driven design method. *Renewable and Sustainable Energy Reviews,* 88, 123–132. https://doi.org/10.1016/j.rser.2018.02.036

MOHURD, & AQSIQ. (2015). Design standard for energy efficiency of public buildings GB50189-2015.

Ramesh, T., Prakash, R., & Shukla, K. K. (2010). Life cycle energy analysis of buildings: An overview. *Energy and Buildings,* 42(10), 1592–1600. https://doi.org/10.1016/j.enbuild.2010.05.007

Shadram, F., & Mukkavaara, J. (2018). An integrated BIM-based framework for the optimization of the trade-off between embodied and operational energy. *Energy and Buildings,* 158, 1189–1205. https://doi.org/10.1016/j.enbuild.2017.11.017

Shadram, F., & Mukkavaara, J. (2019). Exploring the effects of several energy efficiency measures on the embodied/operational energy trade-off: A case study of swedish residential buildings. *Energy and Buildings,* 183, 283–296. https://doi.org/10.1016/j.enbuild.2018.11.026

Testing and Validation. (2014). EnergyPlus. Retrieved February 17, 2024, from https://energyplus.net/testing

United Nations Environment Programme. (2021). *2021 Global Status Report for Buildings and Construction: Towards a Zero-emission, Efficient and Resilient Buildings and Construction Sector.* Nairobi

Vollmer, M., Theilig, K., Takser, I., Reitberger, R., & Lang, W. (2023). Life cycle-based parametric optimization of buildings towards climate neutrality and its implications for environmental protection. *IOP Conference Series: Earth and Environmental Science,* 1196, 012050. https://doi.org/10.1088/1755-1315/1196/1/012050

Wang, L. (2022). Workflow for applying optimization-based design exploration to early-stage architectural design – Case study based on EvoMass. *International Journal of Architectural Computing*, 20(1), 41–60. https://doi.org/10.1177/14780771221082254

Wang, L., Janssen, P., & Ji, G. (2020). SSIEA: A hybrid evolutionary algorithm for supporting conceptual architectural design. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing,* 34(4), 458–476. https://doi.org/10.1017/S0890060420000281

Yu, H., Yang, W., Li, Q., & Li, J. (2022). Optimizing Buildings' Life Cycle Performance While Allowing Diversity in the Early Design Stage. *Sustainability,* 14(14), Article 14. https://doi.org/10.3390/su14148316

DESIGN OPTIMIZATION THROUGH CFD-ABMS INTEGRATION FOR CONTROLLING VIRUS SPREAD

CHAO-NIEN CHEN[1] and PEI-HSIEN HSU[2]
[1,2]*Graduate Institute of Architecture, National Yang Ming Chiao Tung University.*
[1]*eaglechen811221@arch.nycu.edu.tw, 0009-0002-9315-0202*
[2] *phsu@arch.nctu.edu.tw, 0009-0006-0112-7400*

Abstract. In the wake of the pandemic, the interplay between humans, architecture, and viruses has emerged as a critical area of study. Aerosol transmission particularly within indoor environments is a key factor contributing to infections. However, contemporary research frequently isolates individual strategies such as ventilation, crowd control, protective measures, and emergency protocols, often overlooking a more comprehensive perspective. To bridge this gap, we introduce an integrated design framework for parametric environment that synergizes Computational Fluid Dynamics (CFD) and Agent-based Modeling (ABM). Furthermore, compared to large-scale environments such as cities, there is also less understanding of infection control in interior spaces. This study employs simulations to track the movement of viral particles in airflow and model spatial occupancy by workflow, providing insights into both overall particle transmission trend and spatial occupancy trends. The integration of these dual aspects facilitates a thorough assessment of transmission dynamics. Our findings, based on this integrated approach, offer recommendations for optimizing circulation patterns and spatial zoning, aiming to provide adaptability and flexibility in architectural design.

Keywords. Computational Fluid Dynamics (CFD), Agent-based Modeling (ABM), Infection Control, Post-Pandemic Design, Design Integration, Design Optimization

1. Introduction

1.1. ARCHITECTURE IN THE POST-PANDEMIC ERA

In the current post-pandemic context, spatial usage patterns are undergoing transformations beyond the medical sector. This shift has sparked interdisciplinary discussions, including concepts for future hospitals and the development of rapidly deployable modular wards. Extensive research highlights the pivotal role of architectural configuration and materials in infection control (Zimring et al., 2013).

Contemporary approaches to architecture and infection control predominantly

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 475-484.

focus on isolated strategies like ventilation and crowd management, lacking a comprehensive view. While significant research has explored infection spread in large-scale spaces like cities and regions, most infections occur within buildings (Gomez et al., 2020). In contrast to contact transmission, the dispersion of aerosols has been confirmed as a key transmission route for Covid-19 (Zhang et al., 2020). Direct alterations in architectural design or reduced spatial occupancy for infection prevention are often impractical (Perry et al., 2022). Hence, proactive architectural planning, adaptations in space utilization, and flexible designs are vital (Shangi et al., 2020).

1.2. PARAMETRIC ENVIRONMENT

The analysis of architectural ventilation and human spatial usage has gained significance in the post-pandemic era due to increased awareness of infection control. In domains like ventilation and spatial usage, expertise in Computational Fluid Dynamics (CFD) is essential, usually necessitating the involvement of engineers or professionals skilled in CFD analysis. Quantitative studies and analyses can be performed via data-driven approaches based on Agent-based Modeling (ABM). In recent times, architectural design has evolved into an integrative task, requiring preliminary integration before consulting specialists, ensuring design decisions and receiving immediate feedback (Nguyen and Peter, 2020).

Currently, there is still a gap exists in existing software for integrating these professional insights (Gomez et al., 2021). The advantage of parametric environments such as Rhino Grasshopper (GH) lies in their open-ended architecture, enable comparison of results and design optimization by adjusting design parameters. Compared to other tools, architects find them more accessible for learning and application.

1.3. PROJECT GOAL

To enhance design efficiency by minimizing transitions between different software environments and various specialties, this research aims to establish a comprehensive design framework integrating CFD and ABM methods within a parametric environment. By layering CFD and ABM methodologies, the framework seeks to provide a holistic assessment of transmission trends within spaces. Designers can thus gain insights into the potential impact of design parameters, enabling optimization and informed recommendations (see Figure 1).

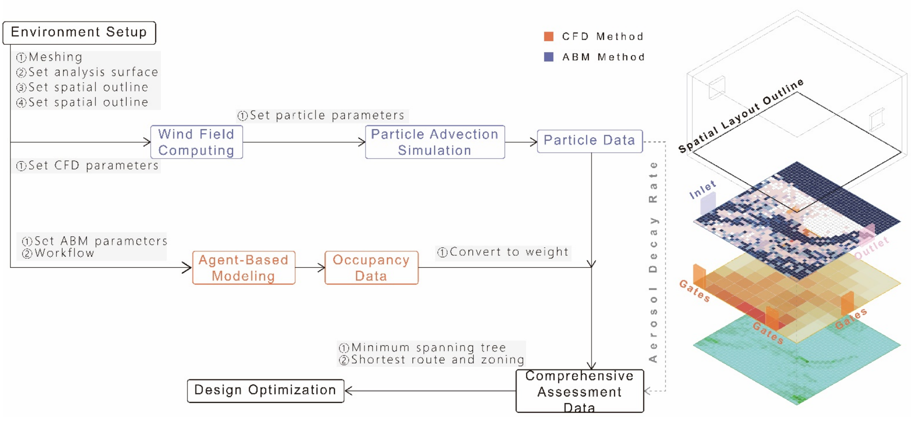


Figure 1. The flowchart and CFD-ABM design framework.

2. Methodologies

Through our literature review, it reveals that post-pandemic architectural design often employs CFD or ABM methodologies to investigate ventilation modes and assess infection risks related to varying behavioural patterns in spaces. In our proposed design framework, we integrate transmission trends of viral particles (SARS-CoV-2) in spatial environments with occupancy trends, overlaying these two layers while considering temporal factors.

In our model, the spatial plane is divided into a two-dimensional grid, representing the width and length of the space. The grid not only signifies the spatial continuity but also allows computational simplification based on the chosen grid size, thereby enhancing computational efficiency (Gomez et al., 2019). Numerical values in each grid cell correspond to computational model results. The chosen space is the ground floor of a campus building, specifically the institute hall, measuring 105.48m² with a height of 3.5m. The model, based on this space (see Figure 2), includes spatial environment parameters like meshing, spatial outline definition, analysis surface setting, and grid setting.

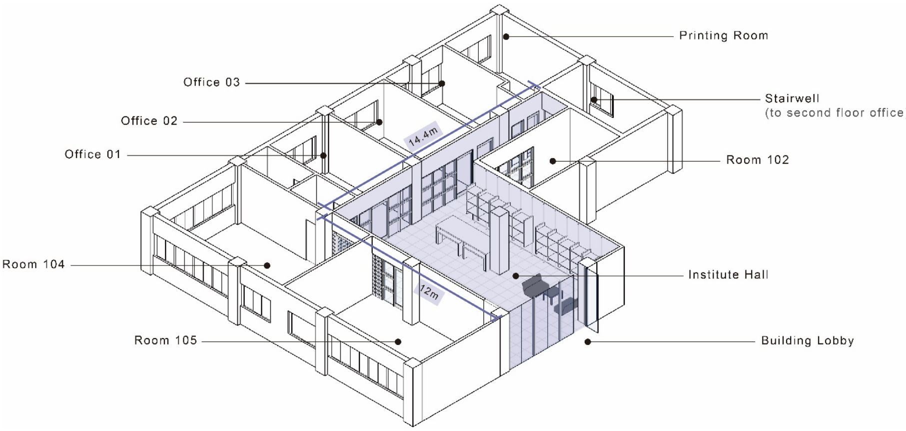


Figure 2. Localized area on the ground floor of the campus building. The shaded part represents the analysis area. The rightmost door connecting the institute hall and the building lobby remains consistently open. The windows in the stairwell and the door connecting it are also consistently open.

2.1. CFD METHOD

CFD, a numerical method for simulating fluid behaviours under specific conditions, offers valuable data and visualizations for designers (Hartog et al., 2000). We use the Butterfly plugin in GH for indoor airflow calculations, employing OpenFOAM for CFD simulations. The first layer of computation involves using this plugin for airflow data, supplemented by custom Python GH functions for particle transmission simulations, capturing the overall trend in the space.

2.1.1. Setting CFD Parameters

Airborne transmission risk is closely linked to ventilation rates, inadequate indoor ventilation can heighten infection risk (Ren et al., 2022). As there are currently no official standards for ventilation rates, we adopt the ideal value suggested by William Bahnfleth, setting the indoor ventilation rate at 6 air changes per hour. The HVAC supply air volume is calculated based on the ventilation rate equation:

$$Q=N*Vol$$

In this equation, where Q is airflow into the room per hour, N is the number of air changes, and Vol is the room volume. Natural ventilation, temperature, and humidity data are sourced from the Central Weather Administration Observation Service (see Table 1).

Ventilation Modes	HVAC Ventilation	Natural Ventilation
Number of Inlets	1	1
Inlet Size(m)	1 * 0.2	1.10 * 2.45
Number of Outlets	2	1
Outlet Size(m)	0.2 * 0.2	0.85 * 3.50
Supply Air Velocity(m/s)	-	1.9
Supply Air Velocity(m^3/s)	0.6	-
Temperature(°C)	24	21.7
Relative Humidity(%)	-	70

Data source from Central Weather Administration Observation Data Inquire Service, extracting data for the average values in November, wind direction(degree) is 60 and wall temperature is set to 18°C.

Table 1. Parameters for CFD simulation.

2.1.2. Particle Transmission Simulation

Building models are created with an analysis surface at 1.7m, representing breathing height. The simulation employs a Semi-Lagrangian method, following these steps (see Figure 3): (1) Using Butterfly GH for computations, we establish fixed wind field vectors. (2) Spreaders, symbolizing individuals, are uniformly placed, each with a 0.3m radius. A specific number of particles are randomly distributed within a 1.8m radius, representing the aerosol transmission range. (3) Particle advection determines their trajectories. (4) Accurate velocity vectors are obtained based on the horizontal component u and vertical component v of the wind field vectors in the surrounding four grids. (5) Particle trajectories update based on time step Δt. In this process, we assume that the fluid is incompressible and has no viscosity. The simulation shows that the space's airflow stabilizes within five minutes. The model, with 243 evenly distributed spreaders and inlet/outlet locations, simulates the overall particle transmission trend.

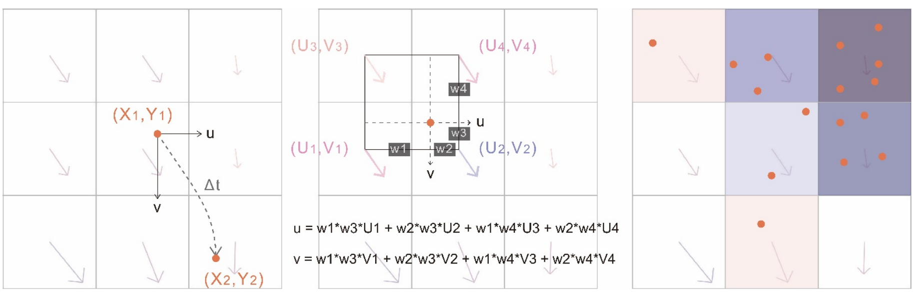


Figure 3. The light-coloured base map represents the wind field vector, with u and v as the horizontal and vertical velocity components of the particle. Particle positions are updated over Δt as time (left). w represents the weight, and the u, v velocity components of particles are computed by weighing against the adjacent four coloured wind field vectors (middle). After the simulation, the particle counts are recorded and visualized within the grid (right).

2.2. ABM METHOD

ABM, a computational model where each agent has unique states, behaviours, and decision-making processes, is widely used to simulate social phenomena in urban and architectural contexts (Batty, 2001). In GH, the Pedsim plugin simulates agent movement and behaviour. ABM in this study models pedestrian flow, forming the computational model's second layer and determining spatial occupancy trends.

2.2.1. Setting ABM Parameters

Agents' parameters include roles (students and office staff) and behaviours (staying and moving). During break times, students exhibit three behaviour modes: (1) Remain in the classroom. (2) Leave and return to the institute hall. (3) Visit a halfway point of interest in the institute hall before returning. Post-class, everyone exits the institute hall. Students without classes (student in room 104) and office staff can leave and return during breaks.

2.2.2. Spatial Occupancy

Workflow adjustments and comparisons aid in understanding flexible spatial usage patterns (Ortner and Tay, 2021). Our simulation uses existing workflow data to model spatial usage, involving: (1) Agents finding the shortest path to their destinations. (2) Agents deciding on additional behaviours en route. (3) Continued movement towards destinations. Given the substantial time intervals in common pedestrian density heatmaps, our ABM model divides time into 30-minute segments to examine spatial occupancy patterns within each time segment. This allows for a comparison with the earlier viral particle decay rates, with a focus on the 13:00 period for detailed assessment (see Figure 4).

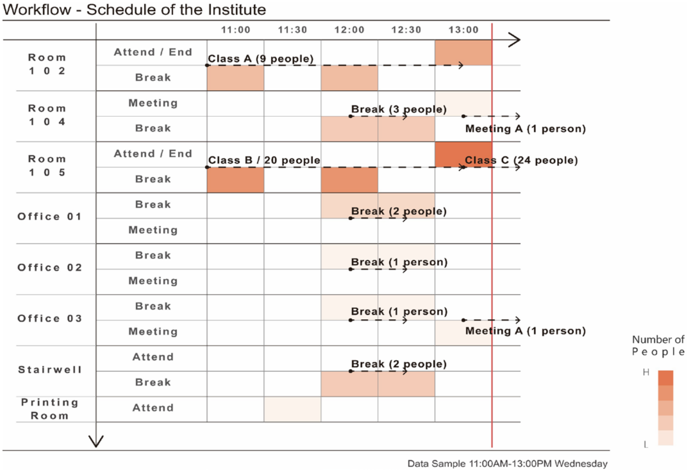


Figure 4. Spatial activity flow sample within the institute, where colours indicate whether agents engaged in entering or exiting the institute hall at that specific time point. The corresponding spaces can be cross-referenced with the earlier architectural model, with colour depth representing the number of agents. Dashed arrows represent the duration of activities, and the red line represents the time segment chosen for assessment in this study.

2.3. CFD-ABM METHOD

The third layer involves overlaying particle transmission and spatial occupancy trends (see Figure 5) to obtain a comprehensive transmission assessment. The framework includes: (1) Converting spatial occupancy data into weights, based on the number of times a single grid has been visited divided by the highest visited count. (2) Refining the grid for data alignment between layers (see Figure 6). (3) Adding the third dimension, time, to the two-dimensional grid, combining the previously influenced virus particles from earlier times, affected by decay rates, and allowing for the comprehensive assessment transmission trend.

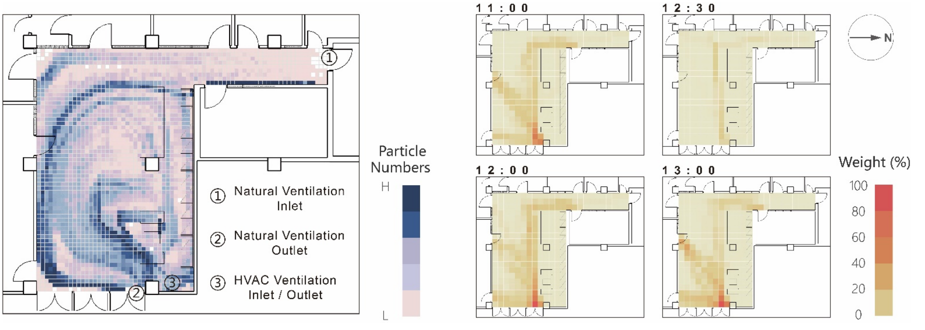


Figure 5. First Layer: Particle transmission trend (left). Second Layer: Spatial occupancy trends in different time segments (right).

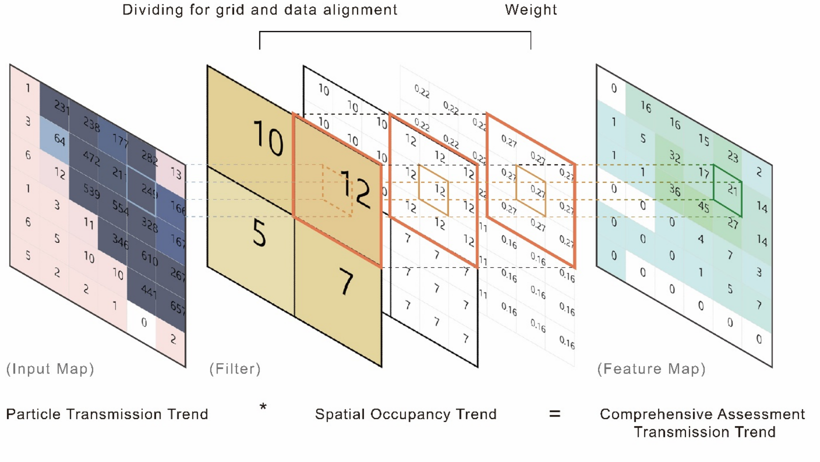


Figure 6. Refining the grid to align the data between the two layers, and obtaining the comprehensive assessment transmission trend by multiplying the particle transmission trend with the spatial occupancy trends transformed into weights.

3. Result

The overlaid layers reveal significant differences from individual layers (see Figure 7). This comprehensive assessment informs design strategies and optimizations. From the trend distribution, we employ the Minimum Spanning Tree (MST) algorithm to identify efficient paths within the space. The process includes: (1) Filtering the overlaid results based on assessed trend levels to form a Spanning Tree, with edges between points representing weights, the graph with the minimum total weight is the MST. (2) Introducing the starting and end point into the layout (spanning tree graph). (3) Edges between nodes, where edge weights can be set based on design considerations, such as distance, slope, lighting, etc. In this study, distance is used as the weight, and derives the most efficient paths from the MST graph, providing trend zoning and design recommendations after the assessment (see Figure 8).

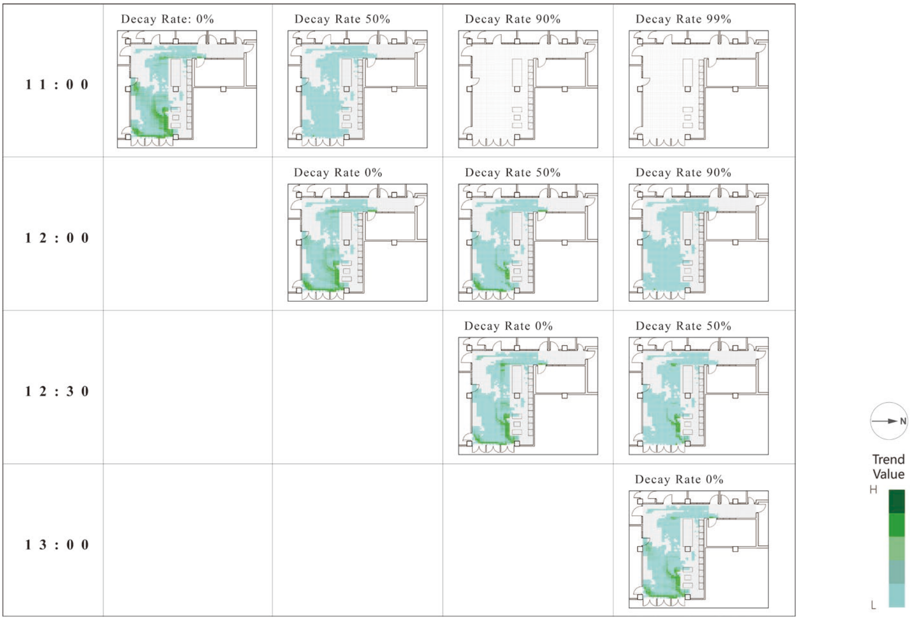


Figure 7. Third layer: Comprehensive assessment transmission trends. The bottom of each straight column, representing a decay rate of 0%, illustrates the current time segment's comprehensive assessment transmission trend, incorporating the influence of trends from previous time periods. The timeline in this figure can be cross-referenced with the earlier workflow figure's timeline.

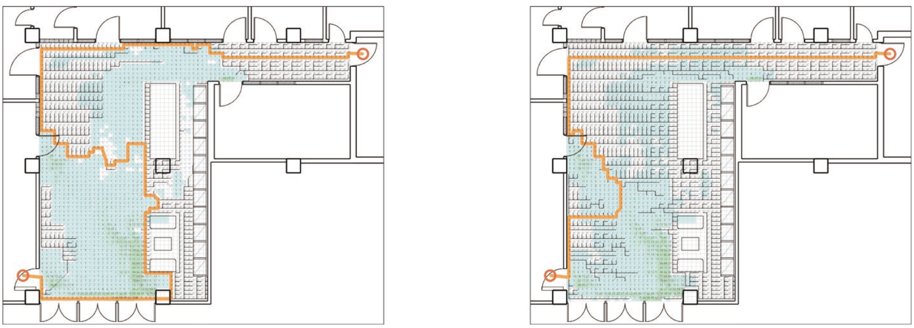


Figure 8. Regions with trend values of 0% (left) and 5% (right) were selected to generate a spanning tree graph. The thin black lines in the graph represent the MST, and recommendations for routes were generated based on specified points (the orange line and points).

4. Conclusion and Future Works

This study introduces a design framework suitable for parametric environments, integrating CFD and ABM methods to comprehend the combined impact of airflow and spatial occupancy on infection trends. CFD methods extend beyond fixed wind field calculations, using Python GH to apply wind field data in particle simulations. While guiding users along these routes may deviate from common sense, the design framework remains critical for early-stage decisions, activity zone planning, and open-space configurations (e.g., mobile cabin hospital layouts). The integration overlays two data layers, offering insights into design impacts on transmission trends, applicable in various architectural projects. On one hand, it discusses the possibility of airborne transmission in the space through CFD. On the other hand, it simultaneously formulates strategies for the utilization of space by individuals and how to avoid the singularity of circulation. Currently, both of these aspects are still under further development.

However, limitations exist. In CFD simulations, precise values for infectious dose and virus shedding are elusive, complicating the determination of necessary virus quantities for infection or released by infected individuals. The intensive computation demands of this hybrid approach currently limit simulations to two-dimensional environments, as such simulations are resource-intensive and time-consuming, this is also the reason why many simulation tools are not accepted by architects or designers (Su and Yen, 2014). Addressing these computational challenges is also a crucial issue.

Future research will focus on overcoming these challenges, comparing transmission trends under different ventilation modes and occupancy strategies, and validating CFD simulations with more agent data to refine model accuracy. This research aims to inform effective infection control across diverse design approaches, inspiring designers to embrace integration across various fields and methodologies, fostering adaptability and flexibility in architectural design.

References

Zimring, C., Denham, M. E., Jacob, J. T., Cowan, D. Z., Do, E., Hall, K., Kamerow, D., Kasali, A., Steinberg, J. P. (2013). Evidence-based design of healthcare facilities: opportunities for research and practice in infection prevention. *Infection Control & Hospital Epidemiology,* 34(5), 514-516. https://doi.org/10.1086/670220

Gomez-Zamora, P., Hadi, K., Kemenova, O., Swarts, M. (2020). Spatiotemporal Modeling of COVID-19 Spread in Built Environments. In *24th Conference of the Iberoamerican Society of Digital Graphics: Transformative design, SIGraDi 2020* (pp.991-996). The International Conference of the Iberoamerican Society of Digital Graphics (SIGraDi). https://doi.org/10.5151/sigradi2020-134

Zhang, R., Li, Y., Zhang, A. L., Wang, Y., & Molina, M. J. (2020). Identifying airborne transmission as the dominant route for the spread of COVID-19. *Proceedings of the National Academy of Sciences,* 117(26), 14857-14863. https://doi.org/10.1073/pnas.2009637117

Perry, I., Xue, Z., Huang, H., Crispe, N., Vegas, G., Swarts, M., Gomez-Zamora, P (2022). Human Behavior Simulations to Determine Best Strategies for Reducing COVID-19 Risk in Schools. In *26th Conference of the Iberoamerican Society of Digital Graphics: Critical Appropriation, SIGraDi 2022* (pp.39-50). The International Conference of the

Iberoamerican Society of Digital Graphics (SIGraDi). https://doi.org/10.5151/sigradi2022-sigradi2022_298

Shangi, Z. A. D., Ahmad, M. I., & Ahmed, S. (2020). Adaptive Design of the Built Environment to Mitigate the Transmission Risk of COVID-19. *American Journal of Civil Engineering and Architecture,* 8(4), 146-153. https://doi.org/10.12691/ajcea-8-4-3

Nguyen, J., & Peters, B. (2020). Computational Fluid Dynamics in Building Design Practice. In *40th Annual Conference of the Association for Computer Aided Design in Architecture: Distributed Proximities, ACADIA 2020* (pp. 574-583). The Association for Computer Aided Design in Architecture (ACADIA).

Gomez-Zamora, P., Hadi, K., Swarts, M., Kemenova, O. (2021). Spatiotemporal Modeling: Spread of COVID-19 in Educational Settings: The Role of Architecture in A Pandemic. In *9th International Conference of the Arab Society for Computer Aided Architectural Design: ARCHITECTURE IN THE AGE OF DISRUPTIVE TECHNOLOGIES: Transformation and Challenges, ASCAAD 2021* (pp. 421-434). The International Conference of the Arab Society for Computer Aided Architectural Design (ASCAAD).

Gomez-Zamora, P., Bafna, S., Zimring, C., Do, E., & Vega, M. R. (2019). Spatiotemporal Occupancy for Building Analytics. In *Architecture in the Age of the 4th Industrial Revolution Proceedings of the 37th eCAADe and 23rd SIGraDi Conference* (pp. 111-120). The Education and Research in Computer Aided Architectural Design in Europe (eCAADe) and the International Conference of the Iberoamerican Society of Digital Graphics (SIGraDi). https://doi.org/10.5151/proceedings-ecaadesigradi2019_153

Den Hartog, J. P., Koutamanis, A., & Luscuere, P. G. (2000). Possibilities and limitations of CFD simulation for indoor climate analysis. *Fifth Design and Decision Support Systems in Architecture and Urban Planning-Part one: Architecture Proceedings,* 152-167.

Ren, C., Zhu, H. C., & Cao, S. J. (2022). Ventilation strategies for mitigation of infection disease transmission in an indoor environment: a case study in office. *Buildings,* 12(2), 180. https://doi.org/10.3390/buildings12020180

Yang, Z., Wang, Q., Zhao, L., & Long, E. (2022). Decay characteristics of aerosolized viruses in the air and control strategy of thermal and humid environment for epidemic prevention. *Indoor and Built Environment,* 31(5), 1287-1305. https://doi.org/10.1177/1420326X211052603

Batty, M. (2001). Agent-based pedestrian modeling. *Environment and planning B: Planning and Design,* 28(3), 321-326. https://doi.org/10.1068/b2803ed

Ortner, F., Tay, J. (2021). PANDEMIC RESILIENT HOUSING: Modelling dormitory congestion for the reduction of COVID-19 spread. In *26th International Conference of the Association for Computer-Aided Architectural Design Research in Asia: PROJECTIONS. CAADRIA 2021* (pp. 589-598). The International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) https://doi.org/10.52842/conf.caadria.2021.2.589

Su, Z., & Yan, W. (2014). Improving genetic algorithm for design optimization using architectural domain knowledge. In *34th Annual Conference of the Association for Computer Aided Design in Architecture: DESIGN AGENCY, ACADIA 2014*

EXPLAINABLE AI AND MULTI-OBJECTIVE OPTIMIZATION FOR ENERGY RETROFITS IN RESIDENTIAL NEIGHBORHOODS

JIAHONG YE [1], YANTING SHEN [2], CHENYU HUANG [3], JINYU WANG[4], RONG QU [5] and JIAWEI YAO [6]

[1,2,3,4,6]*College of Architecture and Urban Planning, Tongji University.*
[5]*China Construction Science and Industry Corporation LTD.*
[1]*yeesun_0v0@tongji.edu.cn*
[2]*sytsyt188@163.com*
[3]*cyhuang@tongji.edu.cn, 0000-0002-6360-638X*
[4]*johnwang1998@163.com*
[5]*qurong@cscec.com*
[6]*jiawei.yao@tongji.edu.cn, 0000-0001-7321-3128*

Abstract. Globally, the carbon emissions of building sector contribute to 40% of all. Therefore, low-carbon retrofits of buildings become pivotal to carbon neutrality. While existing retrofit research focus on individual building without considering whether different neighborhood morphologies would affect the effectiveness of retrofit strategies. Here we propose a framework for optimal retrofitting in diverse urban neighborhoods, balancing energy savings and cost. Based on actual residential neighborhood morphologies, massive parametric sample models have been established and batch simulated before and after the retrofitting for its energy performance. Then, we use XGBoost to establish a predictive model for energy savings, and iGeneS to perform multi-objective optimization. SHapley Additive exPlanations is utilized to reveal the nonlinear contribution of four retrofitting strategies to building retrofit benefits (BRB). Results demonstrate that in different floor area ratio (FAR), the contributions of installing rooftop photovoltaic panels to BRB vary. In neighborhoods with FAR of 1.5-2.8, the contribution of installing photovoltaic panels is not as significant as in other FAR ranges, and its contribution to BRB is not comparable to replacing energy-efficient lights. Moreover, The effectiveness of deep lighting fixture retrofitting may be suboptimal. The proposed framework will offer efficient energy-saving guidance for future residential neighborhood retrofits.

Keywords. Residential neighborhoods, Low-carbon retrofits, Urban morphology, Energy consumption, Explainable AI, Multi-objective optimization

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 485-494. © 2024 and published by the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA), Hong Kong.

1. Introduction

Excessive energy consumption leads to an increase in greenhouse gas emissions, thereby exacerbating the climate crisis. Hence, a consensus has emerged among nations to adopt a low-carbon, and sustainable development model (Huang et al., 2022). Globally, the carbon emissions of building sector contribute to 40% of all carbon emissions (Lu et al., 2023). The urban development of China, the world's largest CO2 emitter, is gradually shifting towards urban renewal. Therefore, low-carbon retrofits become pivotal in attaining carbon neutrality. While current research on building retrofits tends to focus primarily on individual buildings, overlooking the potential impact of interactions among building clusters, and face challenges in reconciling the trade-off between energy savings and economic costs (Ma et al., 2023). So, a framework proposed in this study integrates explainable artificial intelligence (AI) and a multi-objective optimization algorithm to achieve the optimal retrofit solution, balancing energy savings and costs across diverse neighborhood morphologies. Furthermore, the study presents the marginal benefits of different strategies.

2. Related works

2.1. LOW-CARBON RETROFIT OF BUILDINGS.

Currently, existing research has explored low-carbon retrofit strategies for individual buildings (Gustavsson et al., 2022), such as enhancing the thermal insulation of the building envelope (Aruta, Giuseppe et al., 2023), installing rooftop photovoltaic panels, etc. However, the adaptability of various retrofit strategies varies across different neighborhoods. (Hong et al., 2020). For instance, high building density (BD) blocks can lead to shading, impacting photovoltaic efficiency. while, Low BD neighborhoods might increase the amount of heat entering through windows. In such cases, replacing glass with a lower Solar Heat Gain Coefficient (SHGC) is more effective for energy savings. Therefore, it is imperative to broaden the scale of the study of low-carbon retrofit strategies.

2.2. SURROGATE MODELS OF ENERGY CONSUMPTION

Obtaining building energy consumption data through simulation often consumes a significant amount of time. Therefore, many scholars are attempting to accelerate the acquisition of building model energy consumption data through alternative methods. (Liu et al., 2023). For instance, employing machine learning to develop urban energy prediction surrogate models that leverage city-specific feature metrics as input, enable the efficient prediction of building energy consumption (Huang et al., 2022). This approach significantly enhances the efficiency of energy consumption data acquisition.

2.3. EXPLAINABLE MACHINE LEARNING

Model explainability refers to the extent of human understanding of Machine Learning (ML) model predictions and decisions. Currently, the primary application of ML in the field of energy lies in energy consumption prediction. However, some researchers have combined machine learning with explainable analysis to guide design practices. For instance, by utilizing an ensemble algorithm that combines LightGBM with SHapley

Additive exPlanation (SHAP), various urban morphology indicators are explained in terms of their impact and contribution to urban energy consumption and carbon emissions (Zhang et al., 2023).

3. Methods

3.1. RESEARCH WORKFLOW

A framework proposed in this study integrates explainable AI and multi-objective optimization to attain the optimal retrofits solution in neighborhood morphologies. The optimization objectives encompass the maximum energy savings while minimum retrofits costs.

Firstly, this study prototypes real residential neighborhood morphology, constructing a batch of idealized sample models. Secondly, using Urban Weather Generator (UWG), Honeybee, and OpenStudio, the sample neighborhood models undergo two batches of performance simulations before and after retrofitting, and retrofitting costs are calculated. Subsequently, data on urban morphology parameters, retrofit strategy parameters, and energy savings are collected. XGBoost is employed to model the impact of urban morphology and four retrofit strategies on residential neighborhood energy savings, resulting in a predictive surrogate model for post-retrofit energy savings. Different forms of neighborhood retrofit strategies undergo multi-objective optimization to obtain post-retrofit Pareto solutions by multi-objective optimization plugin, iGeneS. Finally, XGBoost is utilized to model the impact of Pareto solutions' four retrofit strategies on residential neighborhood energy savings and retrofit costs. SHAP analysis is then applied to examine the nonlinear contributions of urban morphology and the four retrofit strategies to residential neighborhood energy savings and retrofitting cost when energy savings is the same. The research workflow is shown in the Fig 1.

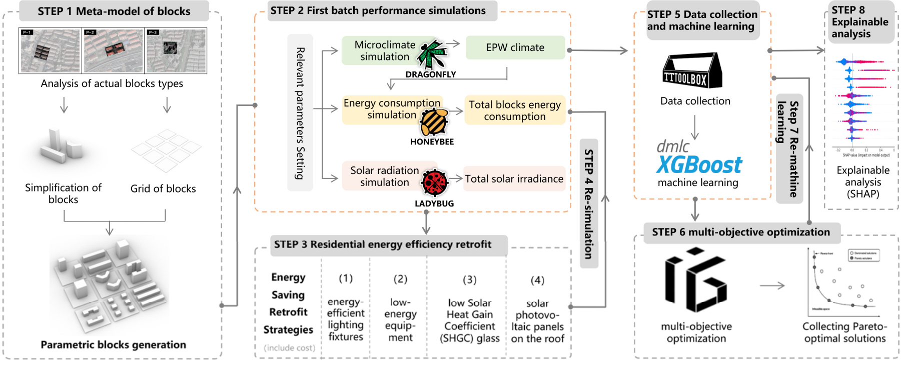


Figure 1. Workflow

3.2. RESEARCH OBJECT AND THE CONSTRUCTION OF A BATCH RESIDENTIALMODELS.

Based on real urban morphology, an ideal residential neighborhood measuring 240m x 240m is planned within the Grasshopper. Subsequently, the land is divided into a 3 x 3 grid with 10m-wide roads. As shown in Figure 2, this study analyzes the residential building types in Shanghai and identifies nine typical morphologies. These include five categories of point-style residential clusters (P-1~P-5), two categories of slab-style residential clusters (S-1, S-2), and two categories of enclosure-style residential clusters (C-1, C-2). Next, these morphologies are simplified and randomly placed within the residential grid to generate a host of diverse sample models. The generated building heights are controlled between 3 and 25 floors, with a uniform floor height set at 3m.

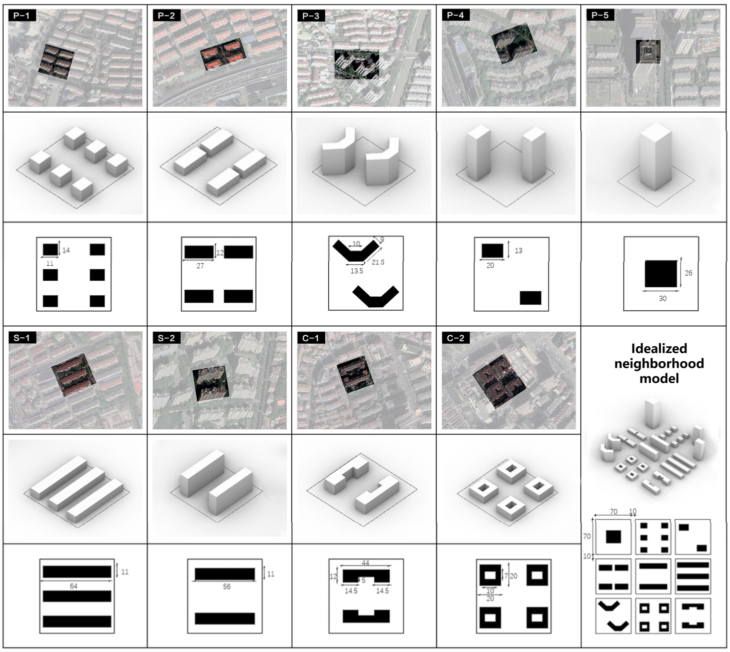


Figure 2. Samples of typical residential building types in Shanghai

3.3. INITIAL PERFORMANCE SIMULATION OF NEIGHBORHOODS

Using the Dragonfly and Honeybee, a large batch of sample models were transformed into UBEM. Based on previous research and the actual proportion of the actual distribution of residential functions (Liu et al., 2023), the idealized residential functional ratios are configured as follows: bedrooms accounting for 40%, living rooms for 20%, kitchens for 10%, bathrooms for 10%, and hallway spaces for 20%.

As shown in Table 1, this study, referencing relevant standards, sets the basic parameters of buildings before retrofitting. Additionally, considering the presence of the urban heat island effect, this research employed the UWG tool in energy simulations to conduct microclimate simulations of sampled neighborhoods for corresponding EPW climate files, which were then integrated with the parameterized UBEM and input into OpenStudio for simulations.

Table 1. Basic parameters of the buildings Before Retrofitting

Basic Information	Indicators	Parameters
Lighting	Lighting power density (w/㎡)	5.0
Equipment	Equipment power density (w/㎡)	3.8
People	Occupancy density (People/㎡)	0.04
Air Conditioning	Summer cooling temperature (℃)	26
	Winter heating temperature (℃)	18
Envelope	External wall heat transfer coefficient [w/(㎡ • k)]	0.8
	roof heat transfer coefficient [w/(㎡ • k)]	0.5
	partition wall heat transfer coefficient[w/(㎡ • k)]	1.5
	SHGC of glass	0.84
	glass heat transfer coefficient [w/(㎡ • k)]	2.7
Window-to-wall ratio	north-south window-to-wall ratio	0.4
	east-west window-to-wall ratio	0.1

3.4. LOW-CARBON RETROFIT AND RE-SIMULATION

This study applied four retrofitting strategies including utilizing energy-efficient lighting fixtures, low-energy equipment, low Solar Heat Gain Coefficient (SHGC) glass, and installing solar photovoltaic panels on the roof. Considering the actual situation, this study defined the post-retrofit parameter changes as follows: a reduction range for residential building lighting power density from 0 to 3.55 W/m^2 , a reduction range for equipment power density from 0 to 1.5 W/m^2 , a reduction range for glass SHGC from 0 to 0.7, and a power variation range for solar photovoltaic panels from 0.275 to 0.6 kW. Additionally, the study further categorized these retrofitting strategies into shallow, moderate, and deep retrofitting, with specific parameter range variation outlined in table 2.

Table 2. Range of retrofit parameter variations

	randomized retrofitting	Shallow retrofitting	Moderate retrofitting	Deep retrofitting
Reduction range of lighting power density (w/m^2)	0-1.5	0-0.5	0.5-1	1-1.5
Reduction range of equipment power density(w/m^2)	0.2-1.7	0.2-0.7	0.7-1.2	1.2-1.7
Reduction range of glass SHGC	0-0.7	0-0.233	0.233-0.466	0.46-0.7
Increase range of photovoltaic panel power (kw)	0.275-0.6	0.275-0.383	0.383-0.492	0.492-0.6

Based on market research, this study obtained retrofit costs (RC) calculation equations for different strategies through regression fitting (Equations 1-4). Following the completion of retrofitting, a second batch of performance simulations was conducted on the sampled models.

$$y_1 = -4.28x_1 + 27.76 \tag{1}$$
$$y_2 = -24.59x_2 + 89.61 \tag{2}$$
$$y_3 = -582.55x_3 + 504.02 \tag{3}$$
$$y_4 = 800x_4 + 331 \tag{4}$$

In the equations, $x_1$represents the post-retrofit lighting power density, $x_2$denotes the post-retrofit equipment power density, $x_3$represents the post-retrofit SHGC value of glass, and x_4 represents the post-retrofit power of the photovoltaic panels.

Additionally, y_1 signifies the cost of replacing energy-efficient lighting fixtures, y_2 represents the cost of replacing low-energy equipment, y_3 is the cost of replacing low-SHGC glass, and y_4 denotes the cost of adding photovoltaic panels.

3.5. CONSTRUCTING AN ENERGY PREDICTION SURROGATE MODEL

The TT toolbox plugin was employed to collect simulation datas, including morphological parameters such as Floor Area Ratio (FAR), Building Density (BD), as well as strategy-related variations like the differences in lighting and equipment power density. Additionally, data on retrofit costs of per square meter(RCM), changes in Energy Use Intensity before and after retrofitting(EUI_D), and Building Retrofitting Benefits (BRB) were collected. The calculation of BRB is as shown in equation 5.

$$BRB = \frac{\text{RCM}}{\text{EUI\_D}}\ (CNY/EUI) \tag{5}$$

Subsequently, XGBoost was applied to model the nonlinear relationships between neighborhood morphology, low-carbon retrofit strategy variables, and EUI_D. The models were evaluated using the coefficient of determination (R^2), resulting in surrogate models that map key design parameters to post-retrofit energy savings.

3.6. MULTI-OBJECTIVE OPTIMIZATION AND SHAP ANALYSIS

In this study, the iGeneS was employed in collaboration with surrogate models to achieve multi-objective optimization. IGeneS is a Grasshopper plugin capable of parallel computation, known for its fast execution speed and high convergence. The optimization parameters in this study included a population size of 100 and 100 generations for the number of iterations. The optimization objectives were set as maximizing EUI_D and minimizing RCM. The adjustable parameters are input parameters for the four retrofit strategies, including the decrease in lighting and equipment power density, glass SHGC reduction, and solar photovoltaic panel power increment.

Upon completion of multi-objective optimization, the Pareto solution sets data was output. Subsequently, XGBoost was applied to model the nonlinear relationships between neighborhood morphology, low-carbon retrofit strategies, and EUI_D and BRB. Following this, the SHAP was utilized to quantify the contributions of different low-carbon retrofit strategies to EUI_D and BRB.

4. Results

This study conducted low-carbon retrofitting and multi-objective optimization on 500 distinct urban morphologies, resulting in a total of 39,700 samples of Pareto solution data.

4.1. THE CORRELATION BETWEEN MORPHOLOGY AND EUI_D AND BRB

As shown in Figure 3, through Pearson correlation analysis, this study reveals the correlation between residential area morphology indicators and EUI_D and BRB. Specifically, BD, SD, OSR exhibit positive correlations with EUI_D and BRB, while FAR, AF exhibit negative correlations with EUI_D and BRB.

Figure 3. Correlation between morphological indicators and EUI_D and BRB

4.2. RETROFIT INTENSITY OF DIFFERENT NEIGBORHOODS

Due to the influence of morphology on EUI_D and BRB, and the presence of collinearity among some indicators, this study selects the commonly used FAR and BD to perform the following gradient division. Low FAR is defined as 0.7 to 1.5, medium FAR as 1.5 to 2.8, and high FAR as 2.8 to 4. The range for low BD is 0.15-0.22, medium BD is 0.22-0.3, and high BD is 0.3-0.4. Then, we further analyze the intensity of low-carbon retrofitting strategy selection under different morphologies.

4.2.1. Low-carbon Retrofitting Intensity under Different FAR

As illustrated by the distribution of Pareto solutions in Figure 4, this study observed minimal variations in the retrofitting intensity of glass, equipment, and lighting fixtures across different FAR for the four retrofitting strategies. Specifically, the retrofitting intensity for glass and equipment remained at a shallow level, while the lighting fixtures exhibited both shallow and moderate retrofitting intensities. The strategy involving the addition of solar panels showed variability in different FAR residential areas; in low FAR neighborhood, a moderate retrofitting level is most suitable, while in medium to high FAR neighborhood, deep retrofitting could be considered.

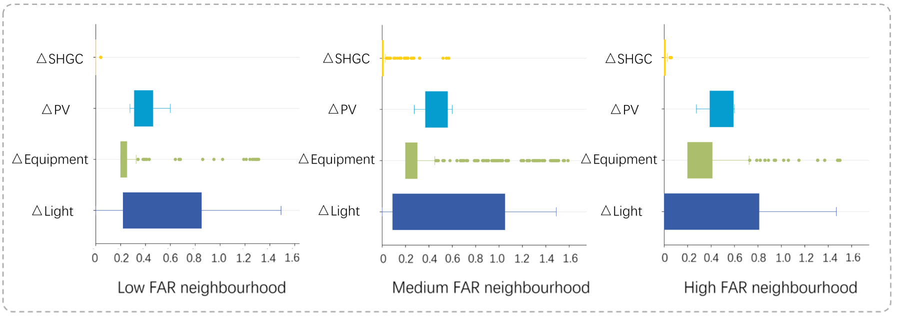


Figure 4. Distribution of retrofitting strategy intensity under different FAR

4.2.2. Low-carbon Retrofitting Intensity under Different BD

As shown in Figure 5, in different BD residential neighborhoods, there are variations in the adaptability of the four retrofitting strategies. Based on the analysis of the Pareto solution sets, this study found the retrofitting intensity for glass and equipment remained at a shallow level, while the retrofitting intensity for photovoltaic panels ranged from moderate to deep retrofitting. The retrofitting of lighting fixtures showed differences across different BD residential neighborhoods, with shallow and moderate retrofitting being suitable for medium to low BD neighborhoods, and medium to deep retrofitting being predominant in medium to high BD neighborhoods.

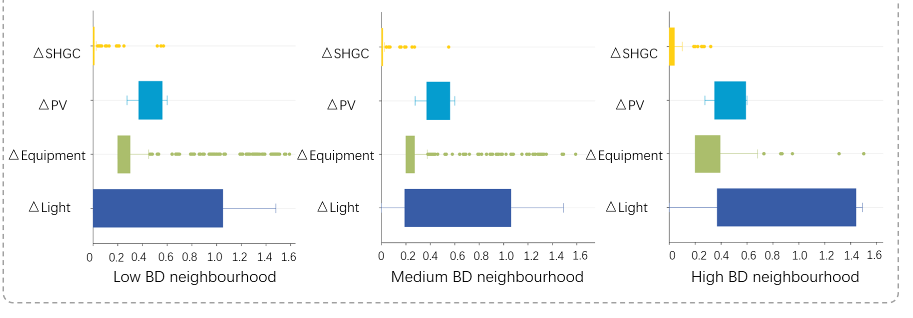


Figure 5. Distribution of retrofitting strategy intensity under different BD

4.3. DISPARITIES IN ENERGY-SAVING CONTRIBUTIONS OF DIFFERENT RETROFITTING STRATEGIES

4.3.1. Disparities in retrofitting strategy contributions in different FAR residential neighborhoods

Considering the optimal compromise between the two objectives in this study, various low-carbon retrofitting strategies demonstrate differences in their contributions to EUI_D across different FAR. As depicted in Figure 6, the strategy with the greatest EUI_D contribution in low FAR and high FAR residential neighborhoods is installing photovoltaic panels. However, in medium FAR residential neighborhoods, retrofitting lighting fixtures stands out as the most prominent contributor to EUI_D.

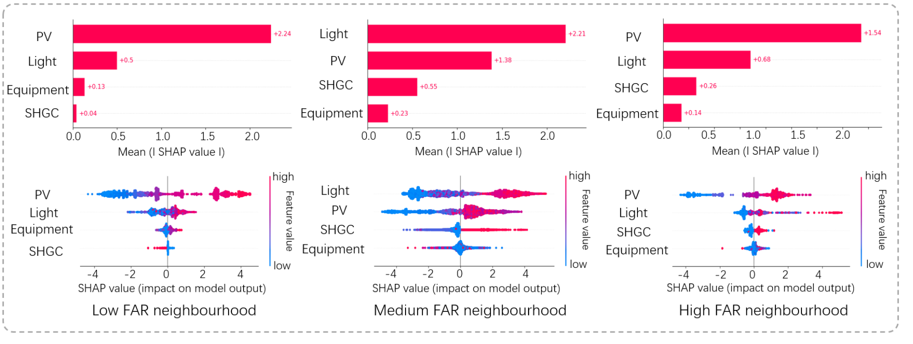


Figure 6. Retrofitting strategy contributions in different FAR

4.3.2. Disparities in retrofitting strategy contributions in different BD residential neighborhoods

Similarly, considering both retrofitting costs and energy savings, various retrofitting strategies exhibit differences in their contributions to energy savings across different BD. As illustrated in figure 7, in low BD residential neighborhoods, the strategy with the greatest energy-saving contribution is the replacement of energy-efficient lighting fixtures. However, in medium BD and high BD residential neighborhoods, the most prominent contribution installing photovoltaic panels.

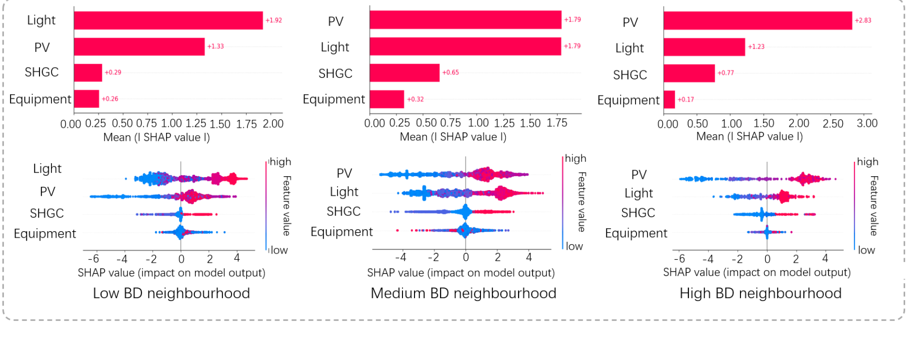


Figure 7. Retrofitting strategy contributions in different BD

5. Conclusion

This study aims to explore the adaptability and energy-saving contribution differences of four strategies in residential neighborhoods. The results indicate that two retrofitting strategies, replacing low SHGC glass and upgrading to high-efficiency equipment, are prioritized for shallow retrofitting, while strategies involving the replacement of energy-efficient lighting fixtures and installing rooftop photovoltaic panels may be considered for moderate retrofitting. Additionally, installing the rooftop photovoltaic panels can be implemented for deep retrofitting in medium to high FAR residential neighborhoods. From an energy-saving contribution perspective, among the four energy-saving strategies in this study, prioritizing installing rooftop photovoltaic panels and the replacement of energy-efficient glass is most advantageous for enhancing the carbon reduction benefits.

The conclusions drawn in this study offer guidance for the retrofitting of neighborhoods, aimed at further enhancing the potential for low-carbon retrofitting. However, this study has certain limitations, such as drawing conclusions solely from simulations and without considering the causal effects between urban morphology, strategies, and energy savings. Therefore, future research will validate the findings using actual energy consumption data and incorporate causal inference into the research process to explore the comprehensive potential of various retrofitting strategies, promoting sustainable urban development.

Acknowledgements

This research was financially supported by the National Natural Science Foundation of China under Grant NO. 52278041 and the Fundamental Research Funds for the Central Universities.

References

Aruta, G., Ascione, F., Bianco, N., & Mauro, G. M., (2023). Sustainability and energy communities: Assessing the potential of building energy retrofit and renewables to lead the local energy transition. Energy, 282, 128377.

Gustavsson, L., & Piccardo, C., (2022). Cost optimized building energy retrofit measures and primary energy savings under different retrofitting materials, economic scenarios, and energy supply. Energies, 15(3), 1009.

Huang, C., Zhang, G., Yin, M., & Yao, J. (2022). Energy-driven intelligent generative urban design. CAADRIA, volume 1, 233-242.

Hong, T., Chen, Y., Luo, X., Luo, N., & Lee, S. H., (2020). Ten questions on urban building energy modeling. Building and Environment, 168, 106508.

Lu, Y., Chen, Q., Yu, M., Wu, Z., Huang, C., Fu, J., ... & Yao, J. (2023). Exploring spatial and environmental heterogeneity affecting energy consumption in commercial buildings using machine learning. Sustainable Cities and Society, 95, 104586.

Liu, K., Xu, X., Zhang, R., Kong, L., Wang, W., & Deng, W. (2023). Impact of urban form on building energy consumption and solar energy potential: A case study of residential blocks in Jianhu, China. Energy and Buildings, 280, 112727

Ma, D., Li, X., Lin, B., Zhu, Y., & Yue, S. (2023). A dynamic intelligent building retrofit decision-making model in response to climate change. Energy and Buildings, 284, 112832.

Zhang, Y., Teoh, B. K., Wu, M., Chen, J., & Zhang, L. (2023). Data-driven estimation of building energy consumption and GHG emissions using explainable artificial intelligence. Energy, 262, 125468

GLARE PREDICTION IN CHECK-IN HALLS OF AIRPORT TERMINALS USING INTEGRATED ALGORITHMS AND TRANSFER LEARNING STRATEGY

A Case Study of Guangzhou Baiyun International Airport

YINYI SHI[1], ZHICHAO MA[2], CUNYUAN WANG[3] and XIONG CHEN[4]
[1]*State Key Laboratory of Subtropical Building and Urban Science, South China University of Technology.*
[2]*College of Architecture and Urban Planning, Tongji University.*
[3]*Harbin Institute of Technology, Shenzhen.*
[4]*Guangdong Architectural Design and Research Institute Co., Ltd.*
[1]*shiyinyi_work@163.com, 0000-0002-5305-1333*
[2]*mazhichao@tongji.edu.cn*
[3]*2426452800@qq.com, 0000-0002-0529-7566*
[4]*chenxiong@gdadri.com*

Abstract. The construction process of large-scale spaces such as airport terminals is often carried out in several phases, so the environmental performance assessment and analysis of completed projects can provide effective reference information for new projects. On the other hand, the layout of large glass curtain walls and skylights in the check-in halls of terminals, while fully introducing natural light, also brings potential glare hazards, and therefore the influence of different design parameters on glare needs to be clarified. However, current research has not yet discussed in detail the prediction of glare performance of terminal buildings and its influencing factors. This study aims to develop a transfer learning strategy and a workflow for predicting glare performance in terminal buildings. The results have proved that the transfer learning strategy can help quickly predict the glare performance between projects with similar spatial characteristics with high accuracy, and the outcomes also help clarify the influencing factors of glare performance and provide designers or managers with support for performance prediction and optimization methods.

Keywords. Check-in Hall, Airport Terminal, Annual Glare, Performance Prediction, Transfer Learning, Integrated Algorithms

1. Introduction

1.1. BACKGROUND

According to the 14th Five-Year Plan for the Development of Civil Aviation in China, the construction of civil airports in China will reach 270 by the end of 2025 (Yan S, 2022). Green and sustainable development goals are also receiving increasing attention during the progress of the civil aviation industry. During the design and construction of a terminal building, a phased approach is often used due to the complexity of the process and the long period it takes. For the same airport, terminals constructed in different periods often have certain similarities, mainly in terms of space and functional composition. This makes the assessment and analysis of the completed projects of the previous phase a high reference for the subsequent phases and other projects of the same type.

Large glass curtain walls combined with roof skylights are often used to ensure that natural light is fully introduced, which helps to reduce the need for artificial lighting and improve the indoor light environment while reducing energy consumption. However, the large lighting area inevitably brings a certain degree of glare hazard, and the glare will mainly affect passenger movement efficiency (Yi L et al., 2019). To adopt appropriate strategies to reduce the glare hazard at the design stage, it is important to investigate the annual glare-influencing factors and obtain the optimal solution.

1.2. RELATED WORKS

Due to the progress of computer simulation technology and artificial intelligence algorithms, building performance prediction has received widespread attention. In recent years, many researchers have realized rapid prediction of building environmental performance indicator values under different parameter combinations by establishing prediction models. For example, Aseel Hussien et al. used a random forest algorithm to predict the long-term energy performance of building envelopes (Aseel H et al., 2023), while He et al. integrated parametric design, performance simulation, image processing, and machine learning techniques to achieve rapid prediction and assessment of the wind environment of building clusters (Yi H et al., 2021). In a previous study, we also used integrated machine learning algorithms combined with a multi-objective optimization process to establish a prediction model between façade shading and indoor light environment and energy consumption for a typical corridor of a terminal building and proposed a preferred solution that could be screened by designers (Yinyi S et al., 2023).

However, the above studies on performance prediction and optimization are only available for specific buildings or certain types of datasets. Many design projects usually need to deal with various types of buildings and may also be constructed in different phases like terminal buildings, which puts a higher demand on the generalization and scalability of prediction models. The strategy of transfer learning has been proposed under such demands, which effectively adapts the training model by using a new building dataset and transferring the useful knowledge of the training model from the original dataset (Giuseppe P, 2022). For example, Fang et al. applied a hybrid deep transfer strategy to achieve the goal of assisting energy prediction of a target building with limited historical measurements (Xi F, 2021). Liu et al. applied the transfer learning strategy to energy systems in the field of data-driven fault detection and evaluation (Jiangyan L et al., 2021).

1.3. OBJECTIVE AND ORIGINALITY

Based on the above background and analysis of related literature, it is evident that the application of transfer learning in the performance prediction of such large public buildings or spaces as airport terminals has not been discussed in existing studies. The influencing factors of glare in airport terminals have also not been systematically compared. To bridge the existing research gap, the objective of this study is to develop a transfer learning strategy and workflow for modeling the prediction of annual glare in the check-in halls of different airport terminals, as well as to clarify the impact of different design factors on the glare performance, and Terminal 1 (T1, built in 2004) and Terminal 2 (T2, built in 2018) of Guangzhou Baiyun International Airport are proposed to be selected as demonstration cases.

2. Methodology

2.1. OVERVIEW OF WORKFLOW

The overall workflow of this study is shown in Figure 1. The workflow can be divided into three main phases. Among them, the first phase involves the establishment of parametric models and the acquisition of different design parameters for both T1 and T2, which include the roofing system (including skylight, metal roof, and ceiling), glass curtain wall, floor, and building orientation, etc. The second phase involves obtaining annual glare performance data through batch simulation and completing the machine learning process for T1. The third stage is to transfer the trained model from T1 to T2 through the transfer learning strategy and then complete the adjustment and prediction.

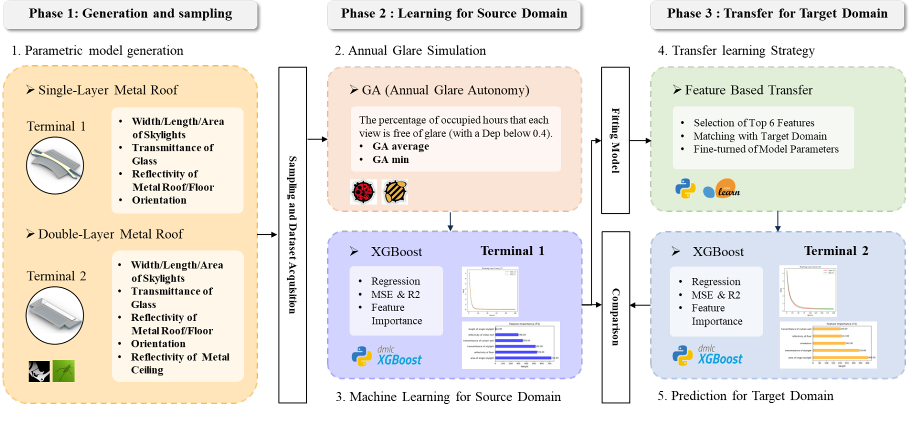


Figure 1. Overview of workflow

2.2. PARAMETRIC MODELLING AND SETTINGS

The determination of the design parameters of the terminal building check-in hall is based on the existing cases. After obtaining the original scheme design parameters, the study focuses on comparing the impact of skylights to simplify and improve the computational efficiency, so other geometric parameters are kept unchanged except for the skylights. In addition, the building orientation and the material parameters of each component will also be set parameter intervals. Figure 2 shows the parameterized model, the interior view, and the schematic of the sectional spatial forms of the selected two terminals. The main difference between them is that T1 has a single-layer roof combined with skylights, while T2 has a double-layer roof with the addition of metal ceilings, and the rest of the spatial forms of the two buildings are similar. Table 1 shows the parameter settings of each part of the model. Meteorological parameters are selected from typical year data of Guangzhou city in China.

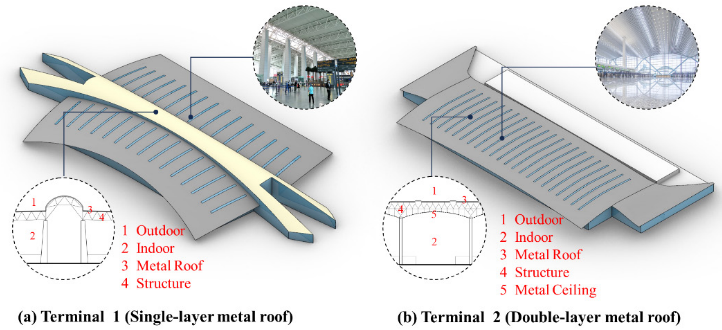


Figure 2. Parametric modeling of check-in halls in different airport terminals

Table 1. Summary of design parameters for terminal check-in hall.

Classification	Design parameter	Range	Baseline	Unit
Geometry	Skylight width	[3, 12]	3 (T2)	m
	Skylight length	[30, 125] (T2); [30, 60] (T1)	125 (T2)	m
	Orientation	[0, 360][a]	15	degree
	Number of skylight columns	16 (T1); 18 (T2)	18 (T2)	-
Glass	Skylight window transmittance	[0.3, 0.6]	0.3	-
	Curtain window transmittance	[0.4, 0.8]	0.5	-
General indoor wall	Reflectivity	-	0.2	-
Roof film (T1)	Transmittance	-	0.15	-
	Reflectivity	-	0.75	-
Floor	Reflectivity	[0.3, 0.6]	0.6	-
Metal roof	Reflectivity	[0.4, 0.8]	0.8	-
Metal ceiling (T2)	Reflectivity	[0.4, 0.8]	0.4	-

a. Interval of orientation will be set as 90° , which means the orientation includes N (north), E (east), W (west), and S (south).

2.3. PERFORMANCE SIMULATION AND DATASET

To balance the performance simulation computation time with the availability of the actual scenarios, the simulation samples of the two terminals are set to 1408 and 3584 respectively. After completing the sampling of model parameters, the calculation of annual glare performance is completed by Ladybug and Honeybee performance calculation plug-ins of Rhina & Grasshopper, and the simulation calculation process is driven by algorithms. Finally, the model parameters and the performance simulation results are recorded in an Excel sheet.

Considering the characteristics of the terminal building's long-period operation mode, the glare performance simulation analysis metrics were determined as Daylight Glare Probability (DGP) and GA (Annual Glare Autonomy). According to the International Commission on Illumination (CIE), the range of DGP values corresponds to different levels of glare. GA, on the other hand, represents the percentage of glare-free time occupied by each viewpoint at each measurement point in the space, and in this study, if the annual average of DGP in the viewpoints is less than 0.4, it will be counted as the cumulative timeshare of GA. The final performance metrics for this study will be defined as the minimum and average values of GA among all measurement points for global considerations.

2.4. MACHINE LEARNING BASED DATA PREDICTION

2.4.1. XGBoost integrated algorithm

For structured data prediction, integrated machine learning algorithms based on decision trees have been proven to be superior by many studies, while XGBoost, as one of the models in this category, is executed recursively during training (Tianqi C and Carlos G, 2016). To further validate the advantages of XGBoost in this study, we also selected the T1 terminal dataset for validation to compare the performance difference between XGBoost and Random Forest as well as other ANN (Artificial Neural Network) algorithms such as MLP (Multilayer Perceptron) and SVM (Support Vector Machines), and the results show that the performance of XGBoost is better, and the specific results are detailed in Table 2. Therefore, the algorithm chosen in this study is the XGBoost integrated learning algorithm, which achieves the desired fitting effect by defining the model parameters in the regression task during training, and the main model parameters involved in this study include learning_rate, max_depth, n_estimators, etc.

For the evaluation of the training effect, this study selects Mean Squared Error (MSE) and R Squared as the evaluation indexes, in which the smaller the value of MSE and the closer the value of R Squared is to 1, the better the training effect of the model is, and the formulas for the calculation of MSE and R Squared are shown as follows. Root Mean Squared Error (RMSE) will be selected as the evaluation index in training loss analysis, which is the arithmetic square root of MSE.

$$MSE = \frac{1}{m}\sum_{i=1}^{m}(y_i - \breve{y}_i)^2$$

$$RMSE = \sqrt{MSE}$$

where m denotes the number of samples and $y_i - \breve{y}_i$ denotes the simulated value minus the predicted value.

$$R^2 = 1 - \frac{\sum_i(\breve{y}_i - y_i)^2}{\sum_i(\bar{y}_i - y_i)^2}$$

Where the numerator part represents the sum of the squared differences between the true and predicted values and the denominator part represents the sum of the squared differences between the true and predicted average values.

Table 2. Comparison of the different machine learning algorithms for the GA_average of T1

Algorithms	XGBoost	Random Forest	MLP	SVM
MSE	0.803	2.693	0.162	3.151
R Square	0.986	0.821	0.974	0.946

2.4.2. Transfer learning strategy

Transfer learning is the process of transferring knowledge from a learned related task to improve the new task from the similarity of data and task, and the two main concepts involved are domain and task. The domain is composed of feature space and probability distribution, and the task is composed of label space and predictive function. For this study, the feature space usually consists of the size of the skylight, the orientation of the building, the material parameter, etc., whereas the label space is the average and minimum values of GA.

The source domain is the dataset defined as T1, while the dataset of T2 is defined as the target domain. The transfer learning we define is as follows: given the source domain D_source and learning task T_source, the target domain D_target, and the target task T_target, the transfer learning aims to improve the target prediction function and D_source $\neq$ D_target but T_source = T_target. Specific transfer learning is divided into the following three steps: (1) Data acquisition and preprocessing. (2) Training in the source domain. (3) Using a small amount of data in the target domain to fine-tune and modify the pre-trained model in the source domain. (4) Completion of predicting in the target domain.

For the source domain, 80% of the data will be used for training and 20% for testing. Whereas in the target domain, only 20% of the data will be selected for fine-tuning and the remaining data will be used for testing.

3. Results and Discussion

3.1. DISTRIBUTION OF GLARE PERFORMANCE DATASET

In Figure 3, we show a comprehensive dataset of the glare performance of the check-in halls of the terminals obtained from the batch performance simulation. Overall, the distribution trends of the datasets for the two terminals are somewhat similar, i.e., the average values of GA are all higher while the minimum values are lower, and the distribution of the minimum values is more dispersed. Specifically, for the average value of GA, T2 generally has values higher than 75%, reaching 90% on average, with the highest even approaching 100%. T1, on the other hand, is mainly concentrated in the range of 20%-40%, with an average of only 30%, which shows that the double-layer roof system with a ceiling is more conducive to controlling the generation of uncomfortable glare. For the minimum value of GA, the lower limit of both terminals is lower than 20%, which shows that although uniformly arranged roof skylights are set up, there still exists a greater risk of glare hazard in some specific spaces, which requires the consideration of certain control strategies.

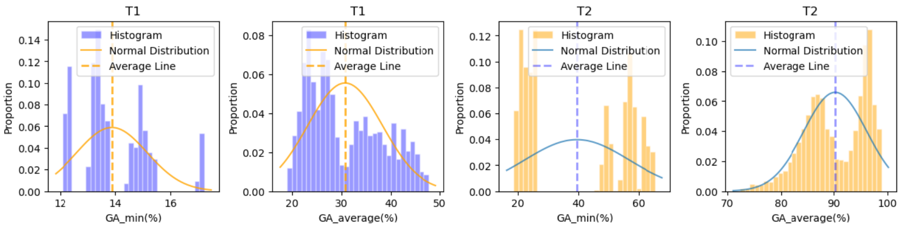


Figure 3. Histogram of GA distribution

The results of the correlation analysis for the two terminals are shown in Figure 4. Among them, the correlation between the glass curtain wall, floor parameter, and area of skylight and the glare performance is higher and shows a negative correlation, indicating that the glare hazard increases as the values of these parameters increase. It is worth noting that the correlation between the design parameters and the average value of GA is higher than the minimum value, which shows that the variation of the design parameters has a lower impact. In addition, the increase in building orientation angles in T1 leads to an increase in the potential hazard of glare.

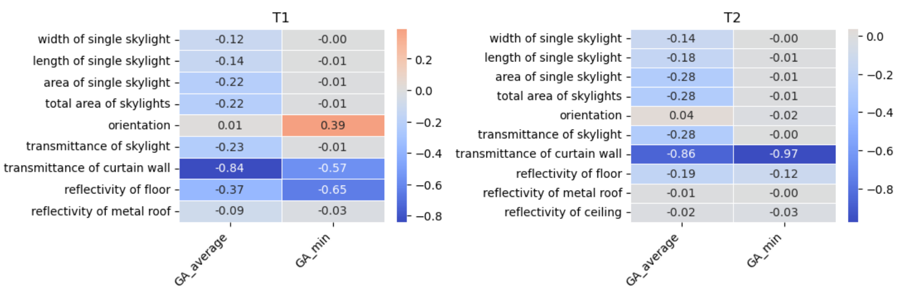


Figure 4. Correlation analysis between model parameters and glare indicators

3.2. EVALUATION AND VALIDATION OF TRANSFER LEARNING

Figure 5 depicts the machine learning in the source domain and the transfer learning training process in the target domain. During the training process, the training set and test loss show a monotonically decreasing trend, implying that the optimization process is stable and converges after 20 and 75 generations, respectively. In addition, we also compared the training effects of different methods or strategies to the transfer learning strategy, and the summary results are shown in Table 3.

Table 3 records the values of R Squared and MSE in the source and target domains by the three strategies of only source, data combination, and transfer learning, and although the different strategies show better training effects in the source domain, the prediction effects of the first two strategies in the target domain are weaker. The R Squared in the target domain under the data combination strategy is only 0.514, while the R Squared in the target domain and the source domain under the transfer learning strategy are 0.986 and 0.921, respectively. Therefore, we believe that the transfer learning strategy developed in this study can effectively improve the prediction effect in the target domain, and thus assist in the prediction of glare.

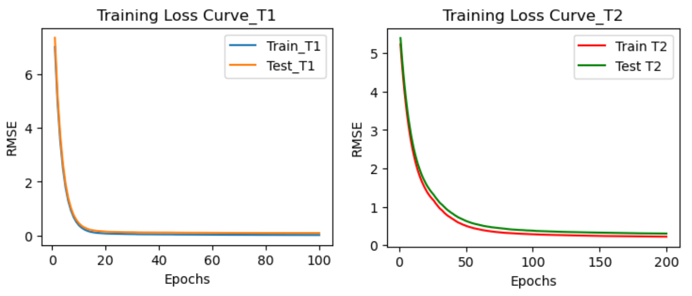


Figure 5. Training loss of source domain (T1) and target domain (T2)

Table 3. Comparison of prediction model evaluation in source and target domains under different strategies

Strategy	Source Domain (T1)		Target Domain (T2)	
	R Squared Score	MSE Loss	R Squared Score	MSE Loss
Only source	0.986	0.803	-56.510	2356.807
Data combination	0.955	23.05	0.514	28.577
Transfer learning	0.986	0.803	0.921	2.758

3.3. FEATURE IMPORTANCE ANALYSIS

Figure 6 shows the results of the feature importance analysis conducted after completing the training in the target and source domains, demonstrating the extent of the influence of different features on glare performance. Overall, the area of a single skylight has the greatest impact on the glare performance of both terminals, which shows that the reasonable design of skylight forms and sizes is important for reducing the glare hazard. Secondly, the material parameters of the skylight and glass curtain wall are also the key features affecting the glare performance, and these features play an important role in the glare control of both terminals.

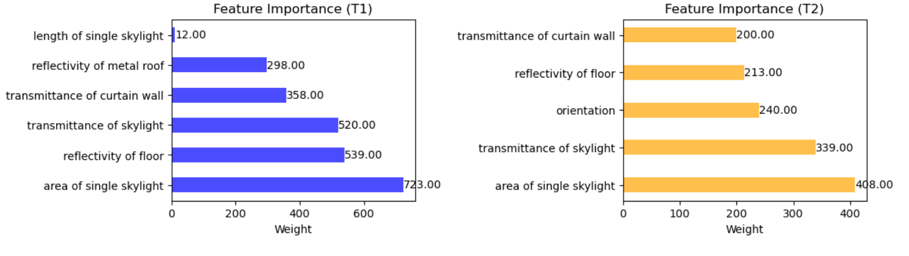


Figure 6. Feature importance analysis in the source and target domain

It is also worth noting that some parameters in the target and source domains have a low impact on glare performance, such as the length, width, and total area of skylights, so it is recommended that the design process focuses on the form and area of individual skylights, as well as the material parameters of the key envelope. Besides, the reflectivity of the floor will also have a certain impact on the glare performance, which is consistent with the results of a previous study (Xingyue H et al., 2021).

3.4. CONTRIBUTION AND APPLICATION

The results of this study demonstrate the feasibility of a machine learning model using a transfer learning strategy for predicting the glare performance of different terminals, which will provide an important reference for the new construction and expansion of other airports in the future. In the practical application of other projects, designers only need to adjust the relevant procedures of parametric modeling, while the batch simulation, machine learning models and transfer learning strategies applied in this study can be referred to directly.

The steps and process of this study can also be referred to when conducting prediction studies for other performance metrics. Meanwhile, the predictive models can be applied directly to the performance optimization process in the next step, which will reduce simulation time driven through the optimizer, and the data can be saved and recalled more flexibly.

3.5. LIMITATION AND FUTURE WORKS

This study also has some limitations. Firstly, the acquisition of design parameters failed to take into account the variation of the overall geometry, and only the skylight parameters were selected as variables. Therefore, in the subsequent study, we will enrich the design parameters to obtain a more complete and representative dataset. Secondly, in the selection of environmental performance targets, the annual glare indicator (GA) was selected in this study, and in terms of future works, we will consider other lighting indicators and combine deep learning to establish an image-based glare prediction study. Thirdly, the T1 and T2 in the study have similar spatial and structural characteristics, and if differentiated spatial types are involved, special attention needs to be paid to the accuracy of the results after transfer learning, and further comparisons can be attempted in conjunction with different transfer learning strategies.

4. Conclusion

This study develops and demonstrates a glare performance prediction and transfer learning strategy for check-in halls in terminals of different spatial types. It aims to obtain the glare performance impact characteristics of different spaces and provide methodological support for future glare performance prediction and transfer learning studies of similar spaces. The main conclusions of the study are summarized as follows.

- Double-layer roofing systems with evenly spaced ceilings are more conducive to less glare hazards and a comfortable indoor light environment.
- The application of transfer learning to glare performance prediction of terminal buildings has good practical value, and in this study, the prediction accuracy in the source domain (R Squared= 0.986) and the application of the XGBoost model to the performance prediction task in the source domain by utilizing transfer learning

is also able to achieve good prediction accuracy (R Squared= 0.921).

- For the case of this study, the area of a single skylight is the most critical feature affecting glare performance. Secondly, skylight glass, curtain wall glass, and floor material parameters are also important features affecting glare performance, and the design process should rationalize skylight forms and sizes to control direct and floor reflectivity to reduce reflected glare.

Acknowledgements

We would like to thank Guangdong Architectural Design and Research Institute Co. Ltd. for offering information on the Guangzhou Baiyun International Airport project covered in this study.

References

Aseel, H, Wasiq K, Abir, H, Panos, L, Ahmed, AL-S & Dhiya, AL-J. (2023). Predicting energy performances of buildings' envelope wall materials via the random forest algorithm. *Journal of Building Engineering,* 69, 106263. https://doi.org/10.1016/j.jobe.2023.10626.

Giuseppe, P, Zhe, W, Abhishek, R, Tianzhen, H & Alfonso, C. (2021). Transfer learning for smart buildings: A critical review of algorithms, applications, and future perspectives. *Advances in Applied Energy,* 5, 100084. https://doi.org/10.1016/j.adapen.2022.100084.

Jiangyan, L, Qing, Z, Xin, L, Guannan, L, Zhongming, L, Yi, X, Kuining, L & Bin, L. (2021). Transfer learning-based strategies for fault diagnosis in building energy systems. *Energy and Buildings,* 250, 111256. https://doi.org/10.1016/j.enbuild.2021.111256.

Tianqi, C & Carlos, G (2016). XGBoost: A Scalable Tree Boosting System. In *22nd International Conference on Knowledge Discovery and Data Mining, ACM SIGKDD 2016* (pp. 785-794). Association for Computing Machinery (ACM).

Xi, F, Guangcai, G, Guannan, L, Liang, C, Weiqiang, L & Pei, P. (2021). A hybrid deep transfer learning strategy for short term cross-building energy prediction. *Energy,* 215, 119-208. https://doi.org/10.1016/j.energy.2020.119208.

Xinyue, H, Xiaoxiao, Z, Junming, L, Yuanjie, P & Xiaopeng, M. (2021). Subjective Evaluation Experiment of Ground-Reflected Glare in Tall and Large Spaces of Airport Terminals. *China Illuminating Engineering Journal,* 32 (5), 130-137. http://10.3969j.issn.1004-440X.2021.05.020.

Yan, S. (2022).The 14th Five-Year Plan for the Development of Civil Aviation. Civil Aviation Administration of China (CAAC). Retrieved November 20, 2023, from http://www.gov.cn/xinwen/2022-01/07/content_5667004.htm.

Yi, H, Xiaohui, H, Hongliang, Z, Wei, Z, Fuyun, Z, Marc, A. S & Yi, M. (2021). Hybrid framework for rapid evaluation of wind environment around buildings through parametric design, CFD simulation, image processing and machine learning. *Sustainable Cities and Society,* 73, 103092. https://doi.org/10.1016/j.scs.2021.103092.

Yi, L, Lin, Z & Rongfeng, C. (2019). Design Research on Spatial Comfort and Green Energy Saving of Airport Terminals. *Architectural Journal,* 09, 18-23.

Yinyi, S., Jinhui, C., Guangzhi, H., & Cunyuan, W. (2023). Prediction and Optimisation of the Typical Airport Terminal Corridor Façade Shading Using Integrated Machine Learning and EvolutionaryAlgorithms. In *28th International Conference on Computer-Aided Architectural Design Research in Asia: Intelligent and Informed, CAADRIA 2023* (pp. 59-68). The Association for Computer-Aided Architectural Design Research in Asia (CAADRIA).

INTEGRATING GENETIC ALGORITHMS AND RBF NEURAL NETWORKS IN THE EARLY DESIGN STAGE OF GYMNASIUM FOR MULTI-OBJECTIVE OPTIMIZATION FRAMEWORK

ZHAOXIANG FAN[1], SHUONING TANG[2]and MENGXUAN LIU[3]
[1,2,3]*Tongji University.*
118346101349@163.com, 0000-0002-2271-8393
[2]*zong20210427@163.com, 0000-0002-9445-9714*
[3]*liumengxuan93@163.com, 0000-0003-4576-7689*

Abstract. The early design phase of the gymnasium's enclosing interfaces directly affects the indoor daylighting and thermal environmental performance. The optimization framework proposed in this study aims to simultaneously balance and optimize conflicting objectives, including the maximum daylight factor (DF), minimum daylight glare index (DGP), and minimum solar radiation (RS) for gymnasium. This approach aims to maximize daylighting performance in hot summer regions while avoiding glare, reducing energy consumption, and ultimately enhancing both daylight comfort and energy efficiency during the sports facility design process. Using the SPEA-2 genetic algorithm, the study explored the Pareto front solutions for three different skylight patterns and established a predictive model for design results based on a Radial Basis Function (RBF) neural network. Compared to traditional Multi-Objective Optimization (MOO) frameworks, this optimization method improves computational efficiency and provides more intelligent decision support for the early-stage design of gymnasiums.

Keywords. Multi-Objective Optimization (MOO), Building Performance Simulation (BPS), Parametric Design (PD), Predictive Model.

1. Introduction

Balancing indoor daylighting, thermal comfort, and energy performance is one of the most critical objectives in the design of sports facilities (Yang, et al., 2018). Existing research has shown that indoor daylighting and thermal environmental performance can directly impact sports experiences (Bale and Vertinsky, 2004), comfort levels (Fantozzi and Lamberti, 2019), and safety and health (Fan, et al., 2023), with continuous effects on the energy consumption of environmental regulation in sports arena interiors (Fan, et al., 2023). However, the impact of different morphological parameters of sports arenas on environmental objectives is often contradictory, and the current challenge lies in ensuring a balance among various targets after complex

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 505-514.

performance goal calculations during the optimization process. Although multi-objective optimization based on building performance simulation has been widely applied, as demonstrated by Wang Pan et al., who established an MOO framework for large sports arena roof structure morphology and spatial layout, evaluating the comprehensive performance of three different roof structures in two cases (Wang, et al., 2019). Nathan Brown from MIT proposed a new approach involving parameterized design space formulation, interactive optimization, and design based on structural morphological diversity (Nathan, 2019). However, existing optimization frameworks often require crossing multiple software platforms and entail significant computational time. This study aims to establish a multi-objective optimization framework to achieve automatic morphological optimization based on performance analysis in the early design stages of gymnasiums. In this process, the SPEA-2 genetic algorithm and Radial Basis Function (RBF) neural network will be introduced for optimization calculations and analysis prediction of result datasets. The framework will predict sports hall morphologies in the early design stages based on performance data, replacing the subjective decision-making process of architects with empirical experience (Figure 1).

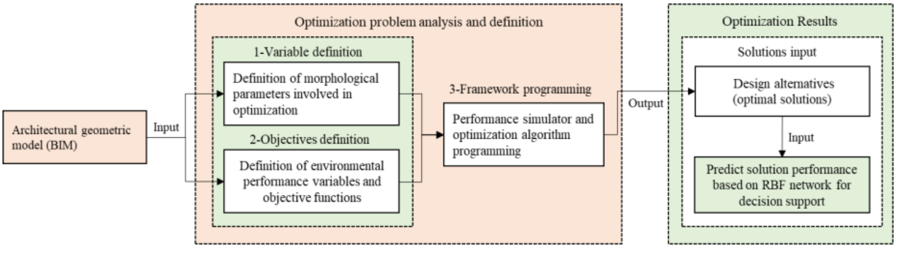


Figure 1. Overall research framework

2. Methodology

2.1. MULTI-OBJECTIVE OPTIMIZATION

In general, the definition of decision variables (inputs) and objective variables (outputs) is a prerequisite and a crucial step in addressing MOO problems. In single-objective optimization, feasible solutions can be ranked based on the value of the objective function, meaning that for any n decision variables $x_1, x_2, x_3, ..., x_n \in X_k$, there exists $f_m(x_n) \geqslant f_m(x_{n-1})$ or $f_m(x_n) \leqslant f_m(x_{n-1})$ to determine their superiority or inferiority. However, in MOO problems, the determination of the final optimization results is based on the relative concepts generated by balancing the contradictions among various optimization objectives. For MOO, with n decision variables, m objective functions, and k constraint conditions, the mathematical function is represented as Eq. (1):

$$\mathrm{Min/MaxF}(x_k) = \begin{vmatrix} f_1(x_1) & f_1(x_2) & \dots & f_1(x_n) \\ f_2(x_1) & f_2(x_2) & \dots & f_2(x_n) \\ \vdots & \vdots & & \vdots \\ f_m(x_1) & f_m(x_2) & \dots & f_m(x_n) \end{vmatrix}, x \in X \quad (1)$$

Where, $F(x_k)$ represents the set of objective functions in the objective function space, while X is the set of decision variables that includes the space containing the optimal solution. Here, $f_m(x_n)$ denotes the individual optimization objective functions, and x is a subset of decision variables. The results of Multi-Objective Optimization (MOO) are categorized into two types: non-dominated solutions and dominated solutions. The non-dominated solutions of the final population are also known as Pareto front solutions, representing the alternative solution set under the combination of decision variable parameters obtained through the algorithm's multi-objective optimization.

2.2. RBF NETWORK

The RBF network is an artificial neural network that uses radial basis functions as activation functions. The output of an RBF network is a linear combination of the radial basis functions of the input and the parameters of the neurons (Chen, Cowan, and Grant, 1991). It can approximate any continuous function with arbitrary precision and is particularly well-suited for solving classification problems. The functional representation of the RBF network is given by Eq. (2):

$$\mathrm{RBF}(x) = \sum_{i=1}^{h} w_{ij} \exp\left(-\frac{1}{2\sigma^2}||x - c_i||^2\right) \quad j = 1,2,3,\ldots,n \quad (2)$$

Where, $\mathrm{RBF}(x)$ represents the output of the radial basis function, where x is the input data vector, c_i represents the center vector of the radial basis function, i denotes the number of input signal source nodes, j represents the number of neurons in each hidden layer, and σ represents the width parameter of the radial basis function, controlling the extent of the function's expansion. RBF is a feedforward neural network with a structure divided into three layers: 1) Input layer, consisting of signal source nodes; 2) Hidden layer, where the transformation function is RBF, and the number of units is determined by the requirements of the problem; 3) Output layer, responding to the input patterns. The transformation from the input space to the hidden layer space is nonlinear, while the transformation from the hidden layer space to the output layer space is linear (Figure 2).

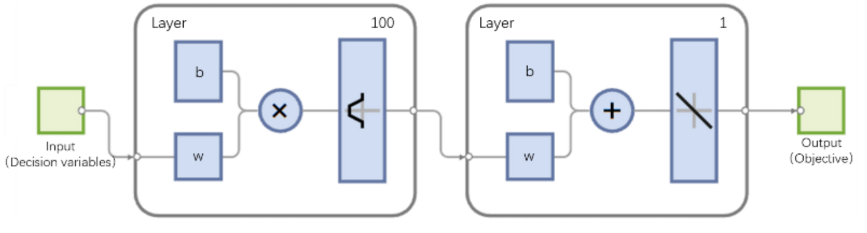


Figure 2. The structure of RBF neural network.

3. Optimization process

The research focuses on the main decision variables of the gymnasium interface morphology parameters, conducting MOO studies with daylighting factor (DF), average daylight glare probability (DGP), and average cumulative solar radiation (RS) as objective variables. This is carried out concerning key interface morphology parameters, including skylight, sidelight, and shading parameters. The aim is to explore different combinations of interface opening and closing degree parameters under the consideration of various daylight and thermal environmental performance goals, along with their performance distribution characteristics. The basic framework of this part is illustrated in Figure 3.

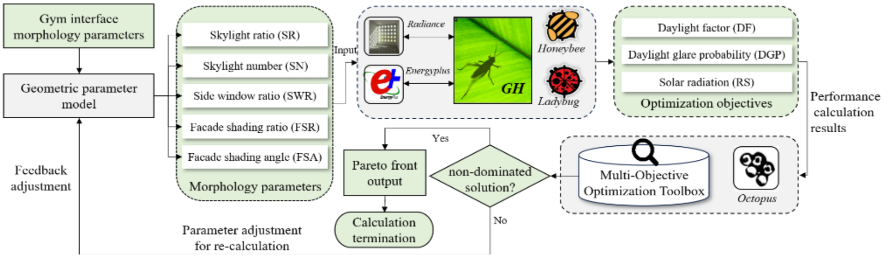


Figure 3. MOO framework for gymnasium interface morphology parameters

In this segment of the study, the gymnasium interface morphology parameters, namely skylight ratio (SR), skylight column number (SN), side window ratio (SWR), facade shading ratio (FSR), and facade shading angle (FSA), will serve as the input genes for genetic algorithms to form the decision space. The optimization will explore the results of the combination of morphology parameters based on DF, DGP, and RS. Furthermore, a comparative analysis of the changing trends between different generations of optimized solution sets will be conducted. This aims to explore and compare the distribution characteristics and variation patterns of different morphology parameters and environmental objectives within the optimization solution set.

3.1. DECISION VARIABLES

In this study, the input decision variables for the interface morphology parameters are as follows: x_1—SR, x_2—SN, x_3—SWR, x_4—FSR, x_5—FSA. The output variables are the DF of the sports hall's activity field, the average DGP from unfavourable viewpoints, and the average cumulative RS in the sports hall.

The study is divided into an optimization process for flat skylights, strip skylights, and point skylights. The decision variables X correspond to the numerical models of gymnasium under different opening and closing forms, where the model supports the continuous combination of values for X within different ranges of morphological variables. Thus, it forms the search space for different decision variables. Different parameter combinations correspond to a relatively extensive set of design configurations (Table 1).

Type	Acronym	Variable Names	Unit	Range
Decision Variables (Inputs)	SR	Skylight ratio	--	0.1-1.0
	SN	Skylight number	--	1-10
	SWR	Side window ratio	--	0.1-1.0
	FSR	Façade shading ratio	--	0.1-1.0
	FSA	Façade shading angle	°	0-90
Objective Variables (Outputs)	$DF_{sports\ hall}$	Daylight factor	%	--
	$DGP_{ave\text{-}court}$	Average daylight glare probability	--	--
	$RS_{ave\text{-}court}$	Average cumulative solar radiation	kWh/㎡	--

Table 1. Decision variables (inputs) and objective variables (outputs)

3.2. OPTIMIZATION OBJECTIVES

In this optimization process, three objectives are set as the DF, the average DGP, and the average cumulative RS within the sports hall. Three conflicting objective functions are used to define the trade-off between maximizing the DF while minimizing both the DGP and the cumulative RS intensity. Therefore, in the optimization objective setting, the objective function $F_1(y)$ corresponds to the negative maximum value of the daylighting factor, $F_2(y)$ corresponds to the minimum value of the daylight glare probability, and $F_3(y)$ corresponds to the minimum value of the cumulative radiation in the sports hall, as shown in Eq. (3)-(5):

$$F_1(y) = -DF_{max}(x) \quad (3)$$
$$F_2(y) = DGP_{min}(x) \quad (4)$$
$$F_3(y) = RS_{min}(x) \quad (5)$$

4. Results

4.1. DECISION VARIABLES DISTRIBUTION

In this section, a MOO exploration was conducted for the combination pattern set of sports arena interface parameters (side facade, roof interface). The optimization process was carried out separately for three typical skylight forms in the gymnasium: strip skylight, point skylight, and flat skylight. Using the intelligent driving of the SPEA-2 genetic algorithm toolbox, the study independently performed performance calculations and iterative comparisons for multiple scenarios using Ladybug tools. During the calculation process, the non-dominated solution ratios for strip, point, and flat skylight patterns all exceeded 80% by the 20th generation. The optimization solutions from the 20th generation were selected as the final alternative solution set.

From the overall distribution of optimization solutions, the three skylight forms exhibit similar performance in terms of DF, DGP, and cumulative RS, with no significant differences in performance distribution. Figures 4 depicts the distribution of Pareto front solutions' objective values for point skylight, one typical pattern under different interface opening and closing parameter combinations. Partial geometric models of typical optimized solutions under different skylight ratios are also presented, ranging from left to right, corresponding to different optimal solutions and combination differences of other morphological parameters as skylight ratio increases. The lower part shows the parallel set plot of the correlation between optimized solution morphological parameters and target variables, as well as a typical set of optimized solutions classification according to skylight ratio.

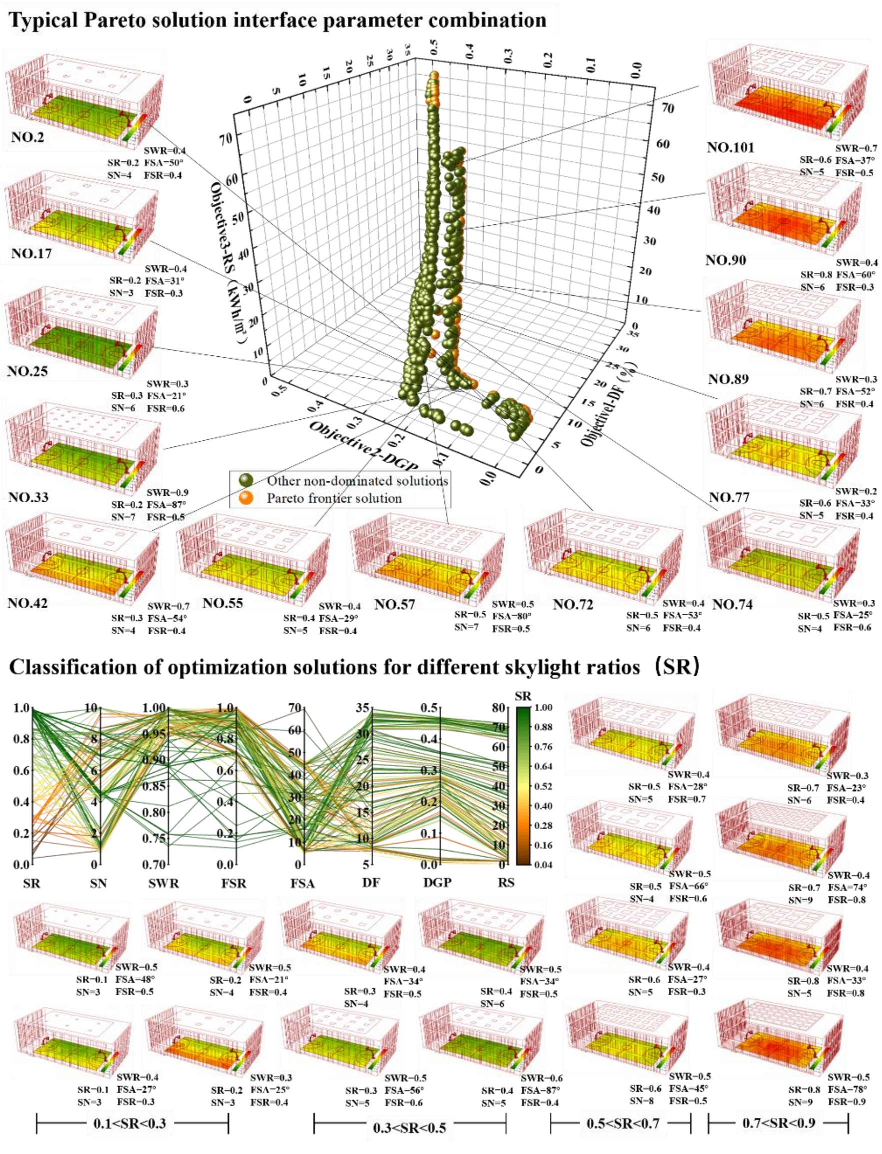


Figure 4. Parameter distribution and classification of optimal solutions under point skylight pattern

4.2. OBJECTIVE VARIABLES DISTRIBUTION

Figure 5 illustrates the distribution and changing trends of objective values of Pareto front solutions in key generations during the optimization process for different skylight patterns. For the strip skylight pattern, with the increase in the number of iterations, different objectives show different changing trends. Considering the range of optimized solutions at the 20th generation, the 10th, 15th, and 20th generations are selected as key generations for comparison. Taking the daylighting factor as an

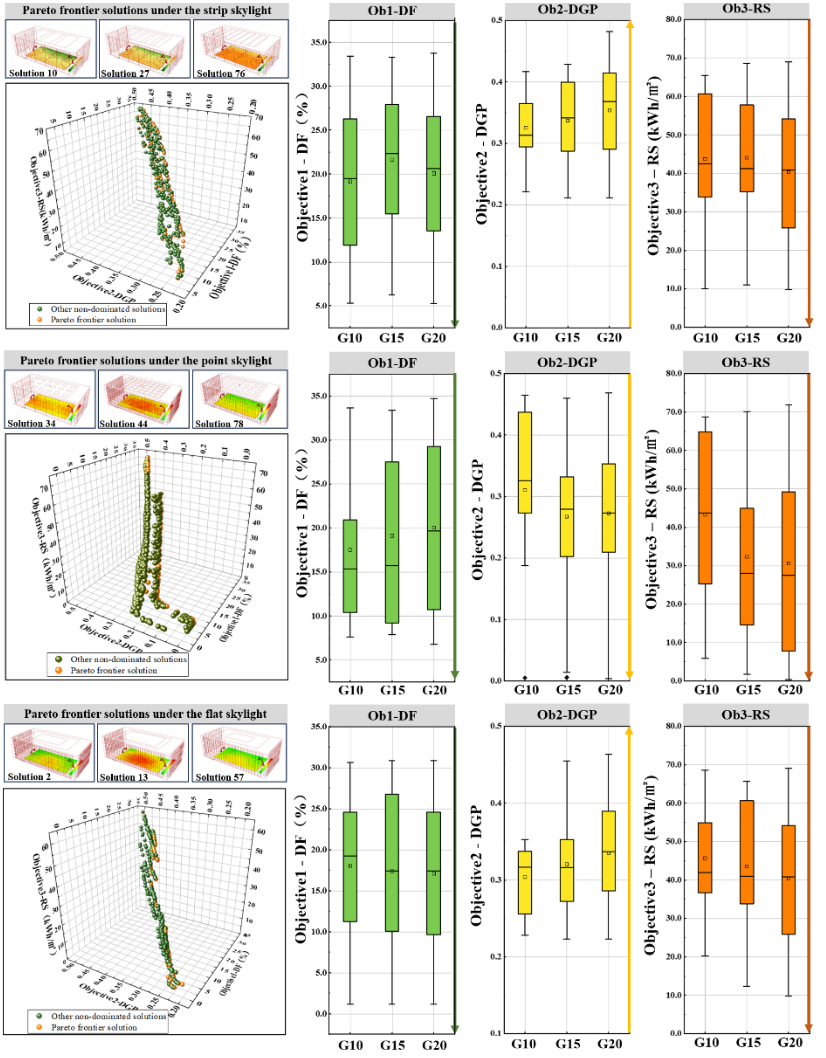


Figure 5. Distribution of objectives in different generations of optimization solutions

example, the range of the Pareto front solutions is primarily between 12% and 26% at the 10th generation, expanding to 15.5%-27.5% at the 15th generation, and finally ranging between 13.5% and 26% in the final generation (20th). The average daylighting factor of the optimized solutions increased from 18% to 20%, representing an improvement of approximately 11%.

In addition, clustering analysis was conducted on the morphological parameters of different schemes. A morphological parameter map was then generated based on the Euclidean distances between parameters (as shown in Figure 6). This facilitates designers in making better choices for further development.

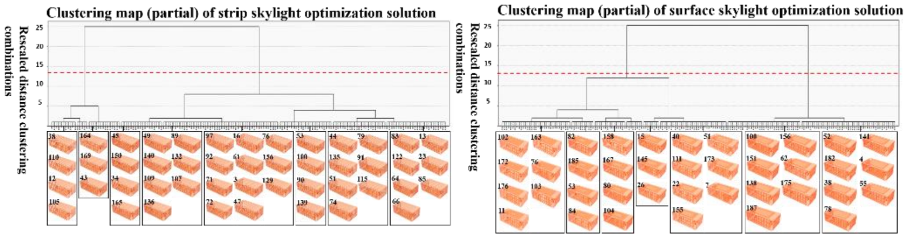


Figure 6. Optimized solution morphological parameter clustering map

4.3. PREDICTION MODEL

In the RBF optimization model, the set of shape parameters outputted by the MOO serves as the independent variables (inputs), while the environmental target dataset serves as the dependent variables (outputs). The hidden layer had 129 neurons; the radial basis expansion rate was set at 500. Learning was conducted separately for different skylight patterns. In this process, 70% of the data samples were used as training samples, and 30% were used for testing. Figure 7 illustrates the prediction results for

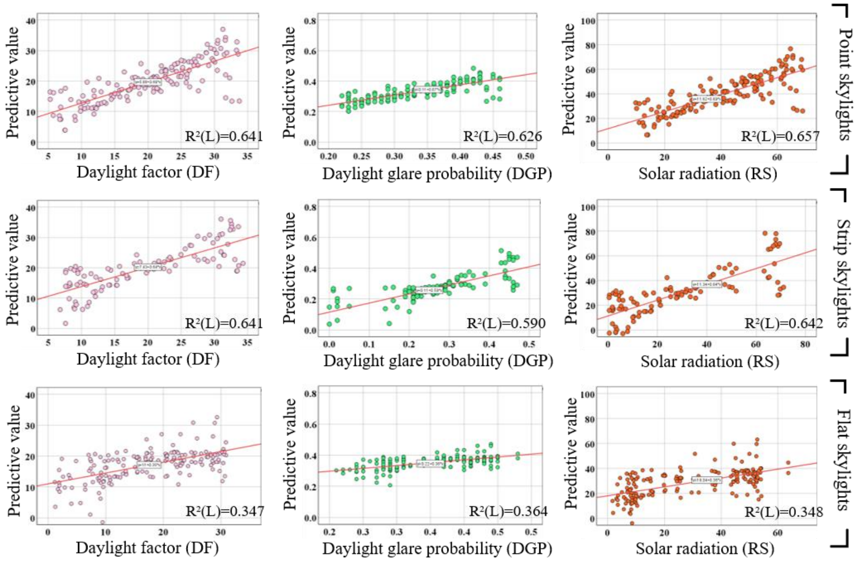


Figure 7. The comparison between the actual values and predicted values of the prediction model

different performance objectives. The prediction model for the point skylight pattern showed the highest fitting degree, achieving an explanatory level (R^2) of 65.7% for solar radiation. In contrast, the prediction model for the flat skylight pattern had a relatively lower fitting degree, approximately 35%. This may be related to overfitting caused by a high expansion rate. Figure 8 shows the predicted RMSE of strip skylights for different objectives, among which the prediction error of daylight comfort is the lowest, with RMSE% of about 17%. In further research, different expansion rates will be adjusted for different morphological parameter sets to improve the explanatory level of the prediction model.

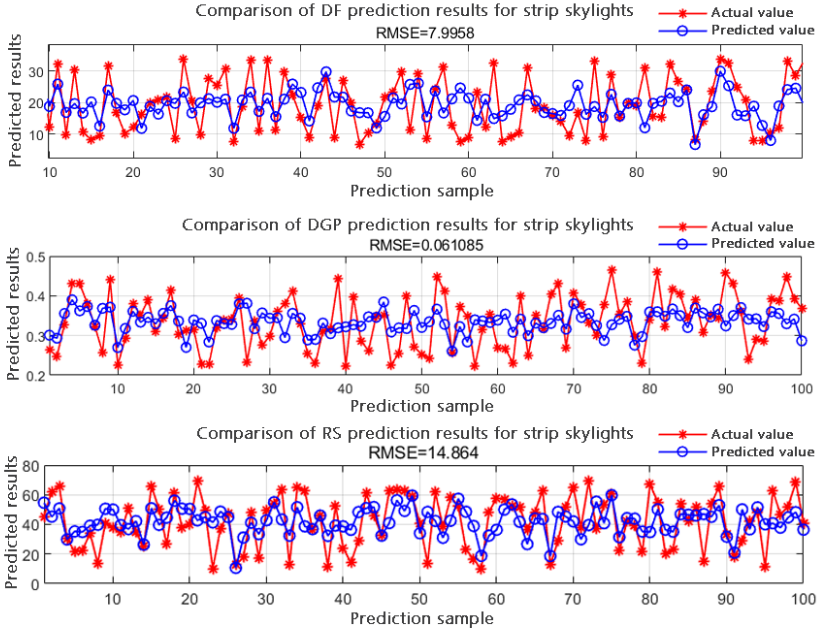


Figure 8. RMSE comparison of strip skylight prediction models

5. Discussion

The study tested the effectiveness of BPS-based MOO in enhancing the daylighting and thermal performance of facades and skylights in the early stages of stadium design. Compared to manual testing methods, this approach is more continuous, explores more comprehensive solution parameters, and can form predictions for new designs based on existing optimization solution sets. However, the study still has some limitations, mainly reflected in:

1) The prototype of the study only optimized the facade and skylights of a square sports hall, lacking discussions on other irregular-shaped gymnasiums. In future research, the scope of parametric prototypes will be expanded to adapt to a wider range of gymnasium optimizations. Additionally, more shading and skylight opening styles should be refined for discussion and comparison to expand the search range of decision variables.

2) The study tested the gymnasium facade and skylights as an integrated system, using a standard CIE sky model. It ignored the impact of different orientations on sunlight and radiation, which typically varies based on the geographic location and climate zone of the case. This difference can be addressed by further comparing simulation results for different directions (east, west, south, north) to provide more refined optimization results.

3) The prediction model in this study was mainly achieved through data modelling. In future research, exploring the correlation mechanism between optimization result maps and morphological variable images using GNN tools can lead to more intuitive and efficient predictions of gymnasium design performance. This can reduce the time cost of long-term simulation and provide a more visual design assistance for the green design and sustainable optimization of sports buildings.

6. Conclusion

This study proposes a multi-objective optimization and prediction method based on genetic algorithms and RBF neural networks through the optimization of morphological parameters of community sports stadium interfaces (facade and skylights). The aim is to collaboratively enhance the early-stage design of stadiums for daylighting, daylight comfort, and energy load optimization. This approach seeks to improve the passive environmental regulation capability of design outcomes. The decision variables in this study include the stadium skylight ratio (SR), skylight numbers (SN), side window ratio (SWR), facade shading ratio (FSR), and facade shading angle (FSA). The optimization objectives are daylight glare probability (DGP), daylight factor (DF), and daylight radiation load (RS). The study compares the multi-objective optimization results of environmental performance in three different skylight modes of sports hall. Additionally, it establishes a design prediction model based on RBF neural networks for the Pareto front solution set. This method provides a technical framework and decision support for intelligent prediction and automatic optimization of environmental performance in the early-stage design of gymnasiums.

References

Yang, D., Ren, S., Turrin, M., Sariyildiz, S., & Sun, Y. (2018). Multi-disciplinary and multi-objective optimization problem re-formulation in computational design exploration: A case of conceptual sports building design. Automation in Construction, 92, 242-269.

Bale, J., & Vertinsky, P. (Eds.). (2004). Sites of sport: Space, place and experience. Routledge.

Fantozzi, F., & Lamberti, G. (2019). Determination of thermal comfort in indoor sport facilities located in moderate environments: An overview. Atmosphere, 10(12), 769.

Fan, Z., Liu, M., Tang, S., & Zong, X. (2023). Integrated daylight and thermal comfort evaluation for tropical passive gymnasiums based on the perspective of exercisers. Energy and Buildings, 300, 113625.

Fan, Z., Liu, M., & Tang, S. (2022). A multi-objective optimization design method for gymnasium facade shading ratio integrating energy load and daylight comfort. Building and Environment, 207, 108527.

Pan, W., Turrin, M., Louter, C., Sariyildiz, S., & Sun, Y. (2019). Integrating multi-functional space and long-span structure in the early design stage of indoor sports arenas by using parametric modelling and multi-objective optimization. Journal of Building Engineering, 22, 464-485.

Brown, N. C. (2019). Early building design using multi-objective data approaches (Doctoral dissertation, Massachusetts Institute of Technology).

Chen, S., Cowan, C. F. N., & Grant, P. M. (1991). Orthogonal least squares learning algorithm for radial. IEEE Trans. Neural Netw, 2, 302-309.

MAKING A CASE FOR DESIGN ANALYTICS:

Complementing Designers' Toolbox for Data-Informed Creative Decision-Making

HALIL ERHAN[1] and AHMED M. ABUZURAIQ[2]
[1,2]*Simon Fraser University.*
[1]*herhan@sfu.ca, 0000-0002-7630-6351*
[2]*aabuzura@sfu.ca, 0000-0002-3604-7623*

Abstract. "Design Analytics" is a methodological approach for engaging with design form and performance data to guide design space exploration. The goal is to integrate design decision-making with data analytics through interactive visualisations. This paper introduces the approach grounded in three distinct patterns of design exploration: parallel development, solution exploration, and collaboration. These patterns are demonstrated through system case studies developed using a design study method, highlighting diverse opportunities and challenges in using design data. In each case, functional prototypes of the systems are presented. We propose a research agenda for this system-agnostic approach, offering a perspective on restructuring design with new design systems to enhance design computing.

Keywords. data inform creativity, Visual analytics, Design Analytics, data-driven design, design collaboration, design space navigation.

1. Introduction

In this paper, we introduce "Design Analytics" (DA), a research program that focuses on developing methods for leveraging design data to improve decision-making in the specification, generation, and evaluation of design alternatives, especially in the early design phases. Design Analytics integrates design exploration with data analytics, drawing from Visual Analytics techniques (Cook and Thomas, 2005). Our focus is on three related areas of design exploration: parallel development, solution exploration, and collaboration within the context of data. We present three case studies to explore the opportunities and challenges of leveraging design data, following a design study approach (Sedlmeir et al., 2012) and resulting in functional prototypes.

Considering that design representation is digitally created, it is reasonable to assert that a design be treated as data, extending beyond computable descriptions of an artefact. For instance, Building Information Modelling captures layers of design data at varying levels of abstraction, amenable to computing diverse performance factors like sustainability or cost. Although traditionally limited to specifications, contractual documents, and design models, design firms generate more extensive types and

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 515-524.

volumes of data. Therefore, they require enhanced tools to derive valuable insights about their designs (Deutsch, 2015; Loyola, 2018).

Through multiple case studies, we identified and addressed specific gaps around using data to inform different design activities. This includes integrating real-time evaluation and comparative analytics of designs within a modelling environment (D-Flow), navigating through large generative design datasets with interactive data visualisation and analysis (D-Sense, or DesignSense), exploring novel interactions for driving parametric exploration with data (D.Star), collaborating in an asynchronously and data-informed fashion with project stakeholders in the early design stages (D-Art). These cases operate on the premise that most design representations are digitally created, including form and performance data. Such representations encompass geometric and substantial amounts of numeric and textual data in structured and unstructured forms, describing processes and structures. We chose cases as examples to narrow the paper's scope rather than presenting exhaustive system solutions. As part of our contributions, we outline a research agenda for Design Analytics (DA) that positions data as an integral (and system-agnostic) component of different design workflows. We envision the possibility of a new perspective on restructuring design by emphasising the impact of data and data-focused design systems.

2. Developing Design Analytics Interfaces

Designers work on multiple alternatives as a core aspect of their creative process (Woodbury and Borrow, 2006), and this process is increasingly being complemented with data. Developing tools that synergise these two elements—creating multiple design alternatives and data-informed design—is an ongoing endeavour in computational design research. One of the goals of this integration is to enhance designers' ability to navigate and uncover the implicit design space, a concept articulated by Aish and Woodbury (2005). Such tools aim to expand the horizons of design possibilities, allowing designers to explore and realise more performant designs that may not be immediately apparent in the conventional design process. It's instructive to consider how those tools are designed since they are part of the structure of a task environment that influences design problem spaces and, consequently, the strategies designers apply when exploring alternatives (Simon, 1996).

In current design task environments, the computational tools delimit working with alternatives and design data (Kasik, Buxton, and Ferguson, 2005; Terry et al., 2004): they are pressed for features to support the generation and management of multiple solutions with their data collectively (Lunzer and Hornaek, 2008; Schneiderman, 2007; Bilal et al., 2016; Touloupaki and Theodosiou, 2017). While most tools focus on form generation—in architectural design—specification evaluation and collaboration aspects are overlooked for several reasons (Kasik, Buxton, and Ferguson, 2005). For example, generating or selecting the designs based solely on their performance is challenging for two reasons. First, the performance usually reflects only partially and far from a complete set of all the design concerns, even if they are the most critical ones. Second, not all criteria can be easily quantified; some are tacitly known to the designers, such as aesthetics. Making sense of design data requires various visual, logical, and temporal structuring in different types of representations.

Computational design tools are creativity-support tools and are expected to enable 'exploratory processes,' facilitate 'collaboration,' manage 'rich history keeping, and be useful for novices and experts (Schneiderman, 2007). In addition, these tools should be 'engaging,' balance 'effort-reward trade-offs,' provide 'transparency' towards achieving tasks and be 'expressive.' For example, the ability to generate solutions creates a need for directly managing, sorting, filtering, and selecting alternatives. Further, evaluating alternatives involves multiple stakeholders with diverse interests, which need means to collaborate and share decisions. Both essentially rely on making design data accessible to decision-makers (Bilal et al., 2016).

3. Adopting a Design Study Methodology

A challenge for DA is to devise approaches and practices for representing, visualising and interacting with design data and workflows. To this end, we use the design study methodology outlined by Sedlmeir et al. (2012), which originates from information visualisation. This methodology is particularly adept at fostering the development of visualisations through an iterative process of design and evaluation. It emphasises a comprehensive approach, starting with analysing real-world problems identified by domain experts and then designing an interactive visualisation system to address these problems. The methodology then moves to validate the design and reflect on the gathered insights to refine guidelines for visualisation design.

The practical advantage of this methodology lies in its structured approach to solution development, allowing us to identify and refine solution features iteratively. It consists of nine stages, distributed across three main phases: the Precondition phase (learning, winnowing, and casting, selecting collaborators), the Core phase (discovery, design, implementation, and deployment), and the final Analysis phase, which involves reflection and evaluation. Each phase is designed with its specific evaluation strategy, ensuring a cohesive and interconnected process throughout all stages. To complement this methodology, we integrated agile and use-case-driven software development. This combination provided a structured yet flexible framework for prototyping, allowing us to develop and refine high-level system requirements in tandem with prototype development. These requirements also served as a basis for the formative evaluations.

Our approach also included extensive research on interactive systems that support data-informed design. In parallel, we engaged with our industry partners to gain insights into their workflows and specific needs in using data in design. This dual approach of theoretical research and practical engagement enabled us to apply learned lessons directly into developing our prototypes, which were enhanced with interactive visualisations to facilitate effective sense-making, as per Cook and Thomas (2005).

Case studies can help with problem characterisation and abstraction, one of the three fundamental contributions of design studies (Sedlmeir et al., 2012). The utility of problem characterisation in the research field lies in its characterisation of patterns for interactive systems requirements on which future efforts can be modelled and, to a lesser extent, the validation of the prototype produced. The cases we present emphasise core stages to provide a structure for research on DA. New technology enables new tasks, altering existing tasks. The discovery and development of these tasks are integral aspects of this problem-driven research. For example, DA should create opportunities for a thorough design evaluation, leading to changes in the approval process.

4. Design Analytics Cases: Generation, Exploration, Collaboration

For structuring the DA case studies, we will follow a workflow consistent with how designs are created in real-world scenarios. In this workflow, the data as input to a design situation is assumed to be given as design constraints or preferences, represented as part of digital models (as parameters or direct input) or implicitly observed by the designers. By taking the risk of simplification for clarity, we divide the workflow into three phases: generation of design alternatives with their form and performance data (both algorithmically and manually), evaluation and filtering of design alternatives, and data-informed collaborative decision-making (Table 1).

Table 1. Design data flow through design exploration and the example tools from practice and research. The tools must be seamlessly connected for data flow coordination and support uninterrupted decision-making.

Phase	Design Analytics	Tool
Generation of design alternatives	Designers generate solutions algorithmically, manually (e.g., by a directly interactive, or mixed, initiative CAD tools). The solutions consist of form and performance data at various levels of abstraction.	D-FlowUI (Erhan et al. 2020) directly interactive modelling and D-CAT (Zarei, 2021) for comparative data analysis D.Star (Mohiuddin and Woodbury, 2020) generative design in a design gallery interface.
Evaluation and Filtering	Generated design alternatives are explored by considering their form and performance metrics. Data used to assess model integrity.	D-Sense by Abuzuraiq (2020), Design Explorer by Thornton Tomasetti (2019) or DreamLens by Matejka et al. (2018)
Collaborative Decision-Making	Curated alternatives are shared with design stakeholders to initiate data-informed discussion and maintain synchronous or asynchronous collaboration	D-ART (Osama and Erhan, 2022) collaborative design-data analytics considering form and performance data presented on interactive data visualizations.

4.1. GENERATION OF DESIGN ALTERNATIVES

Design generation has been widely studied since the availability of computational methods to designers. Detailed reviews of generative or algorithmic design methods can be found in other literature, e.g., (Woodbury, 2010; Krish, 2011; Caetano, Santos and Leitão, 2020). The common characteristic of these methods is their capability to use a parametric description of a design model to generate a large set of alternative solutions. Most such techniques focus on developing ‘form’ with some degree of optimisation considering design criteria. Design generation using parametric models can also occur using design data directly instead of geometry or changing parametric definitions. For example, D.Star by Mohiuddin and Woodbury (2020) uses a parallel coordinate interface (Figure 1) serving two purposes: visualise input and output parameters and generate variations by using the graphs or tracing data points directly on the interface. The form and additional design data are computed and visualised on data graphs. D.Star interfaces combine both DA and design exploration tasks.

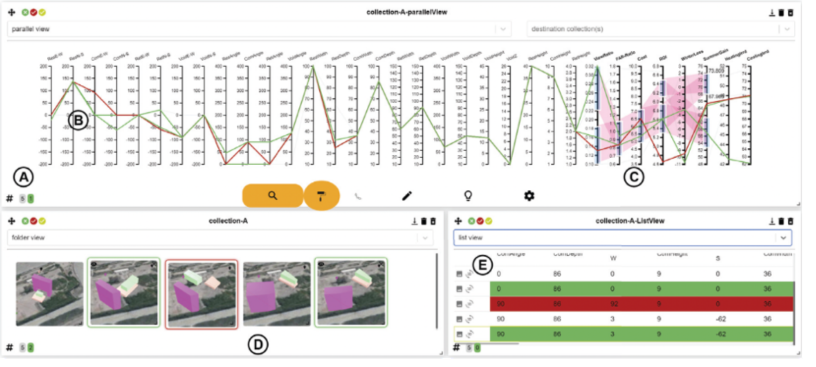


Figure 1. D.Star generates designs (A) on parallel coordinates to assign values and control ranges (B and C). Alternatives are displayed as 3D forms (D) and in a tabular layout (E).

Setting up parametric systems is labour-intensive and cannot be cost-effective in every design case. As a more common method, directly interactive modelling focuses on single-state designs that are amenable to refinement more than exploration, hence the lack of support for the parallel creation of solutions as demonstrated by expert designers (Woodbury, 2010; Erhan, Salmasi, Woodbury, 2010; Kolaric, Erhan, Woodbury, 2017). Design data should be assessed and inform decision-making in design exploration. D-FlowUI provides interfaces for interactively exploring design solutions (Figure 2-left) and D-CAT with D-FlowUI (Figure 2-right) to enable comparative DA (Erhan et al., 2020; Zarei et al., 2021). They are built on a combination of an actively used CAD tool, Rhino, and its add-on for parametric design, Grasshopper. The data visualisations demonstrate close-to-real-time design data, particularly in the concept development phases.

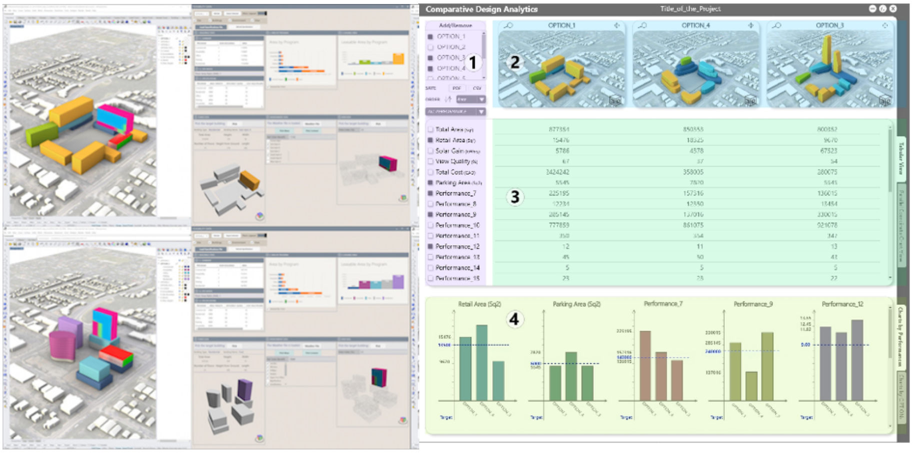


Figure 2. D-FlowUI shows two alternatives to a mixed-use building complex (Top). D-CAT (bottom) proposes interactive data visualisations to focus on performance criteria.

4.2. REVIEW, SELECT, SORT, AND FILTER WITH DATA

Tools for sifting through the generated solutions have been proposed, e.g., Design Explorer by Thornton Tomasetti (2019) and Dream Lens (2018). We developed D-Sense to navigate design space through coordinated and interactive data visualisations. Unlike the others, D-Sense can perform similarity- and set-based interactions (Abuzuraiq, 2020; Abuzuraiq and Erhan, 2020) (Figure 3).

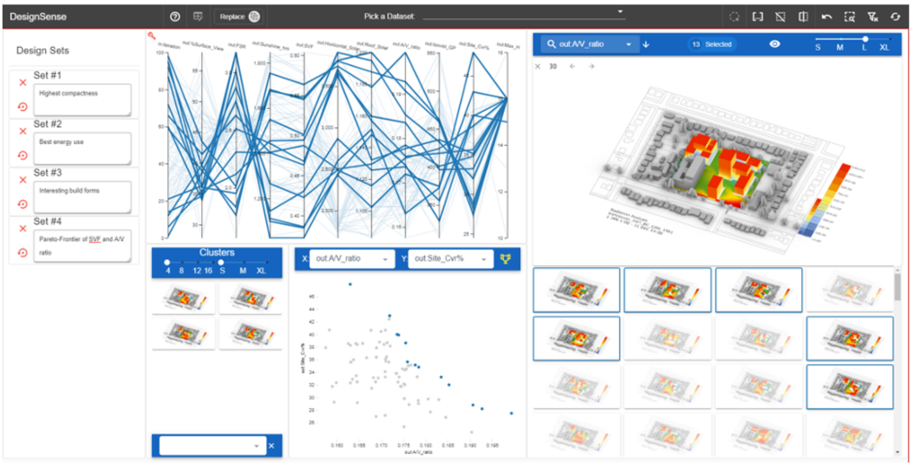


Figure 3. D-Sense (or Design Sense) is an interactive DA dashboard that can adapt any design data type (including form data and images) for exploring alternatives (Abuzuiraq, 2020).

A typical generative design dataset is multi-dimensional, including the input and performance metrics. Both are often quantitative, maybe occasionally categorical or ordinal. 2D images and 3D geometry models represent form. Other images can also be used to visualise the results of performance simulations, such as heatmap colouring to indicate how much sunlight is received. Generated alternatives must be evaluated in a performance-driven process (Erhan, Wang, Shireen, 2014; Anton and Tanase, 2016). The performance computation expects specific geometric fidelity in particular data structures; any inaccuracy, incompleteness, or error will result in unreliable or invalid design data. Considering these, the developers and users of parametric models should be aware of the importance of building reliable, scalable, and reusable models. Therefore, the design models and their parametric descriptions should be tested for their readiness for analysis (Figure 4).

Given the challenges for a large-scale generation of solutions through algorithmic methods, data visualisations can reveal form-performance inconsistencies, avoiding premature commitment. For example, we analysed ten different versions of the parametric tower designs in a workshop. After creating visualisations of the initial 250 designs in Tableau (2022) and using our custom analysis tools, we discovered unexpected patterns in building performance (Figure 5). Exploration involves design space segmentation and simplification for finding satisficing solutions by deciding on trade-offs and recording insights for sharing design decisions. In D-Sense, designers can compose sets of alternatives to reduce, eliminate or mark solutions for further analysis. Selections can be kept as a subset of solutions to be revisited or used for detailed analysis with set operations (e.g., the intersection of two sets of alternatives).

In such a scenario, D-Sense enables the selection of data dimensions and segments them into clusters based on their similarity. The evaluation of design alternatives generates design data as assumed earlier; each design act, whether related to specification, generation, or evaluation, is part of DSE.

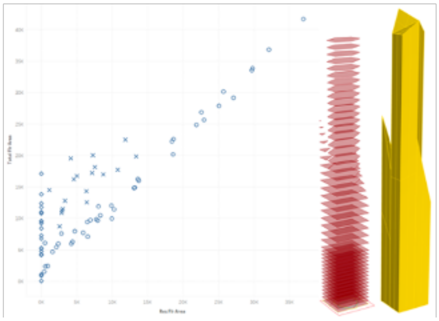


Figure 4. Left: Scatter plot of Total Floor Area vs. Residential Floor Area. Crosses show unexpected outliers. The vertical strip (diamonds) on the marks on the vertical axis are alternatives with no residential floors. Right: Floor planes with random segmentation or orphan floor plates.

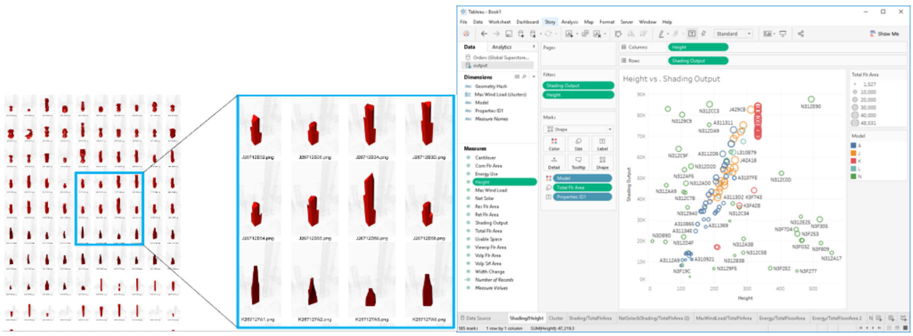


Figure 5. Performance and geometry present two different design aspects that complement each other. Evaluations capture only a fraction of the criteria that may be in action, while geometry at low fidelity can be misleading without details.

4.3. DATA-INFORMED COLLABORATIVE DESIGN-EVALUATION

Evaluating designs, as an integral part of the design process, involves multiple stakeholders with diverse backgrounds. Although there are computational systems for supporting evaluation tasks, they are either highly specialised for designers or configured for a particular workflow with limited functions. There may be a need to share relevant design data to inform collaborative decision-making. Also, stakeholders' feedback creates another layer of design data as input to the design.

To support stakeholders in reviewing a set of curated design alternatives, interacting with each other, and providing feedback, we propose the Design Alternatives Reporting Tool (D-ART) is a design data dashboard akin to a social web app. It aims to complete the data-informed design decision cycle (Alsalman and Erhan, 2022) (Figure 6). It aims to enable the presentation of each design alternative with personal visualisations of data while engaging stakeholders in conversation and feedback sharing. The stakeholders can visually compare and provide input on design alternatives to their data. Being an online platform decouples it from the design systems to manage the interface complexity and to accommodate different stakeholders with different interests and backgrounds. The stakeholders can interact with each other when they compare design alternatives.

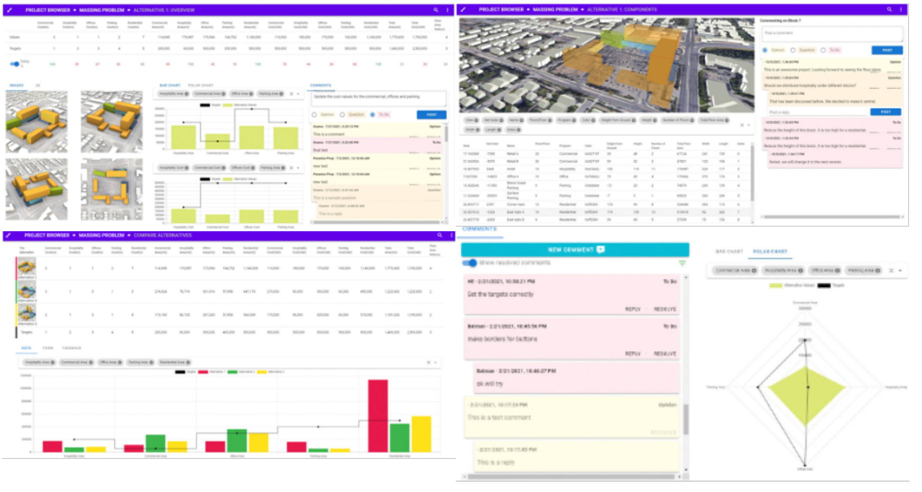


Figure 6. D-ART presents design data in alternatives, component, and comparison views.

5. Discussion and Conclusions

Design Analytics aims to engage in data-informed decision-making on interactive visualisations throughout the design decision-making lifecycle. With the decision-making cases and the tools we proposed, we demonstrated several ways for how DA can better serve in design workflow and create better-built environments with the caveat that neither the cases nor the tools are complete nor may be ideal. In this position paper, we also presented a set of tools and a conceptual process workflow that centres DA as a core activity in data-informed design beyond the conventional approaches with limited interactivity between the data, designers, and design stakeholders.

An example design workflow is presented in Figure 7, where a parametric model is specified to set the design constraints and preferences. Once the model is adequately defined, the system can help generate design alternatives. The designer employs DA to evaluate the alternatives and selects a subset that aligns with the predefined goals. At this stage, the designer can either regenerate additional alternatives or continue working with the existing ones. If the decision is to proceed and the designer deems the alternatives promising, it is necessary to choose metrics for performance assessment, e.g. using simulation or surrogate models. The designer can select plausible alternatives for further consideration upon aggregating the performance data with their design data counterparts. If the performance of the alternatives necessitates improvement, the system regenerates additional alternatives. Conversely, if the performance is deemed satisfactory, the system archives the selections in a repository, which can also serve as a training model. This can lead to generating similar or exploring entirely distinct design possibilities. Through the design study method, each of the presented tools went through several iterations of formative evaluation to better answer the question of what the other cases are or how such tools can improve the design workflow. However, we predict that there will be emerging issues as such tools become part of the real world.

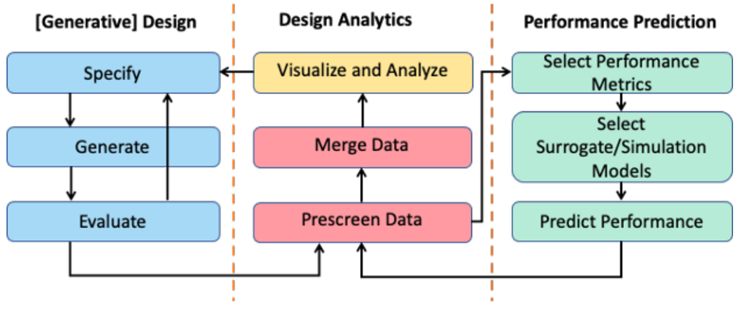


Figure 7. A conceptual model of DA in generative design and performance prediction.

We envision two salient obstacles. The first is about the difficulty of disrupting the established processes or changing beliefs. In time, the demonstration of effective use of design data can gradually trigger changes. The second is developing suitable tools for design tasks considering human-cognitive capabilities. Design-data dashboards can become quickly overpopulated with data visualisation and complex interaction techniques that can affect visual encoding. For example, while D-Sense is highly specialised and affords complex operations on design data, its capability should be revisited to understand how it can scale to respond to different design scenarios with varying levels of abstraction. D-ART presents opportunities for design-data democratisation for projects involving the public or input from different communities. Although the dashboard view may be simplified, transitioning from one view to another requires focused attention. We recommend striving for simplicity and clarity while making DA accessible to different stakeholders.

References

Abu Zuraiq, A. (2020). DesignSense: A Visual Analytics Interface for Navigating Generated Design Spaces (Master's thesis). *Simon Fraser University, Surrey*, BC, Canada.

Abu Zuraiq, A., & Erhan, H. (2020). The Many Faces of Similarity: A Visual Analytics Approach for Design Space Simplification. In *25th International Conference of the Association for Computer-Aided Architectural Design Research*, Thailand.

Aish, R., & Woodbury, R. (2005). Multi-level interaction in parametric design. In SmartGraphics, *5th International Symposium*. Springer.

Alsalman, O., & Erhan, H. (2022). D-ART for Collaboration in Evaluating Design Alternatives. *International Journal of Architectural Computing*, Special Issue 2021.

Anton, I., & Tănase, D. (2016). Informed Geometries: Parametric Modelling and Energy Analysis in Early Stages of Design. *Energy Procedia*, 85, 10-1016.

Bilal, M., Oyedele, L. O., Qadir, J., et al. (2016). Big data in the construction industry: A review of present status, opportunities, and future trends. *Adv. Eng. Inf.,* 30(3), 500-521.

Caetano, I., Santos, L., & Leitão, A. (2020). Computational design in architecture: Defining parametric, generative, and algorithmic design. F of Architectural Rech.., 9(2), 287-300.

Cook, K. A., & Thomas, J. J. (2005). Illuminating the path: The research and development agenda for visual analytics. *Pacific Northwest National Lab* (PNNL), Richland, WA, US.
Deutsch, R. (2015). Data-driven design and construction: 25 strategies for capturing, analysing and applying building data. John Wiley & Sons.
Erhan, H., Abuzuraiq, A., Zarei, M., Alsalman, O., Woodbury, R., & Dill, J. (2020). What Can Data Reveal About Your Design Model? A Case Study on Reliability and Validity. In *25th International Conference of the Association for CAADRIA*.
Erhan, H., Wang, I., & Shireen, N. (2014). Interacting with Thousands: A Parametric-Space Exploration Method in Generative Design. *In Proc. of the 34th ACADIA*, CA., USA.
Hartmann, B., Yu, L., Allison, A., Yang, Y., & Klemmer, S. R. (2008). Design as exploration: creating interface alternatives through parallel authoring and runtime tuning. In *Proceedings of the 21st Annual ACM Symposium on User Int. Soft. and Tech.*, NY, NY.
Kasik, D., Buxton, W., & Ferguson, D. R. (2005). Ten CAD Challenges. *IEEE Computer Graphics and Applications,* 25(2), 81-92.
Kolaric, S., Erhan, H., & Woodbury, R. (2017). CAMBRIA: Interacting with Multiple CAD Alternatives. In CAAD F. Traj., Comm. in *Comp. and Inf. Science. Springer*, pp. 191-211.
Loyola, M. (2018). Big data in building design: a review. J. of Inf. Tech. in Const., 23, 259.
Lunzer, A., & Hornbæk, K. (2008). Subjunctive interfaces: Extending applications to support parallel setup, viewing and control of alternative scenarios. *ACM Trans. on HCI*, 14(4).
Matejka, J., Glueck, M., Bradner, E., Hashemi, A., Grossman, T., & Fitzmaurice, G. (2018). Dream Lens: Exploration and Visualization of Large-Scale Generative Design Datasets. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18*). Association for Computing Machinery, New York, NY, USA.
Mohiuddin, A., & Woodbury, R. (2020). Interactive Parallel Coordinates for Parametric Design Space Exploration. In *Extended Abstracts of the CHI, Human Factors in Computing Systems, Cultural Variation and Heredity*. Springer, pp. 15-32.
Sanchez, R., & Erhan, H. (2014). Design ReExplorer: Interactive Design Narratives for Feedback, Analysis and Exploration. In *Proceedings of the 32nd Education and Research in Computer Aided Architectural Design in Europe Conference*.
Sedlmair, M., Meyer, M., & Munzner, T. (2012). Design Study Methodology: Reflections from the Trenches and the Stacks. *IEEE Trans. on Vis. and Computer Graphics*, 18(12).
Shneiderman, B. (2007). Creativity support tools: accelerating discovery and innovation. *Communications of the ACM*, 50(12), 20-32.
Simon, H. A. (1996). The Sciences of the Artificial (3rd ed.). *The MIT Press*, MA.
Tableau. (2022). https://www.tableau.com/ (Accessed December 2, 2022).
Terry, M., Mynatt, E. D., Nakakoji, K., & Yamamoto, Y. (2004). Variation in element and action: supporting simultaneous development of alternative solutions. In *Proceedings of the SIGCHI Conference on Human Factors in Com. Sys.* (CHI '04). ACM, pp. 711-718.
Thomas, J. J., & Cook, K. A. (Eds.). (2005). Illuminating the Path: The Research and Development Agenda for Visual Analytics. *IEEE Computer Society Press.*
Thornton Tomasetti, C. S. (2019). Design Explorer.
Touloupaki, E., & Theodosiou, T. (2017). Performance sim. integrated in parametric 3D modeling as a method for early-stage design optimization: *A Review. Energies*, 10.
Woodbury, R. (2010). Elements of Parametric Design. Taylor and Francis.
Woodbury, R. F., & Burrow, A. (2006). Whither design space? *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 20, 63-82.
Zarei, M., Erhan, H., Abuzuraiq, A. M., Alsalman, O., & Haas, A. (2021). Design and development of interactive systems for integration of comparative visual analytics in design workflow. *In Prof. of CAADRIA 2021.*

MIXED-USE ECO-COMPUTATIONAL DESIGN IN HOT CLIMATES: A TWO-STEP CROSS-CLIMATE PARAMETRIC WORKFLOW

EILAM SKLAR[1] SHACHAR LORBERBOM[2] and JONATHAN NATANIAN[3]
[1,2,3]*Faculty of Architecture and Town Planning, Technion – Israel Institute of Technology, Haifa, Israel.*
[1]*eilams.sklar@campus.technion.ac.il, ORCID:0009-0009-2759-2556*
[2]*shacharl@campus.technion.ac.il, ORCID:0009-0005-9641-464X*
[3]*jonathann@technion.ac.il, ORCID:0000-0002-8563-3346*

Abstract. Urbanisation trends predict that by 2050, 68% of the global population will reside in cities, requiring substantial urban densification, especially in rapidly developing regions like Israel. This paper integrates mixed-use considerations into the environmental design process through a two-step analytical parametric sequence using Grasshopper and Ladybug in Rhino. In the first step, the solar performance of 11,288 geometrical forms based on six distinct building typologies is analysed and evaluated, considering shading and exposure factors. The second step involves simulating the energy performance of the optimal iteration across different mixed-use scenarios and distributions. The case studies explored in Jerusalem and Eilat, two distinct hot climatic regions in Israel, emphasise the significance of solar-driven considerations in achieving optimal energy efficiency in mixed-use configurations. For example, the study found higher energy efficiency in mixed-use scenarios when office spaces were in the lower half as opposed to the upper half of the building mass in the hot, arid climatic conditions of Eilat. While the computationally inexpensive approach proposed in this study is transferable to other hot climates, the specific built form-related inferences for Eilat and Jerusalem can inform mixed-use design decisions by local designers and policymakers.

Keywords. Building Energy Performance, Mixed-Use, Urban Design, Parametric Design, Solar Design.

1. Introduction

In recent years, the challenge of improving urban energy performance has been at the forefront of the global environmental discourse. Considering the projections for rapid growth in urban dwellers—estimated to increase to 68% of the world population by 2050, up from 50% in 2007 (United Nations, 2019)—the relevance of this challenge is heightened even further. These forecasts should raise concerns in light of the fact that urbanisation has been identified as a major contributor to the progression of climate

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 525-534.

change across multiple spatial scales (Pielke, 2005). In response to rising demands, urban density and morphology undergo modifications, which in turn lead to elevated energy consumption, decreased potential for energy generation within the city (Martins et al., 2016), limited availability of daylight (Capeluto, 2003), and impacts on outdoor thermal comfort (Natanian et al., 2020). In addition to modifications in urban density and morphology, there has been an increased implementation of mixed-use developments within the urban fabric in recent times. These developments integrate various environmental advantages, such as improved walkability, energy efficiency, and infrastructure reduction, among others (Hirt, 2016). While most workflows for optimising environmental performance typically focus on the impact of diverse building types and uses on individual buildings, there is a need for workflows that consider blocks and districts with diverse typologies and uses. This approach allows architects to assess the environmental performance of their designs in the context of rapid urbanisation.

Driven by the predominant influence of solar radiation on energy performance in hot climates, multiple studies have demonstrated a correlation between solar metrics and energy performance in early-stage energy-driven urban design. Solar parametric design has improved over the last few years, evolving from considering only one parameter, as shown in (Capeluto & Plotnikov, 2017), to generative multi-objective key performance indicators (KPIs), as presented by (Natanian et al., 2021). The potential of utilising solar-driven workflows in hot climates is shown in (Natanian et al., 2019) in which they present the potential of implementing parametric workflows for solar-driven urban design in its early stages. As well as at (Natanian & Wortmann, 2021) in which the solar metrics are utilised to evaluate the energy balance potential of larger districts. While solar-driven parametric workflows have proven effective in hot climates, there is a lack of research specifically focusing on these regions. These regions are characterised by significant economic and environmental pressures, including the majority of global urbanisation and demographic growth. As a result, the planning of new cities or districts in these areas must be approached with careful consideration.

Due to the diversity of its hot climates and its future development plans, Israel is a suitable case study for environmental performance in hot climates. According to (Hason et al., 2016), Israel's built environment is projected to double by 2050, in tandem with the country's population, which is expected to increase by more than twofold. Furthermore, with Israel's high urbanisation rate of 92% and limited available land for urban expansion, the anticipation is for urban areas to become more densely populated. Israel, despite its relatively modest area, is characterised by three primary hot climates as classified by the Koppen-Geiger climate system: the hot-summer Mediterranean climate (Csa), the hot semi-arid steppe climate (BSh), and the hot desert climate (BWh). While current strategies address the necessity for urban densification, they overlook the unique climatic conditions across Israel's different climatic zones, which are not adequately reflected in urban development plans throughout the country.

This study addresses the need for block-scale solar parametric workflows in hot climates that take into account mix-use scenarios. This is accomplished by a parametric, two-step workflow that incorporates energy performance simulations and solar indices. Phase one consists of an initial classification utilising solar analysis, from which the iteration with the highest performance is selected. Phase two entails the simulation of the iteration under various distributions and use scenarios. The proposed workflow enables the integration of solar metrics, which evaluate the environmental performance of various forms, with energy evaluation of the most suitable iteration with different use distributions. The following sections describe the analytical workflow and KPIs applied in a case study encompassing two climatic regions in Israel: Jerusalem and Eilat. They discuss the study's results and conclude by highlighting potential future developments.

2. Methodology

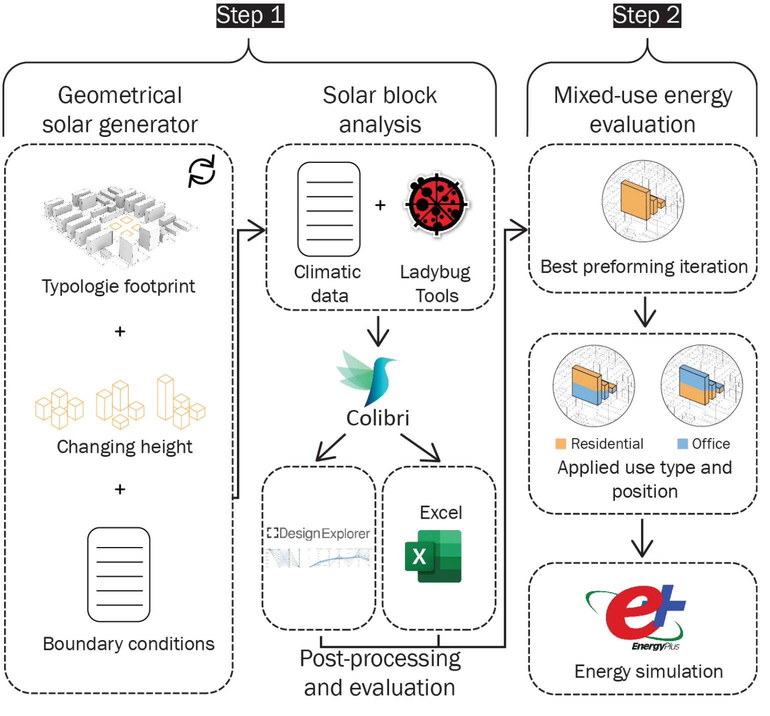


Figure 1. Analytic workflow.

Figure 1 illustrates the study's workflow, comprised of two steps aimed at determining climate-responsive built forms for mixed-use scenarios in hot climates. Both steps utilise Grasshopper-Ladybug within Rhino environment. The first step involves the analysis and evaluation of the solar performance of 11,288 block scenarios based on solar KPIs. These scenarios are automatically generated by modifying two design parameters under two distinct climatic conditions. In the second step, the energy performance of the iterations that achieve the highest performance in each climate is simulated across various distribution and use scenarios. The following sections elaborate on the workflow steps and describe the climatic conditions in the presented case study.

2.1. STEP 1 - GEOMETRICAL GENERATOR

The analysis framework is situated within a representative urban environment in Israel, characterised by rows of slab-block dwellings constructed in substantial numbers during the 1950s and 1960s. The geometry generated in this study originates from six base typologies: (1,2) tower-type with 0- and 45-degree North orientation; (3,4) slab-type with long axis oriented to North-South and East-West; (5) courtyard-type; and (6) scatter-type (Figure 2). To address future densification requirements, the following geometrical variables were employed: the building's typology footprint, varying height with restrictions of 21 stories, a module (voxel) dimension of 3.4 metres, and a floor area ratio (FAR) of 5. The automated parametric study generated a total of 11,288 iterations, evenly split with 5,644 iterations generated for each climate.

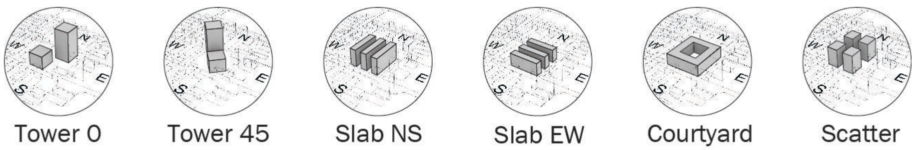


Figure 2. Six base typologies.

Subsequently, each iteration was analysed through Ladybug (Grasshopper plugin) based on five KPIs. This evaluation aimed to address both the external surface exposure and shading performance of the proposed block and its surrounding buildings during the winter and summer seasons (see Table 1). The exposure metrics consist of the following: (1) Context Exposure Index (CEI), (2) New block Exposure Index (NEI) aligned with the guidelines of the Israeli green building code - SI 5281 (Standards Institute of Israel, 2015), and (3) Sky Exposure (SE) metric. Alongside these, the shading metrics include: (4) Outdoor Shading Index (OSI) (Natanian et al., 2020), and (5) East-West Facades Shading Index (FSI) (Natanian & Wortmann, 2021).

The aggregated results were then transmitted via Colibri (a Grasshopper plugin), to Excel for post-processing, and to the online graphic interface Design Explorer for visualization. These results were evaluated by emphasising specific criteria vital to the distinct climatic conditions of each city. In Jerusalem's climate, our focus was on ensuring significant solar exposure on the south facades as indicated by the NEI and CEI, aiming for a minimum exposure of 1.68 kWh/m2 on December 21st in compliance with SI 5281 guidelines. Conversely, for Eilat's climate, we stressed high values for FSI and OSI shading metrics while aiming to minimise sun exposure on the south facade NEI on December 21st, due to the lack of passive heating requirements in Eilat's climate. The design iterations aligned with the emphasised solar KPIs for each climate are visually presented in Figure 3.

Table 1. Evaluation metrics used for the evaluation of each iteration based on (Natanian et al., 2021).

metric	definition	Calculation method	Analysis period	units
CEI	Context Exposure Index	The average solar exposure compliance percentage of the following surfaces according to the respective thresholds taken from SI 5281: Outdoor surfaces - 0.9 kWh/m2 solar irradiation on at least 30% of surfaces South facades - 1.68 kWh/m2 solar irradiation Roof surfaces - 4 hours of solar exposure on at least 50% of each roof surface.	Outdoor surfaces \ south facades - 21st of December 08:00 – 16:00 Roof Surfaces 21st of December 09:00 – 15:00	[%]
NEI	New Block Exposure Index			
SE	Sky Exposure	The average percentage of the sky visible from each of the test point across all the vertical surfaces - averaged between existing and new masses.	N/A	[%]
OSI	Outdoor Shading Index	The average of the annual irradiation ratio between exposed and obstructed configurations of each point across the outdoor surfaces (in and around the site), subtracted from 1	Annual	[/]
FSI	East-West Facades Shading Index	East and West facades' summer irradiation values (in MW/m2), subtracted from 1-averaged between existing and new masses.	1st of June – 31st of October	[1MW/m2]

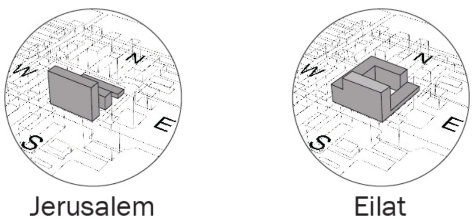


Figure 3. Selected iterations for Jerusalem (left) and Eilat (right).

2.2. STEP 2 - MIXED-USE ENERGY EVALUATION

After identifying suitable iterations for each climate zone, we proceeded with three energy simulations encompassing various use types and distributions within each climate zone. Additionally, we simulated Tower typology iteration, frequently designed in similar development scenarios, using the same inputs and constraints as a reference design. In total, eight energy simulations were conducted, categorised as follows: (1-4) Establishing a baseline evaluation by conducting energy simulations on the selected iterations and the reference design across both climate conditions, in compliance with the Israeli energy rating code IS 5282 (Standards Institute of Israel, 2020) for residential construction and schedule. (5-8) Conducting a second set of simulations to evaluate the energy impact of mixed-use programmes and their placement within the building for each chosen iteration. This time, with a 50% office programme following IS 5282 standards for office construction and schedule, the office spaces were stacked either above or below the residential areas. All simulations were performed using EnergyPlus-Honeybee. The Energy Use Intensity (EUI) for each simulation, including heating, cooling, equipment, and lighting loads, was assessed annually and on a hot summer day (August 1). These values were measured in kWh/m2 and Wh/m2, respectively.

2.3. CLIMATIC CONTEXT

Israel, located between 30° and 33° north latitude, has a diverse climate despite its small size. This diversity derives primarily from altitude, latitude, and proximity to the Mediterranean Sea. The study focused on two of Israel's three main climates, as classified by the Koppen-Geiger climate system: Jerusalem's climate (Csa) represents a subtropical, hot Mediterranean, dry-summer climate with moderate seasonality. In contrast, Eilat's climate (BWh) is arid, low-latitude, and dry, receiving minimal precipitation. Figure 4 showcases the distinctions in air temperature and humidity between Jerusalem and Eilat. Jerusalem experiences notably lower temperatures, dropping to lows of 5°C in winter, while Eilat exhibits significantly higher temperatures, peaking at 41°C during summer. Both climates exhibit significant diurnal temperature fluctuations. Jerusalem encounters on average 62 days annually with temperatures over 25.5°C and 37 hours annually with both temperatures above 25.5°C and relative humidity exceeding 60%. Conversely, Eilat records 203 days and 11 hours, respectively, under similar conditions. Due to these climatic differences, passive heating is crucial in Jerusalem to reduce energy consumption, whereas solar shading is prioritised in Eilat. In the following sections, these climatic differences will serve as a reference point to define the solar shading and exposure indices for comparing various block iterations and use distribution configurations.

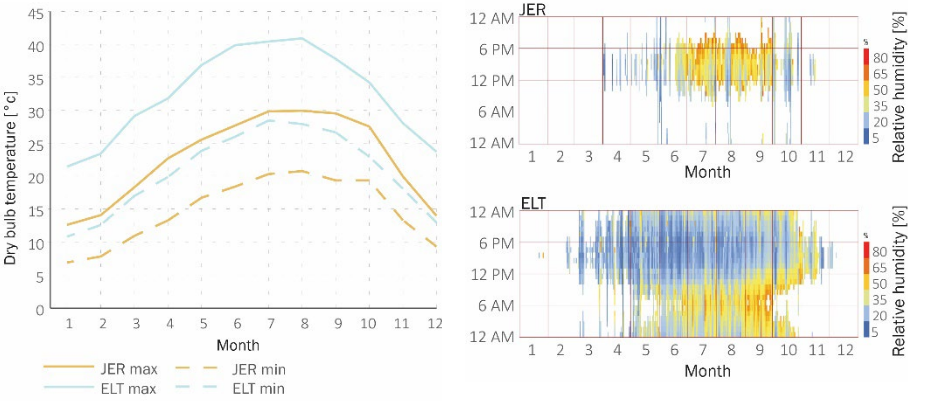


Figure 4. Different climatic conditions in Jerusalem and Eilat. Dry bulb temperature (left) and relative humidity on days with a dry bulb temperature exceeding 25.5°C (right).

3. Results and Discussion

This section will discuss the results obtained from both the initial solar screening evaluation (step 1) and the detailed energy performance evaluation (step 2).

3.1. STEP 1 - SOLAR EVALUATION

Figure 5 illustrates the analysed solar data of the chosen iterations in Eilat (Courtyard) and Jerusalem (Slabs EW), derived from their respective mass configurations. Both iterations exhibit high shading values (FSI and OSI) due to self-shading. The East-West Facades Shading Index (FSI) holds significance in both hot climates to mitigate excessive radiation's adverse impact on thermal comfort during cooling seasons. Additionally, the outdoor shading index (OSI) is crucial in ensuring year-round outdoor thermal comfort in the immediate vicinity of the building block, particularly in hot climates where solar radiation predominantly influences outdoor thermal conditions, as also shown in (Natanian et al., 2020). The chosen iteration for Jerusalem displays elevated radiation values on the south facade of both the new block (NEI) and the surrounding context (CEI). This results from its high facade ratio facing south and the gradual decline in height, allowing solar radiation to reach the south facades of the surrounding context. South facade exposure proves vital for passive heating in winter, contributing to enhanced indoor thermal comfort. Conversely, passive heating is undesirable in Eilat's hot and arid climate, where excessive heating may occur; thus, lower values of south facade exposure (NEI) are preferable.

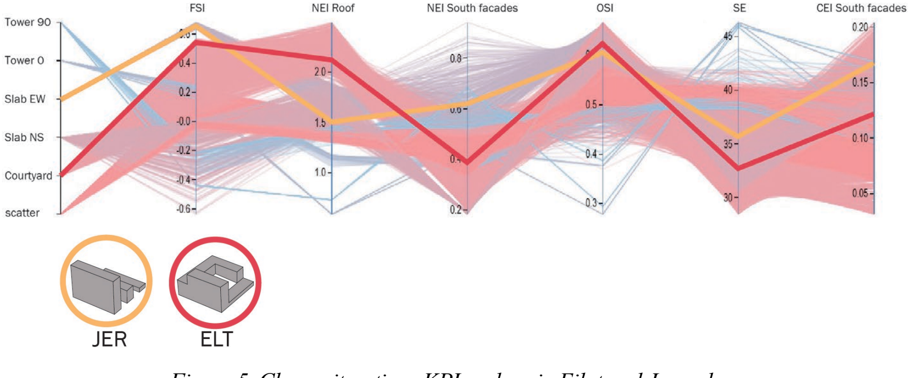


Figure 5. Chosen iterations KPIs values in Eilat and Jerusalem.

3.2. STEP 2 - ENERGY PERFORMANCE EVALUATION

Examination of the annual Energy Use Intensity (EUI) presented in Figure 6 shows that Eilat's energy consumption is significantly higher than that of Jerusalem, both in the tested iteration (Courtyard in Eilat and Slab in Jerusalem) and in the reference Tower design. This disparity arises due to the considerable demand for cooling in Eilat's climate, resulting in a significant proportion of the EUI being allocated to cooling loads. Notably, the Courtyard iteration in Eilat demonstrated lower energy consumption in comparison to the reference Tower design iteration, aligning with findings from previous research (Natanian et al., 2019). This reduced consumption can be attributed to the Courtyard typology's compact mass shape and self-shading features. The Slabs EW iterations in Jerusalem showcase reduced energy consumption in comparison to the reference Tower design iteration. This could be attributed to the high East-West Facades Shading Index (FSI) driven by its narrow east and west facades and the self-shading effect of the higher south mass on the two smaller masses in the Slabs EW iteration. This self-shading may also account for the difference in heating demands between the reference Tower design iteration with residential use (1.5 kWh/m2) and the Slabs EW with residential use (5.5 kWh/m2) in Jerusalem.

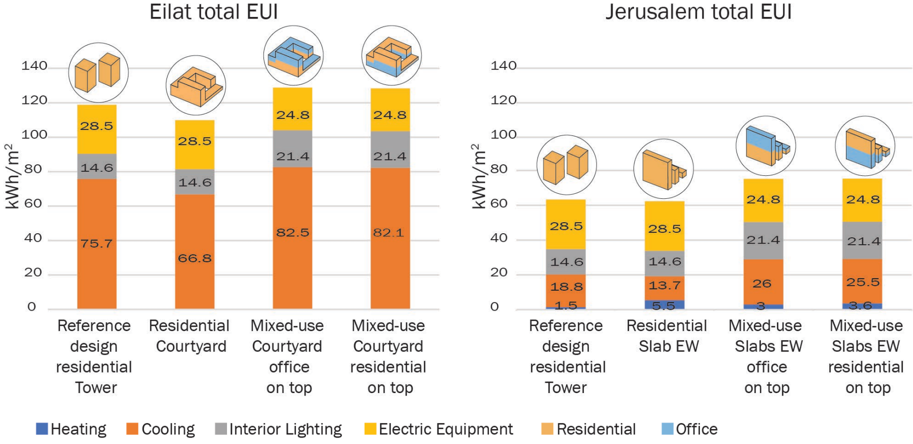


Figure 6. Annual Energy Use Intensity (EUI) of the different chosen typologies and programmes in Eilat and Jerusalem.

Figure 7 presents the daily energy consumption analysis of the selected iterations, encompassing all simulated use types and distributions on August 1, a hot summer day. It visually represents how the occupancy hours of each programme and their location within the block mass influence the daily energy consumption. In both climates, the reference Tower design iteration allocated for residential use demonstrates higher energy consumption during daylight hours compared to the selected iterations with the same use. This disparity is attributed to the Tower iteration's low East-West Facades Shading Index (FSI) values, indicating excessive heat absorption during the summer months. Further comparison of office placement in the upper and lower sections across both climates reveals that iterations with offices in the upper section consume more energy during the daylight hours due to increased solar exposure. The analysis of mixed-use scenarios in both iterations and climates suggests that placing offices in the lower half proves more energy-efficient for cooling. Conversely, for heating, positioning them in the upper half yields better results, owing to extended occupancy hours during daylight, influenced by solar exposure. As a result, in Eilat's climatic conditions, where cooling demands outweigh heating demands, offices are more efficiently situated in the lower half. However, findings in Jerusalem are inconclusive due to a more even distribution of heating and cooling demands throughout the year.

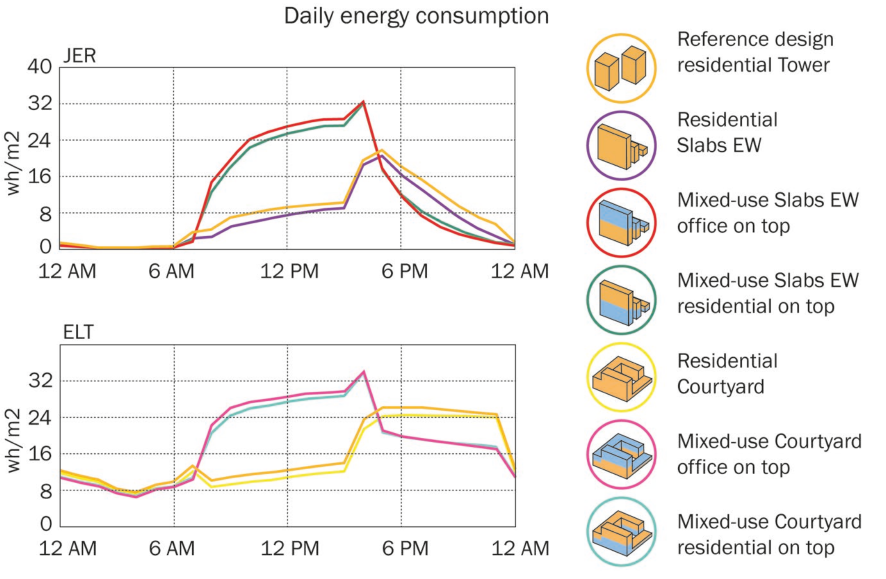


Figure 7. Daily energy consumption analysis of the two selected iterations on August 1 in all simulated use types and positions.

4. Conclusions

Motivated by the need to integrate mix-use scenarios into block-scale solar parametric workflows in hot climates, this study introduced a two-step parametric workflow that integrates energy performance simulations and solar indices. Initially, we analysed the solar KPIs of 11,288 geometrical iterations under two distinct hot climate scenarios. Subsequently, in the second step, considering climate-specific parameters, we selected and evaluated the best-performing iteration in each climate across a range of use profiles and distributions. The study underscores the importance of incorporating mixed-use scenarios and their distributions in urban block environmental assessments. The recommendations regarding built form in this study are applicable to architects and urban designers in Eilat and Jerusalem.

Notably, only the options selected in phase one underwent analysis in the subsequent energy assessment conducted in step two. Had this analysis been conducted earlier, it might have led to different choices in the building typologies. Additionally, this study did not account for window shadings, energy generation potential, microclimatic impacts, vegetation, wall materiality, and thermal comfort parameters, which should be considered in future studies.

The method described in this study is easily adaptable for use in other hot climates. This adaptability comes from its use of a parametric environment and open-source tools, which are increasingly gaining popularity among designers. By utilising the presented workflow, designers can access indicators that enhance the effectiveness of incorporating solar-driven considerations during the initial phases of design.

Acknowledgements

This study is part of the project Energy performance in heterogeneous urban environments: challenges and opportunities in mixed-use and typologically diverse districts in Israel, supported by the Israeli Ministry of Energy [grant number 221-11-037].

References

Capeluto, I. G. (2003). The influence of the urban environment on the availability of daylighting in office buildings in Israel. *Building and Environment*, 38(5), 745–752. https://doi.org/10.1016/S0360-1323(02)00238-X

Capeluto, I. G., & Plotnikov, B. (2017). A method for the generation of climate-based, context-dependent parametric solar envelopes. *Architectural Science Review*, 60(5), 395–407. https://doi.org/10.1080/00038628.2017.1331334

Hason, S., Kotock, O., Drukman, D., & Roter, D. (2016). Israel 2048: Strategic Thinking towards regional planning and development. *Israel Planning Authority and the Hebrew University of Israel*.

Hirt, S. A. (2016). Rooting out mixed use: Revisiting the original rationales. *Land Use Policy*, 50, 134–147. https://doi.org/10.1016/j.landusepol.2015.09.009

Martins, T. A. de L., Adolphe, L., Bastos, L. E. G., & Martins, M. A. de L. (2016). Sensitivity analysis of urban morphology factors regarding solar energy potential of buildings in a Brazilian tropical context. *Solar Energy*, 137, 11–24. https://doi.org/10.1016/j.solener.2016.07.053

Natanian, J., Aleksandrowicz, O., & Auer, T. (2019). A parametric approach to optimizing urban form, energy balance and environmental quality: The case of Mediterranean districts. *Applied Energy*, 254, 113637. https://doi.org/10.1016/j.apenergy.2019.113637

Natanian, J., Kastner, P., Dogan, T., & Auer, T. (2020). From energy performative to livable Mediterranean cities: An annual outdoor thermal comfort and energy balance cross-climatic typological study. *Energy and Buildings*, 224, 110283. https://doi.org/10.1016/j.enbuild.2020.110283

Natanian, J., Luca, F. D., Wortmann, T., & Capeluto, G. (2021). The Solar Block Generator: An additive parametric method for solar driven urban block design. *Journal of Physics: Conference Series*, 2042(1), 012049. https://doi.org/10.1088/1742-6596/2042/1/012049

Natanian, J., & Wortmann, T. (2021). Simplified evaluation metrics for generative energy-driven urban design: A morphological study of residential blocks in Tel Aviv. *Energy and Buildings*, 240, 110916. https://doi.org/10.1016/j.enbuild.2021.110916

Pielke, R. A. (2005). Land Use and Climate Change. *Science*, 310(5754), 1625–1626. https://doi.org/10.1126/science.1120529

Standards Institute of Israel. (2015). SI 5281. *Sustainable Buildings*.

Standards Institute of Israel. (2020). SI 5282. *Sustainable Buildings*.

United Nations. (2019). World Urbanization Prospects: The 2018 Revision. *UN*. https://doi.org/10.18356/b9e995fe-en

MULTISPECIES BUILDING ENVELOPES

Adopting plant habitat suitability modelling for ecological design decision-making

SURAYYN UTHAYA SELVAN[1], LAURA WINDORFER[2], MARIASOLE CALBI[3], SOULTANA T. SAROGLOU[4], ENRICA ROCCOTIELLO[5], SHANY BARATH[6] AND YASHA J. GROBMAN[7]
[1,4,6,7]TECHNION - Israel Institute of Technology, [2]Technical University of Munich, [3,5]University of Genoa.
[1]surayyn@campus.technion.ac.il, 0000-0001-9401-1170
[2]laura.windorfer@tum.de, 0009-0001-3355-1015
[3]mariasole.calbi@edu.unige.it, 0000-0001-6018-4022
[4]saroglou@campus.technion.ac.il, 0000-0002-0322-4670
[5]enrica.roccotiello@unige.it, 0000-0003-3701-9154
[6]barathshany@technion.ac.il, 0000-0002-6490-1380
[7]yasha@technion.ac.il, 0000-0003-4683-4601

Abstract. Urbanisation catalyses habitat loss, impacting humans and biodiversity. To mitigate this, the ECOLOPES research project proposes multispecies building envelopes that enhance ecosystem services provision through cohabitation. Initial envelope designs are optimized and evaluated with a hybrid multi-criteria decision-making model informed by key performance indicators for human and non-human objectives. This paper proposes an architectural approach to derive indicator proxies for the plant stakeholders by adopting aspects of habitat suitability modelling by correlating environmental conditions with species functional traits. Using the hybrid decision-making model, we utilise these proxies to optimise multispecies objectives for a residential building envelope and evaluate the resulting alternative. This alternative is compared with one optimised using indicators inferred from general ecological correlations. Results show the effectiveness of applying the proposed habitat suitability approach in accounting for variations in plant trait values and improving multi-objective trade-offs for multispecies envelope design decision-making.

Keywords. multispecies building envelopes, hybrid multi-criteria decision-making, plant habitat suitability, key performance indicators

1. Introduction

Urbanisation catalyses habitat loss and fragmentation, negatively impacting biodiversity and human well-being (IPBES, 2019). Multispecies building envelopes alleviate this by enhancing ecosystem services provision and cohabitation opportunities between four stakeholders: humans, animals, plants, and microbiota (Weisser et al., 2022). However, computational decision-making challenges arise due to insufficient ecological knowledge for building envelope design (Grobman et al., 2023). This often leads to the use of human-centric and simplified ecological indicators to evaluate design decisions (Selvan et al., 2023). In the ECOLOPES research project, we aim to develop methodologies and computational tools for multispecies building envelope design (Weisser et al., 2022). The proposed methodology includes a hybrid multi-criteria decision-making (MCDM) model that generates and ranks optimised envelope alternatives (Selvan et al., 2023). The model is informed by a framework, termed nested hierarchies, that deconstructs stakeholder objectives to derive directional constraints and key performance indicators (KPIs) (Saroglou et al., 2024). This framework also unmasks potential multispecies trade-offs by exploring KPI relationships between the stakeholders. To effectively evaluate these trade-offs, KPIs must be explicitly computed for all the multispecies stakeholders.

In ECOLOPES, this is facilitated by several novel technologies including expert ecological modelling integrated into a computer-aided design (CAD) environment, as outlined by Vogler et al. (2023). Building upon technological concepts initiated and developed by ECOLOPES, our paper proposes a parallel approach that adopts principles of habitat suitability modelling (HSM) to derive KPI proxies specifically for the plant stakeholder. HSM, a statistical method common in ecology, computes environmental data to predict the presence or absence of a given species in a study area (Hirzel et al., 2006). We test the applicability of the proposed approach with the hybrid MCDM model by optimizing a generic residential building envelope for human and plant stakeholders. The resulting alternative was compared with one informed by KPIs derived from general ecological correlations. Finally, we discuss potential extensions of the HSM approach and outlooks for multispecies envelope design decision-making.

2. Adopting Habitat Suitability Modelling for Design Decision-making

Climatic deterioration and species extinction are catalysed by urbanization, requiring comprehensive understanding of species distribution (Mohammady et al., 2021). This can be achieved by characterising the biotic and abiotic conditions necessary for a species to persist, known as ecological niches. (Polechová & Storch, 2019). In ecology, HSMs are spatially explicit models that assess these niches by computing biotic and abiotic factors, e.g., climate, topography, and geology, with species occurrence data to predict the absence or presence of a species in a given geographical area (Hirzel et al., 2006). HSMs produce spatial-temporal maps used to visualize habitat suitability and species occurrence probability and favourability (Sillero et al., 2021).

For example, Shen et al. (2021) predicted climate change impacts on the habitat suitability of a medicinal plant, considering 19 bioclimatic variables including temperature, precipitation, and topography. Chin et al., (2022) prioritized avian diversity maintenance using avian functional traits and factors such as connectivity, patch quality, and land cover. In an architectural study by Zimbarg (2023), building envelope microclimates were utilized to map plant species based on shading, incident radiation, and humidity - implicitly using habitat suitability. As reflected, the choice of variables used to model habitat suitability depend on the research objective and target species, varying influences from the regional to local scales (Bradley et al., 2012).

2.1. PROPOSED PLANT HABITAT SUITABILITY APPROACH

Drawing insights from Zimbarg (2023), we propose a mapping approach based on HSM principles. This approach utilizes environmental simulations to predict the presence of selected plant species on an envelope geometry. We also adopt the use of species functional traits from Chin et al., (2022), which are "any traits that impact fitness indirectly via their effects on the growth, reproduction, and survival of a species" (Violle et al., 2007). Therefore, functional traits have strong associations with ecological niches and in turn, habitat suitability. The proposed approach, constructed in Grasshopper [ver. 1.0.0007], leverages open-access plugins that facilitate interdisciplinary design support, real-time simulation, and visualization of algorithmic design decisions. This approach simulates abiotic factors on a mesh geometry and correlates associated plant functional traits to generate a 3D habitat suitability map (*Fig. 1*). The map provides proxies for widely used KPIs to measure plant species diversity, e.g., species richness and abundance. This proposed approach enables these KPI proxies to be integrated into schematic design decision-making processes.

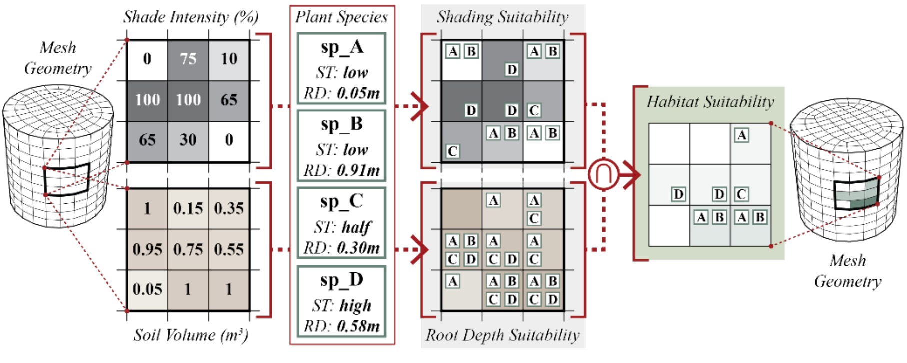


Figure 1. Approximating habitat suitability by associating abiotic factors (shade intensity and soil volume) to plant species functional traits, i.e., shade tolerance (ST) and rooting depth (RD).

In *Fig. 1*, the logic of the proposed plant HSM is illustrated using a 3x3 segment from an example envelope mesh geometry and four mock plant species *A*, *B*, *C*, and *D*. The habitat suitability of these plant species is predicted by correlating the abiotic factors, shade intensity and soil volume, to the functional traits, shade tolerance (ST) and rooting depth (RD). First, the abiotic factors are simulated using existing Grasshopper plugins and components. For example, shade can be simulated using Ladybug tools while soil volume can be distributed using native components. Then, the functional trait values for each species are obtained from ecological databases, e.g., TRY Plant Trait Database (Kattge et al., 2011). Next, these trait values are compared with the abiotic simulation results to assess suitability. For example, a cell with 75% shade might only be suitable for *sp_D* with a high shade tolerance value, while a soil volume of $0.35m^3$ is suitable for the rooting depth requirements of *sp_A* and *sp_C*.

Ultimately, the intersection of these comparisons generates a 3D plant habitat suitability map, defining the ecological niche of each species relative to the selected abiotic factors. The map also provides proxies for commonly used plant species diversity KPIs. Namely, overall and local plant species richness (*number of species*), which are approximated by counting the number of plant species present across all cells or within individual cells, respectively. In *Fig. 1*, the overall species richness is four and the local richness in the last cell is two. Likewise, plant species abundance (*number of individuals*) can also be approximated by counting the total number of cells suitable for each species. In *Fig. 1*, the abundance of *sp_A* is three while for *sp_B*, it is two.

3. Research Methodology

The proposed plant HSM approach was evaluated through a comparative experiment using the hybrid MCDM model to achieve selected multispecies objectives for a generic residential building envelope situated in Tel Aviv, Israel. The 17m-high building features five floors with a floor-to-ceiling height of 3m (*Fig. 2*). The building geometry was converted into mesh cells to facilitate the environmental simulations and application of the proposed plant HSM approach. The building was described by window and soil cell distribution variables, creating inherent conflicts suitable to generate MOO trade-offs. For example, an increase in window cells results in a decrease in available cells for soil distribution. The window cells, randomly distributed, are constrained by local architectural standards for window-to-wall ratio where the North, East and West, and South facades have maximum ratios of 12%, 8%, and 20%, respectively (*Fig 2a*). Similarly, soil cells were randomly distributed using soil-to-wall ratios across different floor groups: ground floor, 1st and 2nd floor, 3rd and 4th floor, and roof level. To generate soil volume, discrete values ranging from 0.15 to 1m were randomly assigned and extruded from the 1x1 soil cells in the z-axis (*Fig. 2b*).

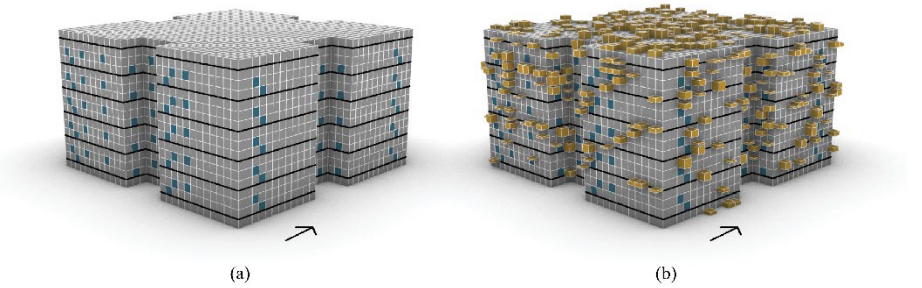


Figure 2. a) Default state of the generic residential building envelope and b) an example distribution of soil volumes from ground floor to roof level with 5%, 15%, 20%, and 25% soil-to-wall ratio.

For the multispecies objectives, the human and plant stakeholders were selected. The overarching human objectives were to improve daylighting and to reduce envelope structural loads, aligned with human-centric themes of human comfort and building performance. Simultaneously, the main plant objective aimed to enhance species diversity on the envelope, focusing on species native to the Mediterranean climate (*Table 1*). Their respective mean functional trait values, associated to the abiotic factors for shade intensity and soil volume, were obtained from the TRY Plant Trait Database (Kattge et al., 2011). Shade tolerance was valued from 0 for low tolerance to 5 for high tolerance while rooting depth was measured in meters.

Table 1. List of plant species and respective functional trait mean values from the TRY database.

No.	*Name*	*Scientific Name*	*Functional Traits (Mean Values)*	
			Shade Tolerance (0 to 5)	*Rooting Depth (m)*
1	Common Ivy	*Hedera helix* L.	4	0.1624
2	Annual Bluegrass	*Poa annua* L.	1	0.1881
3	Salad Burnet	*Sanguisorba minor* Scop.	1	0.6127
4	Common Chickweed	*Stellaria media* (L.) Vill.	2	0.3
5	Hare's-foot Clover	*Trifolium arvense* L.	0	0.05

The experiment was performed using the hybrid MCDM model developed under the ECOLOPES framework (Selvan et al., 2023). The model integrates multi-objective optimisation (MOO) using Wallacei [ver.2.65] and multi-attribute decision-making (MADM) using the Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) scripted in Grasshopper (Hwang & Yoon, 1981; Makki et al., 2019). First, the model optimises an initial envelope alternative using fitness objectives, defined by directional constraints and associated KPIs, outlined in a nested hierarchy for stakeholder objectives. Next, the model ranks the resulting Pareto front solutions according to KPI weights distributed based on stakeholder priorities. TOPSIS calculates scores from 0 to 1, ranking solutions from the best to worst performing based on the priorities. This sequential process of hybrid MCDM allows the generation of optimized alternatives and the identification of the most appropriate solution.

3.1. COMPARATIVE EXPERIMENT DESIGN

Two nested hierarchies were defined to derive different directional constraints and KPIs aimed at achieving the principal plant stakeholder objective "*to increase species diversity*". The first nested hierarchy (NH1), detailed in Section 3.1.1, was based on inferred correlations of plant growth and survival. The second nested hierarchy (NH2), detailed in Section 3.1.2, drew on the principles of the proposed plant HSM approach.

The model was initiated with the default configuration of the building envelope (*Fig.2a*). For MOO, the default algorithm parameters of Wallacei were used for a generation size and count of 10 each, producing 100 alternatives. Window and soil cell distribution variables, as detailed in Section 3, were used as the gene input. After the deconstruction of the primary objectives, the fitness objective for the human stakeholder was to maximize the total window-to-wall ratio and minimize total soil volume (*Fig.3&4*). The plant stakeholder fitness objectives differed in the two optimisations aligned to the nested hierarchies. For MADM, the Pareto front solutions were ranked with equal stakeholder objective priorities that distributed the KPI weights accordingly. Ultimately, the best-performing alternatives of each nested hierarchy were identified and compared based on the TOPSIS score, gene and fitness objective values, total number of suitable cells, and resulting species diversity KPI proxies.

3.1.1. Nested Hierarchy 1 (NH1): General Ecological Correlations

NH1 was based on general ecological correlations about plant growth and survival. As seen in *Fig. 3*, the first-level plant objective "*to increase plant species richness*" diverged to two lower-level objectives "*to maximise total soil volume*" and "*to maximise shade heterogeneity*". This was defined based on general ecological assumptions that higher soil depths/volumes and heterogenous shade conditions offer better opportunities for improved plant species diversity. These objectives were deconstructed into one KPI and directional constraint, each. The KPI, total soil volume, was also shared with the human objective "*to reduce envelope structural loads*" but framed in the opposite direction. This KPI was computed using the Grasshopper script constructed to distribute soil depth values. The second KPI, direct sun hours, was associated with shade heterogeneity by computing the standard deviation of the values. This KPI was computed using Ladybug [ver.1.6.0] and the Statistics component from Dodo [ver.03]. However, to streamline the species diversity comparison, the proposed plant HSM approach was performed on the best-performing optimized alternative.

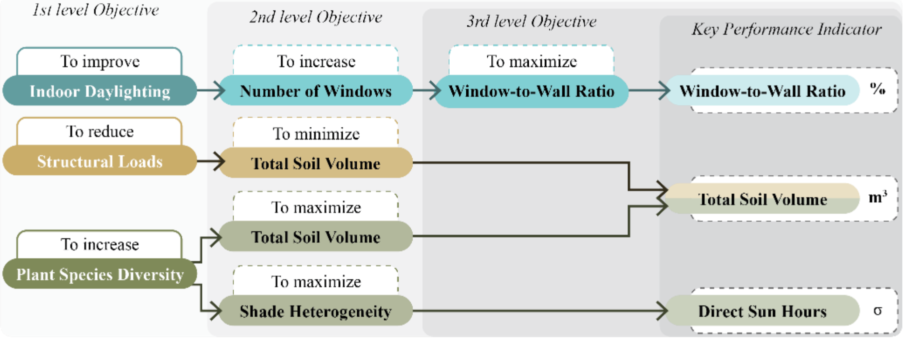


Figure 3. NH1 illustrating the human objectives (blue and brown) and the plant objective (green). The plant objective is associated with two KPIs derived from ecological correlations.

3.1.2. Nested Hierarchy 2 (NH2): Habitat Suitability Approach

NH2 was based on the proposed plant HSM-inspired approach described in Section 2.1. As shown in *Fig. 4*, the first-level plant objective remained unchanged but was deconstructed into a second-level objective "*to increase habitat suitability*". This was formulated on the prediction of habitat suitability using selected abiotic factors and correlated plant functional traits. Therefore, to achieve the objective, the number of cells with rooting depth and shade suitability must be maximized. This drives the fitness objective to accommodate the list of species across all the suitable cells. Finally, these objectives converged into the species diversity proxy, describing the richness and abundance. These proxies for the number of species and individuals on the envelope, respectively, were represented under one fitness objective to be maximised.

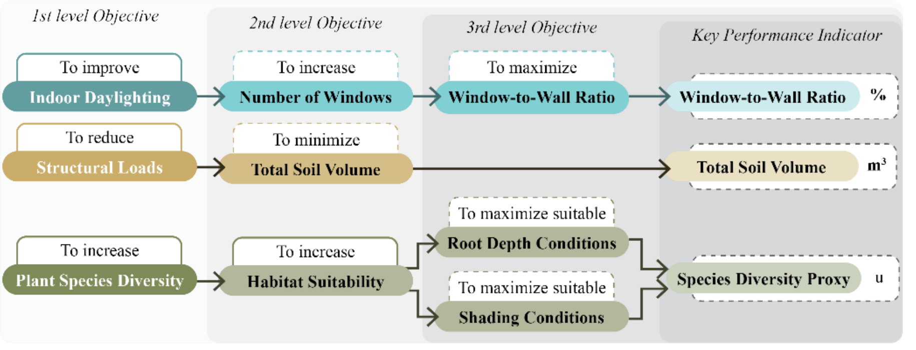


Figure 4. NH2 illustrating the human objectives (blue and brown) and a plant objective (green). The plant objective is associated with the KPI proxies computed by the proposed plant HSM approach.

4. Results and Discussion

NH1, with general ecological correlations, yielded 44 Pareto front solutions while NH2, based on the proposed plant HSM approach, had 30 solutions. As seen in *Fig. 5*, the best-performing alternative from NH1 was the third individual in the second generation: Alternative {1-2}, with a TOPSIS score of 0.5225. From NH2, it was the second individual in the second generation: Alternative {1-1} with a score of 0.6526. Referring to *Table 2*, Alternative {1-2} showcased optimal window distribution, allowing a gradual decrease in soil-to-wall ratios from the ground to roof level. A total soil volume of 818.48 m^3 was distributed across 1420 cells with habitat suitability. Across all cells, the optimized envelope supported an overall species richness of five. For each cell, the envelope had a mode and mean local richness of one and 1.14, respectively. Conversely, Alternative {1-1} had an optimal window distribution with substantially higher soil-to-wall ratios, except on the 3rd and 4th floor, which only achieved 30%. Hence, a higher total soil volume of 1260.3 m^3 was distributed across 2152 suitable cells. Similarly, across all cells, the overall species richness remained at five, with a mode local richness of one with a slightly higher mean of 1.42, per cell.

Based on the results, NH2 produced 32% fewer Pareto front solutions compared to NH1, suggesting a reduced design search space. This facilitates decision-making for selecting potential alternatives. While both alternatives achieved window distributions close to the architectural standards, Alternative {1-2} had the lowest soil volume to achieve reduced structural loads. However, Alternative {1-1} had 34% more suitable cells and 35% higher soil volume aligning with the plant objective of increasing species diversity. Moreover, while both alternatives had the same overall species richness, the mean local richness for Alternative {1-1} was 20% higher than in Alternative {1-2} due to higher species abundance (*Table 2*). Except from Common Chickweed, the abundance of Common Ivy, Annual Bluegrass, Salad Burnet, and Hare's-foot Clover in Alternative {1-1} exceeded that of Alternative {1-2} by 21%, 72%, 77%, and 68%, respectively. This suggests that the hybrid MCDM model using the proposed plant HSM achieved better trade-offs for the plant objective without compromising the human objectives by accounting for the functional trait variations in each species.

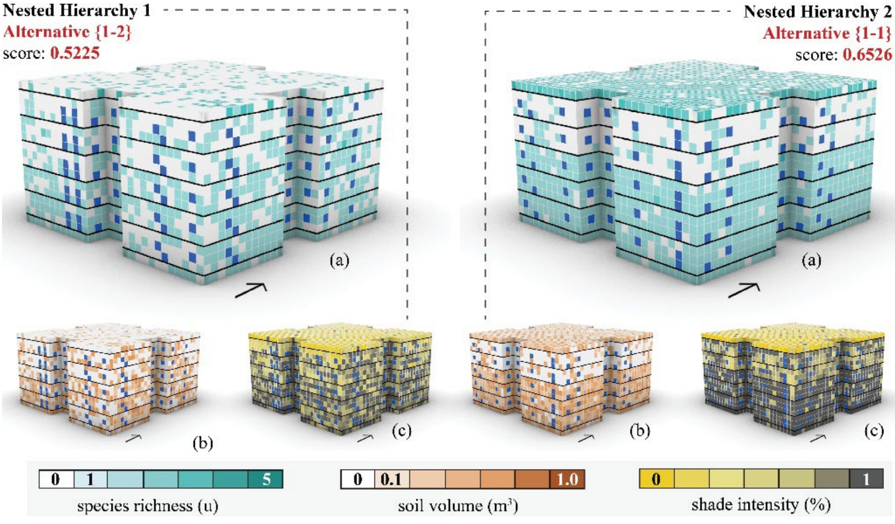


Figure 5. Best-performing alternatives from NH1 and NH2 with the a) plant suitability mapping, b) soil volume distribution, and c) shading intensity visualised.

Table 2. TOPSIS score, optimized gene and fitness objective values, overall and local plant species richness, and species abundance of the resulting best-performing alternatives.

Criteria	*Additional Notes*	*Unit*	*Experiment 1*	*Experiment 2*
			Alternative {1-2}	*Alternative {1-1}*
Performance score	-	-	**0.5225**	**0.6526**
Window-to-wall ratio	*N-S-E-W facades*	%	9, 20, 7, 6	10, 18, 7, 6
Soil-to-wall ratio	*Ground to roof level*	%	85, 70, 45, 25	95, 95, 30, 90
Total Soil Volume	-	m^3	818.48	1260.3
Number of Suitable Cells	*On envelope*	u	1420	2152
Overall Species Richness	*Across envelope*	u	5	5
Mode Local Richness	*Per envelope cell*	u	1	1
Mean Local Richness	*Per envelope cell*	u / cell	1.1358	1.1476
Total Species Abundance	*Following sequence in Table 1*	u	1008, 190, 70, 196, 216	1278, 676, 310, 198, 676

5. Conclusions

This paper proposed a mapping approach based on the principles of plant HSM to optimise and evaluate design alternatives using ecologically driven KPIs. The approach employed environmental simulation plugins to analyse abiotic conditions which were correlated with plant functional trait values extracted from an ecological database. The resulting habitat suitability map provided plant KPI proxies for species diversity to achieve a plant objective within a hybrid MCDM case study. A comparison with an alternative informed by general ecological correlations revealed that the plant HSM-informed alternative achieved higher species suitability across the envelope, without compromising the human stakeholder objectives. This highlights the advantage of the proposed plant HSM approach in facilitating improved multispecies stakeholder trade-offs through informed ecological knowledge using the hybrid MCDM model.

The adaptable "plug-and-play" format of the approach allows for the integration of additional plant or animal functional traits, given computable abiotic factors. For example, plant growth forms could be correlated with the surface normal direction of mesh faces, or animal habitat suitability correlations could be explored based on resource availability or potential living spaces on the envelope geometry. Future experiments could leverage the approach to optimize and assess objectives tailored to specific species for conservation support. Additionally, spatially informed objectives could be explored such as human-nature proximity or habitat clustering. In summary, the proposed HSM approach functions as a methodology to obtain schematic ecological results in the early design decision-making phases of multispecies building envelopes. Notably, a limitation of the proposed plant HSM is the static nature of the analysis, lacking temporal dynamics such as competition which is essential for informed design decision-making. Ongoing developments in the ECOLOPES research project bridges this gap by integrating advanced ecological modelling into CAD processes, supporting collaboration between architectural and ecological domains.

Acknowledgements

This work is funded by H2020-EU.1.2 - FET Open [Grant Agreement ID 964414]. The authors express their gratitude to the ECOLOPES consortium members for the interdisciplinary support and collaboration.

References

Bradley, B. A., Olsson, A. D., Wang, O., Dickson, B. G., Pelech, L., Sesnie, S. E., & Zachmann, L. J. (2012). Species detection vs. habitat suitability: Are we biasing habitat suitability models with remotely sensed data? Ecological Modelling, 244, 57–64. https://doi.org/10.1016/j.ecolmodel.2012.06.019

Chin, A. T. M., Ruppert, J. L. W., Shrestha, N., & Fortin, M.-J. (2022). Urban Avian Conservation Planning Using Species Functional Traits and Habitat Suitability Mapping. Land, 11(10), Article 10. https://doi.org/10.3390/land11101831

Grobman, Y. J., Weisser, W., Shwartz, A., Ludwig, F., Kozlovsky, R., Ferdman, A., Perini, K., Hauck, T. E., Selvan, S. U., Saroglou, S. (Tanya), Barath, S., Schloter, M., & Windorfer, L. (2023). Architectural Multispecies Building Design: Concepts, Challenges, and Design Process. Sustainability, 15(21), Article 21. https://doi.org/10.3390/su152115480

Hirzel, A. H., Le Lay, G., Helfer, V., Randin, C., & Guisan, A. (2006). Evaluating the ability of habitat suitability models to predict species presences. Ecological Modelling, 199(2), 142–152. https://doi.org/10.1016/j.ecolmodel.2006.05.017

Hwang, C.-L., & Yoon, K. (1981). Introduction. In C.-L. Hwang & K. Yoon (Eds.), Multiple Attribute Decision Making: Methods and Applications A State-of-the-Art Survey (pp. 1–15). Springer. https://doi.org/10.1007/978-3-642-48318-9_1

IPBES. (2019). Global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. Zenodo. https://doi.org/10.5281/zenodo.5657041

Kattge, J., Díaz, S., Lavorel, S., Prentice, I. C., Leadley, P., Bönisch, G., Garnier, E., Westoby, M., Reich, P. B., Wright, I. J., Cornelissen, J. H. C., Violle, C., Harrison, S. P., Van BODEGOM, P. M., Reichstein, M., Enquist, B. J., Soudzilovskaia, N. A., Ackerly, D. D., Anand, M., … Wirth, C. (2011). TRY - a global database of plant traits: TRY - A GLOBAL DATABASE OF PLANT TRAITS. Global Change Biology, 17(9), 2905–2935. https://doi.org/10.1111/j.1365-2486.2011.02451.x

Makki, M., Showkatbaksh, M., & Song, Y. (2019). Wallacei Primer 2.0. https://www.wallacei.com/

Mohammady, M., Pourghasemi, H. R., Yousefi, S., Dastres, E., Edalat, M., Pouyan, S., & Eskandari, S. (2021). Modeling and Prediction of Habitat Suitability for Ferula gummosa Medicinal Plant in a Mountainous Area. Natural Resources Research, 30(6), 4861–4884. https://doi.org/10.1007/s11053-021-09940-3

Polechová, J., & Storch, D. (2019). Ecological Niche. In B. Fath (Ed.), Encyclopedia of Ecology (Second Edition) (pp. 72–80). Elsevier. https://doi.org/10.1016/B978-0-12-409548-9.11113-3

Saroglou, S. T., Selvan, S. U., Windorfer, L., Weisser, W. W., Joschinski, J., Hauck, T. E., Perini, K., Mosca, F., Grobman, Y. J., & Barath, S. (2024). Utilizing Design Objectives and Key Performance Indicators as a Means for Multi-Species Building Envelopes. Buildings, 14(1), Article 1. https://doi.org/10.3390/buildings14010250

Selvan, S. U., Saroglou, S. T., Mosca, F., Tyc, J., Joschinski, J., Calbi, M., Vogler, V., Weisser, W. W., Grobman, J. Y., & Barath, S. (2023). Multi-species building envelopes: Developing a multi-criteria decision-making methodology for cohabitation. Proceedings of the 28th International Conference of the Association for Computer-Aided Architectural Design Research in Asia, 2, 645–654. https://doi.org/10.52842/conf.caadria.2023.2.643

Shen, T., Yu, H., & Wang, Y.-Z. (2021). Assessing the impacts of climate change and habitat suitability on the distribution and quality of medicinal plant using multiple information integration: Take Gentiana rigescens as an example. Ecological Indicators, 123, 107376. https://doi.org/10.1016/j.ecolind.2021.107376

Sillero, N., Arenas-Castro, S., Enriquez-Urzelai, U., Vale, C. G., Sousa-Guedes, D., Martínez-Freiría, F., Real, R., & Barbosa, A. M. (2021). Want to model a species niche? A step-by-step guideline on correlative ecological niche modelling. Ecological Modelling, 456, 109671. https://doi.org/10.1016/j.ecolmodel.2021.109671

Violle, C., Navas, M.-L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I., & Garnier, E. (2007). Let the concept of trait be functional! Oikos, 116(5), 882–892. https://doi.org/10.1111/j.0030-1299.2007.15559.x

Vogler, V., Joschinski, J., Mosca, F., & Eckelt, K. (2023). D3.3: Interim ECOLOPES Platform Architecture (Public Ares (2023)5374931; ECOLOPES: A Game-Changing Design Approach for Regenerative Urban Ecosystems). European Union. https://cordis.europa.eu/project/id/964414/results

Weisser, W., Hensel, M., Barath, S., Culshaw, V., Grobman, Y. J., Hauck, T. E., Joschinski, J., Ludwig, F., Mimet, A., Perini, K., Roccotiello, E., Schloter, M., Shwartz, A., Sunguroglu Hensel, D., & Vogler, V. (2022). Creating ecologically sound architecture by integrating ecology, architectural, and computational design. People and Nature. https://doi.org/10.1002/pan3.10411

Zimbarg, A. (2023). Mapping Plant Microclimates on Building Envelope Using Environmental Analysis Tools. In P. F. Yuan, H. Chai, C. Yan, K. Li, & T. Sun (Eds.), Hybrid Intelligence (pp. 150–164). Springer Nature. https://doi.org/10.1007/978-981-19-8637-6_13

UNCOVERING THE CIRCULAR POTENTIAL: ESTIMATING MATERIAL FLOWS FOR BUILDING SYSTEMS COMPONENTS REUSE IN THE SWISS BUILT ENVIRONMENT

SHUYAN XIONG[1], EDWIN ZEA ESCAMILLA[2], GUILLAUME HABERT[3]
[1,2,3]*Chair of Sustainable Construction, ETH Zurich.*
[1]*xiong@ibi.baug.ethz.ch, 0000-0002-4563-730X*
[2]*zea@ibi.baug.ethz.ch, 0000-0002-8677-161X*
[3]*habert@ibi.baug.ethz.ch, 0000-0003-3533-7896*

Abstract. The construction industry plays a critical role in global resource consumption and greenhouse gas emissions, highlighting the urgent need for sustainable development practices. However, a key challenge in this area is the lack of effective models for resource use that align with circular economy principles. This gap hinders efforts to achieve sustainable resource management, especially in the face of increasing urbanization and material demand. To address this issue, our study presents a Parametric Predictive Model (PPM) to improve resource efficiency, specifically targeting the often-underestimated building systems. The model takes a bottom-up approach, utilizing local databases to accurately assess material stocks of building systems, thereby improving the granularity of data on material composition. Using advanced machine learning algorithms, the model processes both categorical and non-categorical data. The output, an enriched comprehensive database can support more informed decision making in sustainable resource recovery and allocation, but also contribute to the broader goals of reducing waste and promoting resource efficiency in the built environment.

Keywords. Building Systems, Building Stock Modelling, Predictive Model, Circular Economy, Parametric Model.

1. Introduction

The construction industry significantly impacts global resource consumption, energy use, and greenhouse gas (GHG) emissions and faces challenges due to increasing material demand and the escalating urbanization dilemma. The United Nations and the Intergovernmental Panel on Climate Change (IPCC) (Hamilton et al., 2020; Rogelj et al., 2018) emphasize the urgency of sustainable urban development and transitions in construction practices to mitigate climate change effects. The circular economy (CE) is defined as a regenerative approach aiming to maximize resource utility and value (ARUP, 2016).

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 545-554.

Building Stock Modelling (BSM) serves as the core to fulfill the CE concept and allows for material management (Pasichnyi et al., 2019). It predicts future availability by measuring the materials in a system and facilitates effective management. The top-down approach using generalized archetypes, is fast and scalable but lacks traceability of specific building components and precision and may not suit all cultural and economic contexts. In contrast, the bottom-up 'building-by-building' method provides a detailed analysis of each building's material flows, offering greater precision and potential for future monitoring, though it is more costly and complex (D'Alonzo et al., 2020). Moreover, current study focuses mainly on heavy materials like concrete, often overlooking building systems which are lightweight but play an important role in total building emissions (Hoxha et al., 2021). While repurposing them might not always be economical, the easy disassembly nature of metallic parts offers practical, environmentally beneficial solutions.

The implementation of CE strategies in the building sector is hampered by the lack of comprehensive material inventory data and the lack of modern digital models for buildings facing demolition (Leao et al., 2001). The application of machine learning (ML) for assessing old buildings and facilitating material reuse is impeded by data scarcity. For instance, in Switzerland, although numerous public databases exist, obtaining detailed building data for material estimation and identifying reusable materials remains a challenge. Manual data collection is time-consuming, inefficient, and often impractical(Verellen & Allacker, 2020).

Sustainable building solutions involve integrating ML-based predictors with parametric modelling for decision making (Murphy, 2012). However, data from old buildings should be carefully considered to avoid losing valuable insights. Expanding databases with highly accurate ML predictions will improve material estimation and building stock modelling. Parametric modelling, which relies on parameters and equations (Davis, 2013), is meanwhile built by incorporating expert knowledge of building system, and can provide design flexibility, simplification of the design process, and accuracy by complying with the design rules.

This research was supported and funded by the Future Cities Laboratory (FCL) and is part of a work package focusing specifically on Switzerland case study. In this paper this was addressed by combining comprehensive datasets with advanced modelling, particularly in complementing BSM at the component level. This enabled accurate material estimates and efficient planning for material recovery, and optimizing resource use (Röck et al., 2018). The proposed methodology reduces data dependency, allowing it to be widely used with minimal dataset requirements and to predict larger quantities of materials at lower computational cost. The adaptability of the parametric model makes it a versatile open-source platform capable of integrating a wide range of building components and evolving to meet future needs. The results of the study will be further validated with a smaller sample dataset and a reuse potential framework will be developed to provide insight into material reusability, marketability, and comprehensive reuse guidelines.

2. Methodology

Parametric Predictive Modelling (PPM) development in the Python environment consists of two main components. First, the predictive model generation begin with compiling the existing federal building database (Die Schweiz in 3D, n.d.; Home / GEAK, n.d.; Swiss Geoportal, n.d.; Office, 2023), and the missing data is divided into non-categorical and categorical subsets. For non-categorical data such as building dimensions and year of renovation, a linear regression algorithm from the Scikit Learn (Pedregosa et al., 2011) toolkit is used. For categorization of building types and energy systems, a neural network implemented in PyTorch is used (Paszke et al., 2019). Secondly, the parametric model integrates key parameters from the database and generates material quantities by following building system design principles and expert knowledge. Figure 1 illustrates the general workflow and the relevant information exchanged between each step.

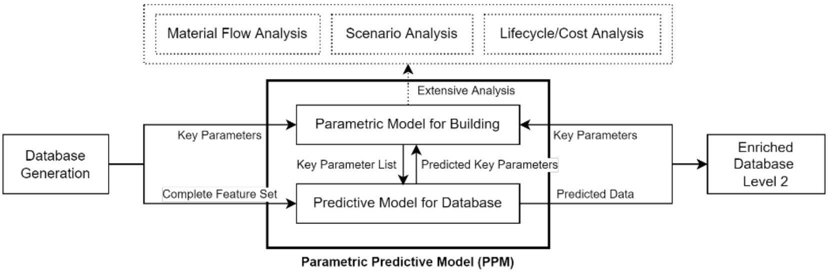


Figure 1. Workflow Diagram of Parametric Predictive Model (PPM) Development

The main categories considered were HVAC, electrical and plumbing systems. PPM initially features selected equipment such as radiators, boilers, ventilation ducts, plumbing, and electrical cables, which are chosen based on reuse experience and metal content, with high reuse potential that could provide significant carbon savings opportunities through reuse.

2.1. PARAMETRIC MODEL DEVELOPMENT

The parametric model development is demonstrated in this subsection, taking radiators as an example. The model assumes its quantity in a heated room is based on the heating system type and presence. The total weight of radiators in a building is calculated by:

$$W_{total\ radiator} = N_{radiator} \times W_{radiator}$$

The number of radiators is estimated using the heated area, number of rooms, or building's heating energy consumption, with a factor indicating one radiator per room (Bornhoft, 2024). The final count of radiators compares the number calculated by the number of rooms with that calculated by the size of the building or heating demand to ensure accuracy.

$$N_{radiator,unit} = Max(N_{radiator,area}, N_{radiator,\#room})$$

$$N_{radiator,area} = A_{unit}/(RC \times HLC)$$

$$N_{radiator,\#room} = \alpha N_{room} + N_{kitchen}$$

$$N_{radiator,building} = Q_{heating,building}/(T_{heating} \times RC)$$

$$N_{radiator} = Max(\sum N_{radiator,unit}, N_{radiator,building})$$

Where:
N_{room}: total number of rooms excluding kitchen
$W_{radiator}$: radiator unit weight
$N_{kitchen}$: total number of kitchens
A_{unit}: area of the living space of apartment
RC: radiator capacity in kW
α: number of radiators per room
HLC: heat loss coefficient in W/m^2
$Q_{heating,building}$: heating demand of the building in kWh
$T_{heating}$: heating period in hours

This approach balances detail with practicality, aiming to closely reflect real-world scenarios. To improve the accuracy of the model, connection matrices for different radiator models and typical installation years can be incorporated into the model, thereby improving the ability to replicate the real world. The approach is also applicable to other building components, thereby improving robustness and generalizability.

2.2. PREDICTIVE MODEL DEVELOPMENT

For the development of the predictive model, it is necessary to create a comprehensive database that integrates various sources with key parameters to ensure flexibility and reliability for practical applications. Key databases used include: *Gebäude- und Wohnungsregister (GWR), Gebäude- und Wohnungsstatistik (GWS), Der Gebäudeenergieausweis der Kantone (GEAK) and 3D Data of Swiss Topography (Swisstopo)* (Die Schweiz in 3D, n.d.; Home / GEAK, n.d.; Swiss Geoportal, n.d.; Office, 2023). The databases cover a wide range of buildings, with available data spanning from about 0.13 to 5 million buildings, encompassing diverse aspects such as building geometry, energy sources, and registered information.

2.2.1. Preparation of Data

To manage the extensive data in this study, various sources are consolidated into a unified database by merging the databases on unique building identifier numbers (EGID), as demonstrated in Figure 2. From the parametric model development section above, essential parameters are identified to estimate equipment quantities and materials. The data is refined and standardized for predictive model development, and harmonized by excluding non-essential details, like the recording year, to focus on relevant features. Included parameters cover aspects such as building geometries, number of floors, residents, apartments, and details about construction, demolition, renovation, building categories, and heating systems.

The data is converted back to the original format from a specialized coding system. In addition, the complexity of the building systems and building categories makes it necessary to summarize the various types of heating systems and buildings by "map" function. The simplified categories allow the study to focus on residential and office buildings, consistent with the primary interest of the study.

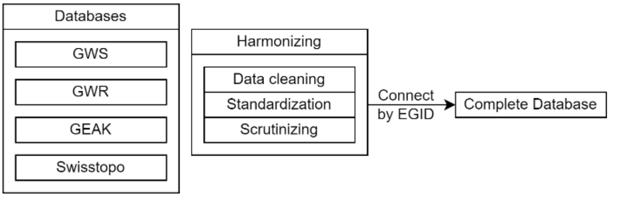


Figure 2. Data Preparation Workflow Diagram

2.2.2. Filling Data Gap

The main challenge in analysing building data lies in the large discrepancies and gaps in existing databases. Switzerland lacks comprehensive data on all buildings, with the GEAK database covering less than 5% of buildings. Even in the more inclusive GWR and GWS databases, data inconsistencies and gaps are evident, which hinders the construction of PPM and the estimation of equipment and material quantities.

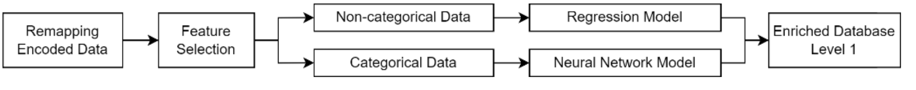


Figure 3. Workflow Diagram for Predictive Model Development and Database Enrichment

To address these gaps, especially for predicting material quantities necessary for further analysis, ML-based predictors are utilized instead of archetype-based clustering. This ML approach allows for the creation of detailed BSM at the building level and is adaptable for exploring specific categories or entire building ranges, thus bridging data gaps, and enhancing predictions for sustainable practices. As shown in Figure 3, building data is categorized into two types for prediction: categorical and non-categorical, each processed using appropriate predictors.

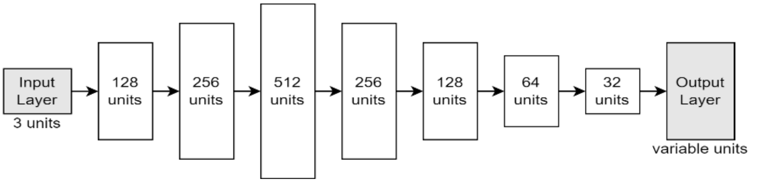


Figure 4. Neural Network Architecture for Predictor of Classification Problem

The data preprocessing step is detailed as follows. Rows with missing values are first removed to maintain the integrity of the dataset. The data is then split into features and target variables. Feature selection is performed using a cross-correlation matrix to identify the top three correlation parameters as key features. These features are then validated by expert knowledge to ensure logical predictions.

Regression methods are suited for continuous or numeric data like unique values per household, including building geometrical and occupancy details. For these non-categorical data, XGBoost regression is employed for data imputation. An XGBoost regression model customized with specific hyperparameters is used to predict and fill in missing values. The validity of the model is measured by the R-squared metric, which assesses how well the model explains the variance of the target variable.

In contrast, neural networks (NN) are highly effective in classification tasks with limited choices, such as categorizing building types and heating systems, due to their proficiency in handling predefined categories. To leverage this strength, a deep feedforward neural network is constructed using the PyTorch framework, specifically tailored for categorized data. The network is composed of several fully connected layers. Each layer incorporates batch normalization and the ReLU activation function for enhanced performance. The architecture's core features 8 hidden layers and dropout regularization is applied to prevent overfitting. Regarding the network's parameters, the batch size is set to 12,800 across 100 epochs. The output feature number of the last layer is set based on the unique category count in the target column. For the optimization process, the Adam algorithm is employed with a learning rate of 1e-3. Additionally, the ReduceLROnPlateau scheduler is utilized for optimizing learning rate adjustments.

3. Results

3.1. PREDICTIVE MODEL AND ENRICHED DATABSE

The enrichment score indicates the percentage of newly filled data fields in the database, which initially had over 40% missing values, especially in non-categorical data, due to sparse Swiss building coverage in the GEAK database from null values and limited size. Post predictive model application, database completeness significantly improved, potentially enriching data up to 95%, especially in previously underrepresented areas.

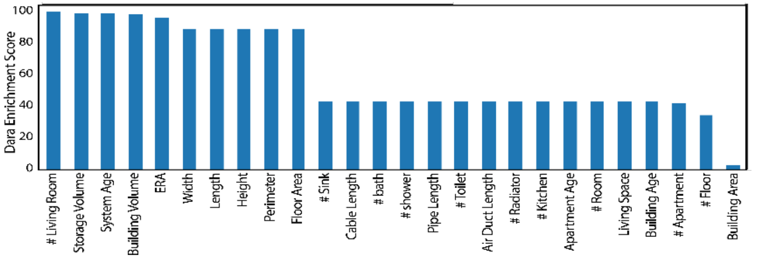


Figure 5. Enrichment Score of the Predictive Model for Non-Categorical Datasets

3.2. PARAMETRIC MODEL

Parametric model is designed to estimate the availability of the building components based on key parameters. Connection matrices are included for addressing equipment availability and variability by year, building categories etc.

3.2.1. Connectivity Matrix

Central to this model is the connectivity matrix, which enriches the database by including assumptions on inter-relationships of energy sources, building categories, and construction periods. For instance, regarding the relationship between building systems and building categories, the prevalence of Swiss residential buildings typically has no HVAC systems. To facilitate understanding, buildings categories are simplified into residential, office, and others, integrating binary availability assumptions from expert knowledge, as sampled in Table 1. Same approach is also applied to analyse relationships of equipment and energy sources, equipment (material) and installation years. This integration enhances the model with detailed layers and predicts the quantities of available equipment and related material amounts, offering flexibility for additional matrices as required.

Table 1. Sample Connectivity Matrix for Building Systems Across Different Categories

Building Systems	Component	Building Category		
		Residential	Office	Other
Heating	Radiator	1	1	0
HVAC	Air Duct	0	1	0
Plumbing	Water Pipe	1	1	1
Electrical	Electrical Cable	1	1	1

3.3. UNCERTAINTY QUANTIFICATION

3.3.1. Parametric Model

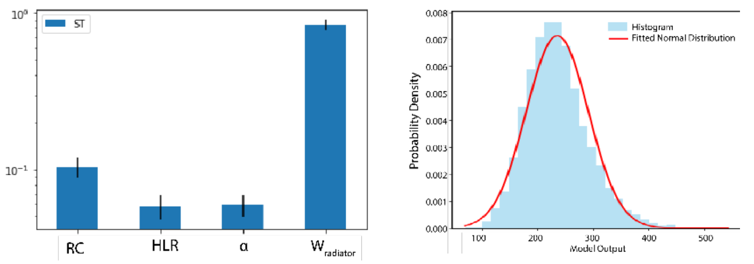


Figure 6. Evaluation of the Parametric Model for Radiator Quantity Estimation. (a) Sensitivity Analysis. (b) Probabilistic Distribution of Model Results.

To calculate radiator quantity with a parametric model, basic data like living area and room count are used, but certain values are based on assumptions, introducing uncertainty. A three-stage sensitivity analysis is conducted to mitigate this. Firstly, assumption-based parameters are refined through iterative analysis, reducing reliance on assumptions. Secondly, parameter ranges reflecting real-life scenarios are chosen, with radiator unit weight being a key factor (Figure 6(a)). Lastly, uncertainties are quantified using Monte Carlo simulation within the SciPy package(Virtanen et al., 2020), resulting in a 20% relative standard deviation. This comprehensive approach enhances the model's reliability by critically evaluating and narrowing down essential parameters and their ranges.

3.3.2. Performance of Predictor

The performance of the predictors, including both regression and neural network models, is evaluated as shown in Figure 7. The regression model's accuracy is determined using the R-squared (R^2) score, which measures the variability of model. The R^2 score for cooking equipment is 0.912 which indicates 9% variance.

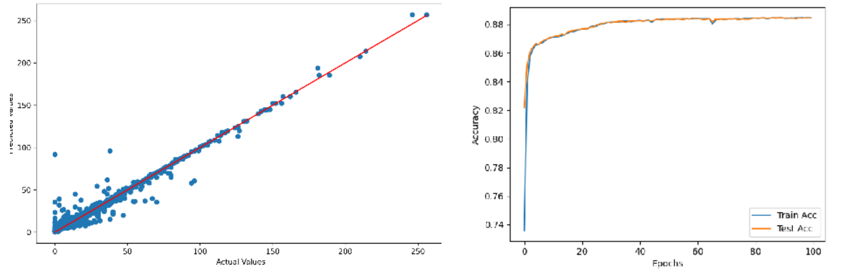


Figure 7. Performance of Predictors. (a) Scatter Plot: Actual vs. Predicted Values for the Regression Model on Cooking Facilities. (b) Evolution of Accuracy for the NN Model on Heat Generator.

For the NN-based classification model, predictor performance is assessed based on the accuracy achieved after all training epochs are completed. The accuracy of the predictor, as shown in Figure 7(b), can reach nearly 90% for determining the heat generator used in each building, crucial for understanding the availability of related heating equipment on a building-by-building basis.

4. Discussion

To accurately estimate resources, the probability of different processing streams and equipment lifetimes must be considered. Model validation involving additional data and expert knowledge is also critical. Through cross-validation and real-world data, model parameters can be adjusted, creating a feedback loop for continuous improvement.

However, the effectiveness of the method depends on the quality of the database used. The comprehensive database includes diverse characteristics of the building stock across different regions and time periods, and how the variation may affect the predictive accuracy of the model further investigated. Switzerland is unique in that it has very detailed public resources, including extensive federal register information and detailed data on building performance and energy systems. While this level of detail may not be replicable in other countries, the robustness of the PPM has been proven. Even with minimal requirements, the model improves the completeness of data characterization with a high degree of accuracy. In addition, PPM is adaptable to a wide range of situations and can be customized based on local expert knowledge and region-specific building system design rules. For example, in Singapore, where HVAC systems are prevalent, PPM can be adapted to focus on the components associated with these systems.

5. Conclusion

Current BSM approaches, primarily archetype-based, often lack detailed information on individual buildings, limiting their utility in assessing reusable materials on a local scale. This gap affects the development of decentralized markets for the circular economy. BSM has traditionally focused on structural materials and building skins, which are challenging to reuse due to labour intensity and limited environmental savings potential. However, building system materials, mainly metals, offer a longer service life and high reuse potential, yet their environmental impact and reuse potential are often overlooked.

This study highlights the importance of data-rich integrated databases in developing parametric models for the building sector. By combining various data sources and applying machine learning tools, a data-driven predictive model is developed, capable of accurately predicting building systems and materials embedded in the current building stock. While our focus is on leveraging these robust models for filling the missing values, other variants may exhibit comparable performance, which opens avenues for future research.

Despite inherent data gaps and uncertainties, this research lays the groundwork for data-driven decision-making in construction. Efforts are made to reduce uncertainty, such as minimizing assumption-based parameters and enhancing accuracy through machine learning predictions. Future potential lies in refining these models with real project data, offering flexibility to extend to various building components and aligning with design principles. In summary, this research contributes significantly to the construction industry's sustainability by enabling informed decisions in reuse, recycling, and renovation, thus promoting resource efficiency and sustainable building practices.

Acknowledgements

This research was conducted at the Future Cities Lab Global at ETH Zurich. Future Cities Lab Global is supported and funded by the National Research Foundation, Prime Minister’s Office, Singapore under its CREATE programme and ETHZ, with additional contributions from the NUS, NTU, SUTD. The processing of geometrical is conducted by ESD (ETHZ). Appreciation is extended to Huiying Zhang and Yunzhi Lin for supervision on ML and hardware support.

References

Bornhoft, A. (2024). The surprising world of swiss apartment life. Retrieved from https://www.wanderwego.com/home/14-things-about-swiss-apartments

D’Alonzo, V., Novelli, A., Vaccaro, R., Vettorato, D., Albatici, R., Diamantini, C., & Zambelli, P. (2020). A bottom-up spatially explicit methodology to estimate the space heating demand of the building stock at regional scale. Energy and Buildings, 206, 109581. https://doi.org/10.1016/j.enbuild.2019.109581

Davis, D. (2013, August 7). A History of Parametric. Daniel Davis. https://www.danieldavis.com/a-history-of-parametric/

Die Schweiz in 3D. (n.d.). Bundesamt für Landestopografie swisstopo. Retrieved Novem-ber 30, 2023, from https://www.swisstopo.admin.ch/de/karten-daten-online/karten-geodaten-online/3d-viewer.html

Hamilton, I., Kennard, H., Rapf, O., Kockat, J., Zuhaib, S., Abergel, T., Oppermann, M., Otto, M., Loran, S., Steurer, N., & others. (2020). Global Status Report for Buildings and Construction: To-wards a Zero-Emission. United Nations Environment Programme, Efficient and Resilient Buildings and Construction Sector: Nairobi, Kenya.

Home / GEAK. (n.d.). Retrieved October 23, 2023, from https://www.geak.ch/

Hoxha, E., Maierhofer, D., Saade, M. R. M., & Passer, A. (2021). Influence of technical and electrical equipment in life cycle assess-ments of buildings: Case of a laboratory and research building. Inter-national Journal of Life Cycle Assessment, 26(5), 852–863. https://doi.org/10.1007/s11367-021-01919-9

Leao, S., Bishop, I., & Evans, D. (2001). Assessing the demand of solid waste disposal in urban region by urban dynamics modelling in a GIS environment. Resources, Conservation and Recycling, 33(4), 289–313. https://doi.org/10.1016/S0921-3449(01)00090-8

Murphy, K. P. (2012). Machine learning: A probabilistic perspective.

Office, F. S. (2023, September 26). Gebäude- und Wohnungsstatistik (GWS), Geodaten 2022 | Publication. Federal Statistical Office. https://www.bfs.admin.ch/asset/en/27905171

Pasichnyi, O., Wallin, J., & Kordas, O. (2019). Data-driven building archetypes for urban building energy modelling. Energy, 181, 360–377. https://doi.org/10.1016/j.energy.2019.04.197

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., … Chintala, S. (2019). PyTorch: An Imperative Style, High-Performance Deep Learning Library. In Advances in Neural Information Processing Systems 32 (pp. 8024–8035). Curran Associates, Inc. http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research, 12, 2825–2830.

Röck, M., Hollberg, A., Habert, G., & Passer, A. (2018). LCA and BIM: Visualization of environmental potentials in building construc-tion at early design stages. Building and Environment, 140(December 2017), 153–161. https://doi.org/10.1016/j.buildenv.2018.05.006

Rogelj, J., Shindell, D., Jiang, K., Fifita, S., Forster, P., Ginzburg, V., Handa, C., Kheshgi, H., Kobayashi, S., Kriegler, E., & others. (2018). Mitigation pathways compatible with 1.5 C in the context of sustain-able development. In Global warming of 1.5 C (pp. 93–174). Inter-governmental Panel on Climate Change.

Swiss Geoportal. (n.d.). Geo.Admin.Ch. Retrieved October 23, 2023, from https://map.geo.admin.ch

Verellen, E., & Allacker, K. (2020). Using data-driven models to estimate the energy use of buildings based on a building stock model. IOP Conference Series: Earth and Environmental Science, 588(3), 032006. https://doi.org/10.1088/1755-1315/588/3/032006

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., … SciPy 1.0 Contributors. (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. Nature Methods, 17, 261–272. https://doi.org/10.1038/s41592-019-0686-2

INDEX OF AUTHORS

A

Abuzuraiq, Ahmed M. 1-455, 1-515, 2-181, 2-191
Adelzadeh, Amin 3-339
Akbaylar Hayreter, Ipek 2-283
Alkadri, Miktha Farid 2-253
Altun, Sevgi 2-313
Alva, Pradeep 2-365
Antonio, Rishan 2-435
An, Yutong 2-445
Arisman, Arman 3-29
Arjaghi, Niousha 1-353
Arlene, Vania 3-499
As Cemrek, Handan 2-283
Ataman, Cem 2-201
Aung, Ye Yint 1-363, 3-109
Austin, Matthew 2-211

B

Bai, Zishen 2-9
Bao, Ding Wen 1-293
Barashkov, Julia 2-241
Barath, Shany 1-535, 3-349
Bassing, Stefan 3-59
Bauscher, Erik 1-159
Behrouz, Tina 2-191
Belek Fialho Teixeira, Müge 3-79, 3-191
Berger, Yuval 3-349
Bittermann, Michael S. 1-19
Böhm, Michael 1-435
Brambilla, Arianna 3-241, 3-459
Burry, Mark 1-49
Buš, Peter 3-171

C

Caetano, Inês 3-9
Cahyadi Agung, Muhammad Rafif 2-253
Cai, Chenyi 2-375, 3-19
Calbi, Mariasole 1-535
Caldwell, Glenda 3-191
Calixto, Victor 3-39
Castelo-Branco, Renata 3-9
Chang, Ching-Chih 2-49
Chang, Teng-Wen 2-49
Chang, Ya-Chen 2-49
Chatterjee, Shamik Sambit 3-449
Chavan, Tejas 2-435
Chen, Ashley 3-69
Chen, Changyu 2-119
Chen, Chao-Nien 1-475
Chen, Chu-Ding 3-509
Cheng, Cesar 2-435
Cheng, Jack C.P. 1-395
Cheng, Sifan 2-139
Cheng, Tsung-Wei 1-363, 3-109
Cheng, Yiming 1-69
Cheng, Yu-Tung 3-181
Chen, Hsin 3-371
Chen, Jiajian 3-141
Chen, Jianhao 3-171
Chen, Xiang 3-281
Chen, Xiong 1-495
Chen, Yong 3-431
Cheung, Lok Hang 1-59
Chiujdea, Ruxandra 3-311
Choi, Jin Woo 2-303
Choi, Tae Ho 3-539
Cho, Ju Eun 3-391
Cimillo, Marco 1-465
Cohen, Avraham 3-349
Croffi, Juliana 3-39
Crolla, Kristof 3-181
Cuartero, Javier 1-201

D

Dabas, Yoav 3-349
Dai, Anni 1-159
Dai, Shenzhou 2-425
Das, Avishek 3-361
Dedeler, Elif Gamze 1-19
Deng, Yingxin 2-355
Donovan, Jared 3-79, 3-191

Dritsas, Stylianos 1-211, 1-231, 2-465

E
El Antably, Ahmed 1-129
El Mesawy, Mohamed 1-129
Elshani, Diellza 1-159
Eppinger, Carl 3-311
Erhan, Halil 1-455, 1-515, 2-181, 2-191

F
Fan, Bingbing 3-161
Fang, Iuan-Kai 1-189
Fang, Rong 2-505
Fan, Zhaoxiang 1-505, 2-129
Fialho Teixeira, Frederico 1-313
Fich, Lars Brorson 3-361
Fukuda, Tomohiro 2-59, 3-421, 3-469

G
Gao, Jin 1-343
Gao, Naixiang 2-415
Gao, Tianyi 1-415, 3-151
Gao, Wen 1-169
Gao, Xiangyu 3-411
Gardner, Nicole 1-1, 2-1, 2-515, 3-1
Gegana A, Gregorius A 3-499
Geropanta, Vasiliki 3-549
Gerothodoros, Chris 3-549
Globa, Anastasia 3-241, 3-459
Gocer, Ozgur 3-459
Gomaa, Mohamed 3-121
Gong, Jueqiu 3-59
Grobman, Yasha 1-445, 1-535
Guan, Kaitong 3-529
Gui, Peng 1-251
Gulec Ozer, Derya 2-283
Guo, Rizen 1-99
Guo, Yuhan 2-119
Guo, Zhe 3-281, 3-479, 3-529
Gu, Sijia 1-415, 3-151

H
Habert, Guillaume 1-545, 2-335
Hadiatmadja, Juniarto 1-261
Haeusler, M. Hank 1-303, 3-121
Han, Zhen 1-29, 1-139
Hao, Xinyuan 2-345
Harsono, Kevin 1-363, 3-109
Hartanto, Elissa 3-69
Hein, Carola 2-253
Herr, Christiane M. 3-441
Herthogs, Pieter 3-19
He, Shanshu 2-445
Heywood, Kate 3-271
Hölscher, Christoph 2-405
Hong, Cheng-Chun 2-49
Hong, Yi Yeon 3-539
Horvath, Anca-Simona 3-49
Hou, June-Hao 3-131, 3-181, 3-371
Hsain, Houssame Eddine 2-435
Hsiao, Wei-Ting 3-509
Hsieh, Po-Yu 3-131
Hsieh, Tzu-Hsin 1-271
Hsu, Pei-Hsien 1-475, 3-181
Huang, Chenyu 1-9, 1-485
Huang, Hsin-Yi 2-49
Huang, Huei-Ying 3-231
Huang, Jeffrey 2-273
Huang, Jingfei 2-395
Huang, Qixuan 3-59
Huang, Qiyue 3-59
Huang, Xiaoran 2-385, 2-425, 3-411
Huang, Yulu 2-99
Hu, Anqi 3-421
Hu, Qian 2-505
Hu, Wei 1-119
Hu, Xinyi 3-281

I
Ikeda, Yasushi 1-425
Indraprastha, Aswin 3-29
Innes, Daniel 2-211
Ishizawa, Tsukasa 1-241

J
Jahn, Gwyllim 3-99
Jeng, Tay-Sheng 1-271
Jeong, Moon-Young 1-435
Jia, Muxin 2-39

Jian, Yixin 1-9
Ji, Guohua 1-333
Jin, Chenxi 3-329
Jin, Zihao 3-329
Ji, Seung Yeul 3-489, 3-539
Johanes, Mikhael 2-273
Jordan, Mathers 1-353
Joyce, Sam Conrad 2-495
Jun, Han Jong 2-303, 3-391, 3-489, 3-539

K

Kang, Se Yeon 3-391
Karagianni, Anna 3-549
Karimian-Aliabadi, Hamed 3-339
Kasai, Ai 1-241
Katsangelou, Smaro 1-149
Kayasü, Sena 2-313
Khademi, Seyran 1-201
Khan, Md Zishaan 2-181
Kim, Dongyun 2-159
Kim, Eonyong 2-303
Kim, Frederick Chando 2-273
Kim, Hanjun 2-159
Kim, Il Hwan 2-79, 3-263
Kimm, Geoff 1-49
Kim, Mi Kyoung 3-489
Kimura, Shun 1-425
Kladeftira, Marirena 3-291
Koh, Immanuel 1-221, 3-69
Kondratenko, Aleksei 2-435
Kong, Yuhang 3-479
Koupaei, Afshin 3-301
Kozlovsky, Roy 1-445
Krncevic, Monika 1-353
Kuang, Zheyuan 2-29
Kwok, Helen H.L. 1-395

L

Lahtinen, Aaro 2-515
Lai, Ih-Cheng 3-401
Lee, Cheryl 1-211, 1-231, 2-465
Lee, Rachel 3-241
Leitão, António 3-9
Levy, Hanna 1-445
Liang, Jiadong 1-221, 2-19
Lian, Guoliang 1-415
Li, Biao 1-373, 3-19
Li, Bin 3-161
Li, Ce 3-281, 3-479
Li, Guanhong 2-169
Li, Jia-Rong 2-49
Li, Jinmin 1-99
Li, Jinze 2-375
Li, Jiongye 2-69, 2-89
Lim, Chor-Kheng 3-89
Li, Meikang 2-475
Li, Mingming 1-89
Lim, Joie 2-405
Lin, Bo-Young 1-405
Ling, Ban Liang 1-323
Lin, Shuiyang 2-109
Lin, Wei 1-405
Lin, Ying 1-79
Li, Te 2-445
Liu, BoSheng 3-231, 3-321
Liu, Gang 1-29
Liu, Huizhao 3-59
Liu, Jiaxin 2-385
Liu, Mengxuan 1-505, 2-129
Liu, Shengyu 2-19
Liu, Siyan 2-445
Liu, Yongkang 2-149
Li, Xiaoqian 1-29, 1-139
Li, Yang 2-355
Li, Yilin 1-383
Li, Yuechun 2-221
Li, Yunqin 2-29
Li, Zao 3-529
Loh, Paul 3-99
Loke, Lian 3-381
Long, Yini 2-445
Lorberbom, Shachar 1-525
Lotfian, Shabnam 3-191
Lo, Tzu-Hsien 3-321
Lo, Wei Che 3-201
Luke, Thom 1-39
Lu, Owen Zhiyuan 1-303
Luo, Yilu 1-99
Lu, Shuai 1-99, 1-169
Lu, Weisheng 2-231

M

Madsen, Claus B. 3-361
Makki, Mohammed 1-353
Mao, Yujun 1-179
Masuda, Lynn 3-381
Matheou, Maria 1-435
Mavros, Panos 2-405
Ma, Zhichao 1-495
Meng, Leo Lin 1-303
Meng, Yu 2-325, 2-455
Meng, Yuan 2-505
Miao, Shuhan 1-109
Middlemiss, Aleisha 2-211
Mo, Fangshuo 1-415
Mohd Taib, Rusdy Bin 3-519
Mok, Jun Wei 2-405
Morgan, Tristan 2-211
Mostafavi, Fatemeh 1-201
Mostafavi, Mehdi 2-293
Mosteiro-Romero, Martín 2-365
Mottaghi, Esmaeil 1-455
Moulaii, Mohammadmehdi 2-293
Mo, Yichen 1-373

N

Nagakura, Takehiko 1-109, 1-179
Nakapan, Walaiporn 3-519
Narahara, Taro 2-39, 2-293
Natanian, Jonathan 1-525
Nazim, Ibrahim 2-495
Ng, Provides 2-221
Nicholas, Paul 3-271, 3-311
Nisan, Alon 3-349
Nisztuk, Maciej 2-435

O

Özkar, Mine 2-313
Öztürk, Berkay 2-313

P

Pados, Dimitris A. 1-149
Palamas, George 3-49
Parascho, Stefana 3-291
Park, Jong Jin 2-303
Parthenios, Panagiotis 3-549
Patel, Sayjel Vijay 1-343, 2-435
Pei, Lulu 3-441
Pei, Wanyu 2-335
Peng, Wenzhe 1-109, 1-179
Peng, Yufei 1-415
Peng, Ziyu 2-231
Perrault, Simon 2-201
Pia, Fricker 2-19
Pittiglio, Alexandra Nicole 3-291
Pouliou, Panagiota 3-49

Q

Qian, Xidong 3-329
Qiu, Waishan 2-99
Qu, Rong 1-485

R

Rafizadeh, Hamidreza 3-79
Rahman, Muhammad Arif 2-253
Ramos Jaime, Cristina 1-303, 2-515
Ramsgaard Thomsen, Mette 3-311
Reinhardt, Dagmar 3-381
Robeller, Christopher 3-339
Roccotiello, Enrica 1-535
Rössle, Reiner 1-435

S

Şahin, Yusuf 2-313
Sanatani, Rohit Priyadarshi 3-449
Saroglou, Soultana T. 1-535
Sasatani, Masamichi 3-141
Sawodny, Oliver 1-435
Schork, Tim 3-79
Shafiei, Maryam 1-313
Shao, Tong 1-333
Shen, Yanting 1-485
Shiel, John 3-121
Shih, Shen-Guan 1-189, 1-363, 3-109
Shi, Ji 1-373
Shimizu, Sayuri 3-469
Shinohara, Hiroyuki 3-141
Shi, Yi'an 1-59
Shi, Yinyi 1-495
Shi, Zewei 1-9
Shu, Kun Di 1-293

Skevaki, Eleni 3-291
Sklar, Eilam 1-525
Song, Qiwei 2-99, 2-475
Song, Zhehao 2-375
Sonne, Konrad 3-311
Stouffs, Rudi 1-29, 1-139, 2-69, 2-89, 2-335, 2-365, 2-405
Sudradjat, Iwan 3-29
Sun, Chengyu 1-383, 2-325, 2-455, 3-431
Sun, Jiapan 3-161
Sun, Kenan 3-171
Sun, Mingli 3-171

T
Tan, Benjamin 3-519
Tan, Chuheng 3-211
Tang, Chao 3-441
Tang, Peng 2-263, 2-375, 2-505
Tang, Shuoning 1-505
Tan, Linus 1-39
Thiodore, Jacky 3-499
Tian, Dehui 3-59
Tian, Runjia 1-69, 2-169
Tian, Shulin 3-221
Tohidi, Alexander 3-121
Tong, Ziyu 2-345, 3-221
Trivic, Zdravko 2-405
Tsai, Daniel 1-109
Tseng, Jui-Ho 3-231
Tsiliakos, Marios 3-59
Tu, Han 2-395
Tunçer, Bige 1-323, 2-201, 2-405

U
Ünal, Gözde 2-313
Urroz, Celso 3-253
Uthaya Selvan, Surayyn 1-535
Uzun, Fatih 2-313

V
van Ameijde, Jeroen 2-109, 2-139, 2-221, 2-475
Vermisso, Emmanouil 1-149

W
Wagiri, Felicia 1-363
Wang, Cunyuan 1-495
Wang, Jinyu 1-9, 1-485
Wang, Jun 1-99
Wang, Likai 1-59, 1-333, 1-465
Wang, Rixin 1-415
Wang, Ruishu 1-415
Wang, ShaoMing 1-99
Wang, Shih-Yuan 3-201
Wang, Weili 1-251
Wang, Wenjie 2-445
Wang, Xiao 2-263, 3-19
Wang, Xiaoxin 1-9
Wang, Xinru 2-325, 2-455
Wang, Yanyu 1-373
Wang, Yi 2-149
Wang, Yu 1-9
Wang, Yueli 1-383
Wang, Yujiao 2-505
Wang, Zhe 1-395
Weijenberg, Camiel 2-435
Welch, Chris 2-211
Wen, Jian 2-375
Wen, Shu 2-445
Wen, Yi 1-283
White, Marcus 1-49, 2-425
Widiastuti, Indah 3-29
Wilmoth, Parker 1-149
Windorfer, Laura 1-535
Wong, Nichol Long Hin 3-181
Wortmann, Thomas 1-159
Wu, Yen-Liang 3-509
Wu, Yufei 1-415, 3-151
Wu, Zhaoji 1-395

X
Xiao, Yijun 2-485
Xia, Shengtao 1-69
Xie, Yuchen 2-29
Xiong, Shuyan 1-545, 2-335
Xu, Hang 2-169
Xu, Jintu 2-169
Xu, Pengjian 1-169

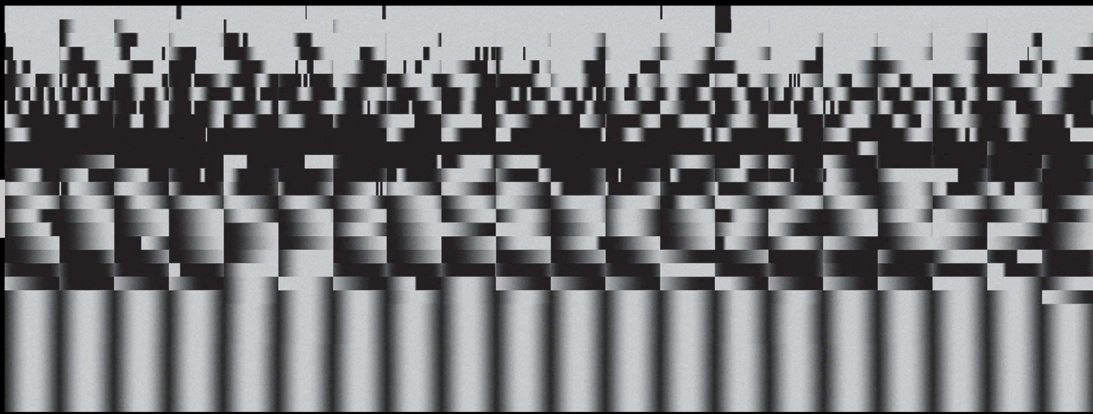
Xu, Qianbin 1-313
Xu, Weishun 1-89, 3-329

Y

Yabuki, Nobuyoshi 2-59, 3-421, 3-469
Yan, Chao 1-415, 2-445
Yang, Chunxia 2-129
Yang, Hong-bin 2-273
Yang, Runyu 1-251
Yang, Xuyou 1-89
Yang, Yang 1-465
Yan, Hongming 2-385
Yan, Xin 1-293
Yao, Jiawei 1-9, 1-485
Ye, Fei 1-79
Ye, Jiahong 1-485
Yeo, Khystelle 1-211, 1-231, 2-465
Ye, Wei 1-283
Yildirim, Muhammed 3-459
Yildiz, Burak 1-201
Yin, Jun 1-169
Yi, Xinyi 3-529
Yousif, Shermeen 3-253
Yuan, Philip F. 1-415, 3-151
Yuan, Sinan 2-485
Yuan, Ye 1-139
Yu, Gang 1-283
Yu, Hanting 2-119
Yu, Jiaxi 2-59
Yu, Kuai 2-515
Yuliana, Yuliana 2-253
Yu, Shufan 3-221
Yu, Zhichao 2-345

Z

Zaher, Nawal 1-129
Zea Escamilla, Edwin 1-545
Zeng, Pengyu 1-169
Zhang, Baizhou 1-373
Zhang, Chao 1-373
Zhang, Chen 2-9
Zhang, Jiahao 2-29
Zhang, Jiaxin 2-29
Zhang, Jingyun 1-99
Zhang, Kaiheng 2-39
Zhang, Lanming 2-445
Zhang, Li 3-411
Zhang, Yi 2-231
Zhang, Yimeng 3-161
Zhang, Yunsong 3-221
Zhang, Zihuan 3-479, 3-529
Zhang, Zixi 3-479
Zhao, Yi 3-381
Zhong, Chuwen 1-59
Zhong, Haipeng 3-211
Zhong, Ximing 1-221, 2-19
Zhou, Quan 3-211
Zhuang, Junling 2-169
Zhu, Shutong 2-221
Zhu, Xuerong 2-263
Zhu, Yue 3-281
Zou, Mingyan 3-431



Can and should design be accelerated? To what extent and to what end? Or perhaps, design must decelerate instead? Accelerated Design is an urgent call for a critical reflection of and creative action by architecture during this challenging time of accelerating climate crisis, unrestricted data surveillances, generative AI copyright infringements, global geopolitical conflicts, hyperconcentration of digital power, post-pandemic mental health deterioration, and widespread disinformation attacks.

CAADRIA2024 seeks contributions in addressing the conference theme by discussing and debating the role of design and designers in the midst of accelerated changes brought about by and on technology, economy, environment, and governance, to construct new ways of thinking, teaching, researching and practising architecture in the age of artificial intelligence and climate change.

The 29th 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 towards an inclusive future for humans and non-humans.

ISBN 978-988-78918-1-9

9 789887 891819

The Association for Computer-Aided Architectural Design Research in Asia (CAADRIA), Hong Kong
ISBN: 978-988-78918-1-9
ISSN: 2710-4257 (print)
2710-4265 (online)
Printed in Singapore, SINGAPORE