Integrating human perception in 3D city models and urban digital twins

Binyu Lei[a], Xiucheng Liang[a], Filip Biljecki[a,b]

[a]Department of Architecture, National University of Singapore, Singapore – (binyul, xiucheng)@u.nus.edu, filip@nus.edu.sg
[b]Department of Real Estate, National University of Singapore, Singapore

Keywords: Data integration, Data standard, Data extension, CityJSON, Architectural perception, 3D GIS.

Abstract

Urban digital twins, and 3D city models underpinning them, provide novel solutions to urban management but tend to overlook the human element. The trending research on human perception reveals people's perspective towards interpreting and experiencing the built environment. Advancing the representation of building physics and descriptive information in 3D city models and urban digital twins, we establish the addition and integration of the notion of how humans perceive buildings. Unlocking a new dimension in our domain, this new concept can facilitate a broader adoption of semantic 3D data in socio-economic fields across various domains, and advance existing use cases in 3D GIS. This work is the first instance of integrating such attributes in 3D city models, which have traditionally been confined to physical and objective measures. The visual perception of each building is evaluated based on building images extracted from street view images. We add such information as new attributes to an existing CityJSON dataset representing thousands of 3D buildings in Amsterdam, the Netherlands. To facilitate a robust and sustainable integration, we develop a CityJSON Extension to accommodate the new data and validate its schema successfully, and we visualise the semantic 3D dataset. Further, we present two use cases to demonstrate the usability of our new data for downstream analysis. One is the concurrent clustering of buildings based on 3D morphology and human perception, while the other is conducting an attribute-based query that enables various stakeholders to identify a particular building of interest combining both traditional and perception attributes.

1. Introduction

Urban digital twins have permeated through multiple domains, from transportation, disaster management, energy consumption to virtual tourism (Dembski et al., 2020, Ferré-Bigorra et al., 2022, Hämäläinen, 2021, Nochta et al., 2021, Ketzler et al., 2020, Rahmadian et al., 2023, Caprari et al., 2022). Integrating a large collection of data and techniques, urban digital twins advance urban analytics towards a dynamic and interactive way of building and managing smart cities. However, they are described to be techno-optimistic but overlook social and human aspects (Lei et al., 2023a, Stoter et al., 2021, Charitonidou, 2022, Nochta et al., 2021). Among various human aspects, perception has been studied in strands of domains, from medicine to social science (Redelmeier and Dickinson, 2011, Galesic et al., 2021). In urban research, humans play a role in delivering their sensations and experience towards the living environments (Goodchild, 2007, Zhang et al., 2018). Exploiting the active and explicit information from a human perspective may bridge the gap in urban digital twins, leading to a human-centric direction for further applications across multidisciplines.

From the standpoint of buildings, which are a key element in both the physical environment and urban digital twins, integrating perception and sentiments as additional building characteristics will be valuable in multiple ways. Buildings serve as the foundation of urban environment for a variety of use cases. Their physical facets, such as economic value, energy consumption and building geometry, are measurable and quantifiable for scientific studies (Zhao and Magoulès, 2012, Biljecki and Chow, 2022, Liang et al., 2023, Helbich et al., 2013). However, the less objective aspects of buildings are difficult to evaluate, e.g. emotions and perceptions towards them, which are associated with the architectural aesthetics (i.e. the art of buildings) and the surroundings. Buildings as objects in the city are identifiable, nevertheless the experience of buildings and individual perceptions are dynamic and personal, reflecting how people define and interpret them (Grütter, 2020, Kotseruba et al., 2016, Lu and Sperling, 1995).

Perception- and sentiment-related urban studies are prevalent topics, contributing to understanding urban environments and reflecting activities in the city through a human perspective (Wei et al., 2022, Ma et al., 2021, Liu et al., 2016). The rise of urban sensing enables a gathering of insights from humans regarding how they perceive their living environments, such as generating urban vitality through volunteered geographical information, e.g. trajectory data, and quantifying urban characteristics from social sensing data, e.g. social media platform, such as Airbnb reviews (Yan et al., 2020, Huang et al., 2020, Galesic et al., 2021, Liu et al., 2015). From a visual perspective, street view imagery (SVI) serves as a growing data source, collecting human perception on a large scale to evaluate urban settings and promote liveability (Kang et al., 2020, Liang et al., 2023). However, incorporating human sensing data in 3D city models and urban digital twins to a certain extent is limited. The state of the art of a human-centric urban digital twin puts a focus on leveraging it as a digital platform to represent cities to collect people feedback, e.g. supporting participatory planning and simulating community resilience (Dembski et al., 2020, Ye et al., 2023, Nochta et al., 2021). Several attempts are made to include human sensing data in an urban digital twin rather than adopting it as a visualisation tool, for example, integrating outdoor comfort data to reflect the impact of urban surroundings (Liu et al., 2023, Lei et al., 2023b). However, there remains a gap in making human sensing data analysable when adopting urban digital twins. In this work, we integrate human perceptual criteria previously used to evaluate building exteriors (Gifford et al., 2000, Heath et al., 2000, Ghomeishi, 2021, Arslan and Yıldırım, 2023) with a computer vision approach. The human perception of buildings is quantified with a score from 0 to 10 based on 5 perceptual indicators (i.e. original, pleas-

https://doi.org/10.5194/isprs-annals-X-4-W5-2024-211-2024

ing, ordered, boring and complex), retrieved from our ongoing perception SVI dataset. Considering the position of buildings, incorporating perception data with traditional building attributes can contribute to a valuable concept, which will enrich semantics of 3D building data and amplify a human-centric direction in the landscape of urban digital twins. Further, combining building physics with human perception into the same dataset have the potential to support downstream studies, such as predicting housing value and promoting human wellbeing.

In this paper, we put forward the idea of integrating such information into semantic 3D city models, and by extension — urban digital twins. Semantics in 3D city models, which are a critical pillar in urban digital twins, are the foundation for urban analysis and simulation. Therefore, we believe that an integration of human-generated information can be a motivation for adopting urban digital twins in the socio-economic realm, e.g. supporting participatory planning and collaborative urban design. We elaborate a motivation and present a comprehensive proof of concept that includes an implementation, together with standardisation, visualisation, and application of such information. In particular, we focus on data standardisation and interoperability, which have perennially been common challenges in 3D city models and urban digital twins amid the growing variation of urban data and their different provenances, impacting operation and adoption. Mitigating such issues, existing standards such as CityGML and CityJSON can be extended through mechanisms such as Application Domain Extensions (ADEs) to facilitate adopting 3D buildings to serve various purposes (Gröger and Plümer, 2012, Ledoux et al., 2019, Biljecki et al., 2018, Kolbe et al., 2008). For example, extensions to these standards have been developed to accommodate additional data pertaining to noise, energy, and metadata, which have enhanced their use in specific application domains (Gröger and Plümer, 2012, Agugiaro et al., 2018, Labetski et al., 2018). However, to the best of our knowledge, the integration of human perception represents a void in the realm of 3D data and international geospatial standards. With the goal of a broad adoption, we believe that perception data, once collected and associated to urban structures such as buildings, should be stored following an international standard and organised with a clear and valid schema, ensuring sustainability and enabling effective information exchange across cities and domains. In this context, we believe that an inescapable task in this line of work is developing an extension for human perception data in semantic 3D buildings to encode such new data.

This paper introduces a novel and robust method to integrate human perception of buildings in 3D city models and urban digital twins, with a focus on standardisation, validation, and usability through visualising such data and employing it in use cases we develop. Two use cases are conducted to demonstrate practical values of integrating perception data with semantic 3D data. The first use case clusters buildings based on their 3D morphology and perception scores, intended to characterise urban buildings using both objective and subjective aspects, a novel type of amalgamation in the 3D GIS body of knowledge. The second use case creates an attribute-based query for identifying a building of interest, based on the introduced set of new perceptual attributes in combination with traditional attributes such as year of construction of a building, which allows a variety of stakeholders (e.g. the scientific community, the general public and decision makers) to identify particular buildings of interest. Investigating the potential of our new data in these two use cases, we warrant its usability to support downstream applications in urban and geospatial domains. Beyond a technical aspect and a focus of physical elements, enriching 3D buildings with subjective building information addresses the current lack of considering a human-centric aspect, thereby encouraging a broader adoption related to social and economic development in the city. With such added values, this multidisciplinary work contributes to various domains spanning from 3D GIS to computational social science, proposing the potential to foster collaborations across disciplines.

2. Methodology

2.1 Workflow

The workflow is straightforward so it can be reproduced and easily adopted in future research. A number of key steps are illustrated in Figure 1, which will be explained in the following sections in detail. The process includes three main parts. First, perception data and 3D building models are prepared for data integration. The data on human perception of buildings is adapted from our ongoing building perception dataset, leveraging street view imagery, crowdsourcing, and computer vision. Regarding 3D buildings, we use CityJSON, a semantic data format, which stores 3D models in a standardised way and allows semantic information, thus enhancing usability for further analysis and allowing us to extend it. Second, the integration process involves matching the location of buildings from SVI with those in 3D building models. Third, to spur adopting the extended 3D building data for applications, it is important to store the new data according to a standard and validate the generated dataset and its schema. Hence, this step ensures that our new data is not only reusable but also interoperable for future use, hopefully fulfilling the promise of human sensing data in 3D GIS. As part of this validation process, later in Section 3.1 we also demonstrate that the generated data can be visualised according to common practices, which further warrants its usability.

2.2 Data collection

In this proof of concept, the Netherlands is considered as the study area considering its good quality and availability of 3D and other data. Two data sources are used in this paper — perceptions of buildings and semantic 3D city models. Figure 1 illustrates an example of the content of both datasets.

First, street view imagery is leveraged as the base source to identify and extract images of individual buildings in Amsterdam. This process collected 30,356 images of buildings from 10,210 Google Street View images. A selected portion of this data is then used to conduct a pairwise comparison survey to evaluate building façade from five aspects: complexity, originality, order, pleasantness, and excitement. As a preliminary experiment, the responses are gathered from 49 individuals, leading to the evaluation of images of 600 buildings on a scale from 0 to 10. Subsequently, ResNet models (He et al., 2016) are developed and trained using these images and their perceptual scores respectively, achieving an average RMSE of 1.82 across all models. The trained models are then employed to generate perception scores for the entire collection of building exterior images extracted from Amsterdam.

Second, 3D buildings from city models and urban digital twins are used for obtaining semantics and visualisation. The 3DBAG is a country-wide open dataset in the Netherlands developed by

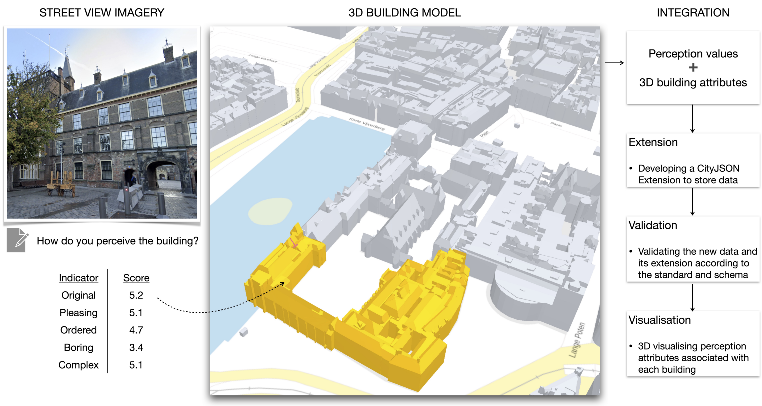

Figure 1. The workflow of integrating human perception of buildings in 3D city models and urban digital twins. Sources of data: Google Street View and 3D BAG.

TU Delft, which provides public access to 3D building models (Peters et al., 2022). Allowing free downloads in different formats, it contains three levels of detail (i.e. LoD1.2, LoD1.3, and LoD2.2), and it is of high quality (Dukai et al., 2020, Dukai et al., 2021). We download 3D buildings from 3DBAG dataset in CityJSON, which provides a compact and use-friendly format, supporting easy visualisation and extensions (Ledoux et al., 2019, Vitalis et al., 2020b).

2.3 Data integration

In processing human perception data from the dataset, each building is characterised by the aforementioned 5 pairs of perceptual indicators and numeric scores: original, pleasing, ordered, boring and complex. Such features describe how people perceive buildings across multiple dimensions of interpreting their architecture and condition. This human perception data is then incorporated as a set of extended attributes in 3D buildings.

To support standardisation and usability, we develop a CityJSON Extension to accommodate the new data (i.e. perception attributes associated with each building) and store the attributes according to the Extension. Following CityJSON Specifications 2.0.0 (Ledoux et al., 2019), we define the Extension as *Perception* to model human perception of buildings. The perception attributes are stored as `extraAttributes` in existing building objects. Each perception feature starts with a '+', structured with the feature name and value, e.g. `+perception-originality` (see in Figure 2). We adapted the 2.0 version of CityJSON to create the Extension schema.

```
{
  "type": "CityJSONExtension",
  "name": "Perception",
  "description": "Extension to model the human perception",
  "url": "https://raw.githubusercontent.com/binyulei/extension/main/perception.ext.json",
  "version": "0.5",
  "versionCityJSON": "2.0",

  "extraAttributes": {
     "Building": {
        "+perception-originality": {
          "value": "number"
        },
        "+perception-pleasing": {
          "value": "number"
        },
        "+perception-ordered": {
          "value": "number"
        },
```

Figure 2. An excerpt of the developed CityJSON Extension for accommodating human perception data.

By referencing the Extension with a publicly available URL and incorporating it into the existing 3D buildings, the new CityJSON files are validated using `cjio` and an online validator[1] (as shown in Figure 3). Hence, the perception data is stored in compliance with a standard and a valid schema as well. The process provides an initial venture to integrate human sensing data into 3D buildings. Further, other types of human-related data can be also incorporated in 3D city models and urban digital twins following this instance.

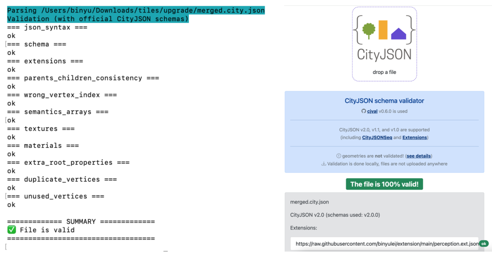

Figure 3. Validating the generated dataset and its schema of the CityJSON Extension of building perception.

2.4 Implementation

The workflow is implemented in Python:

- Google Colab[2] environment for processing perception data, creating an Extension, and conducting use cases;
- `cjio` for processing CityJSON models, upgrading data version, merging files, analysing and validating;
- `scikit-learn` library (Pedregosa et al., 2011) for clustering 3D buildings based on morphology and perception;
- QGIS[3] software for 2D visualisation and spatial analysis (QGIS.org, 2023);
- `CityJSON Loader` plugin for loading CityJSON files in QGIS (Vitalis et al., 2020a);
- `ninja`[4] as a a web application for visualising 3D models and investigating buildings based on attributes.

[1] https://validator.cityjson.org/
[2] https://colab.research.google.com
[3] https://www.qgis.org
[4] https://ninja.cityjson.org/

3. Results

3.1 Generating a sample large-scale dataset, visualisation and investigation

We select 13 tiles from 3DBAG dataset in Amsterdam as our study area, which includes a sample of more than 10,000 buildings. By adding the developed Extension of building perception in CityJSON files, each 3D building incorporates a set of new attributes, with scores indicating the degree of human perception from 5 aspects (ranging from 0 – 10). Figure 4 shows a visualisation of the dataset, using ninja (Vitalis et al., 2020b), in which a building of interest is highlighted and its extended semantic content is revealed, together with the original attributes inherited from the 3DBAG dataset, e.g. type of roof.

Such data may be found useful by a variety of potential applications. For instance, in the realm of urban digital twins, 3D buildings as the backbone play a role in mirroring the reality and are central to conducting various simulations. The practice of urban digital twins is commonly found in, for example, managing energy consumption and simulating urban climate, using objective building information (i.e. geometry and semantics) (Zhao and Magoulès, 2012, Katal et al., 2022, Wang et al., 2021). Integrating building perception data in urban digital twins can further their use in socio-technical applications, such as support participatory planning through crowdsourced knowledge rather than empirical experience or improving walking comfort at the street level (Nochta et al., 2021, Liu et al., 2023). Therefore, to corroborate these promises, we propose two use cases to illustrate the usability of our new data in urban analysis and 3D GIS. The first one is on categorising 3D buildings based on morphological metrics and perceptual values, aiming at a collection of building clusters differentiated from physics and sentiments. The second case will conduct an attribute-based query for identifying particular buildings for further purposes, such as those are constructed before 1980 and considered highly visually pleasing. As such, our new data not only complements building information but also substantiates its adoption in 3D urban and building research.

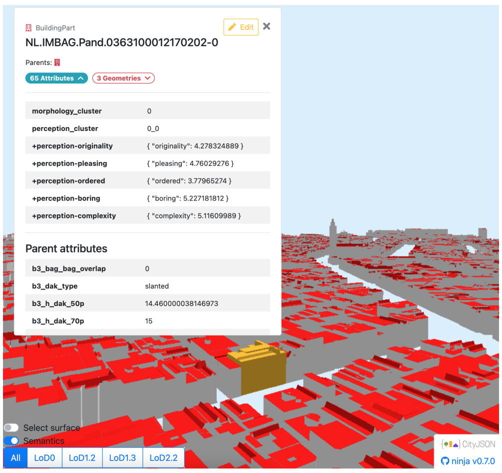

Figure 4. Visualising a 3D building of interest by displaying its existing and extended attributes. Our work presents the first concept and implementation of such data in 3D city models.

3.2 Use case: multi-source clustering

Figure 5 illustrates the density of buildings for each perception feature. As shown, the distribution for 'boring' buildings is broader with a peak near 5, which indicates a greater divergence in individual perception. Some buildings are likely to be perceived as uninteresting or lacking in engaging design, while others are not seen as an unexciting. The density distribution is insightful, reflecting a relationship between building morphology and human perception. Indeed, buildings perceived as complex can correlate with their innovative design and structures, making them intriguing. Meanwhile, it suggests the subjective nature of human perception varies from architectural style and urban interpretation. Therefore, we delve into clustering buildings based on physical morphology and subjective perception.

In contrast to perception of buildings, the study of urban morphology has been explored for decades, providing means to quantify and characterise the geometrical characteristics of buildings and other urban elements (Arribas-Bel and Fleischmann, 2022). In this use case, we develop the marriage of the traditional physical morphology with perception. By combining objective and subjective metrics, we provide a holistic means to characterise buildings and classify them. This use case is a novelty in the body of knowledge, contributing to understanding and comparing building form from a human perspective.

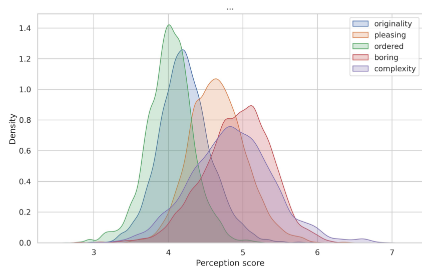

Figure 5. The density distribution of buildings categorised by each perception feature.

First, we cluster buildings using 3D building metrics, retrieved from an openly released dataset containing 2D and 3D metrics for buildings in the Netherlands (Labetski et al., 2023). By applying a hierarchical clustering algorithm, we define 4 clusters of buildings based on a selection of 3D indices. Second, within each morphology cluster, we further divide buildings into 2 subclusters using perception values. It results in a total of 8 building clusters, taking into account both 3D morphology and subjective perception. The cluster labels are subsequently added as additional attributes to the building objects in CityJSON for investigation.

The distribution of clusters based on 3D building morphology is illustrated in Figure 6. From a spatial-explicit perspective, it can be concluded that buildings in cluster 2 tend to be intricate and unique, whereas buildings in the remaining three clusters exhibit more regular forms. However, a detailed comparison reveals nuanced differences. Buildings categorised in morphology cluster 0 occupy larger land areas and are predominantly located in the city centre. Instead, cluster 1 and 3 consist of buildings concentrated in suburban areas (e.g. residential districts), regardless of their footprint sizes.

Subsequently, we visualise the clustered buildings via ninja (Vitalis et al., 2020b), styling them in different colours according to their cluster labels. Taking a building of interest as an example (as demonstrated in Figure 7), it has two attributes related to

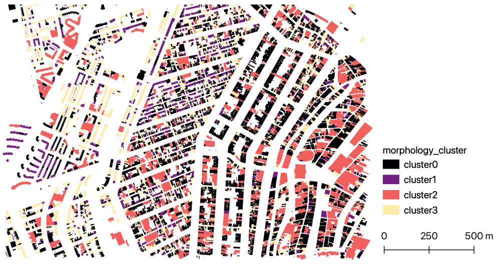

Figure 6. The spatial distribution of morphology clusters in selected areas of Amsterdam.

clusters: 'morphology cluster 2' and 'perception cluster 2_0', where *2* indicates the label of a morphology cluster and *0* suggests the perception cluster label). For example, in cluster 2_0, buildings are likely to be situated on large lots and irregularly shaped from a morphology aspect. Meanwhile, they tend to be perceived highly pleasing from a human perspective, related to iconic and intricate architectural designs. Therefore, urban buildings can be depicted using cluster labels in a comprehensive way, encompassing information from physical elements to aesthetic features.

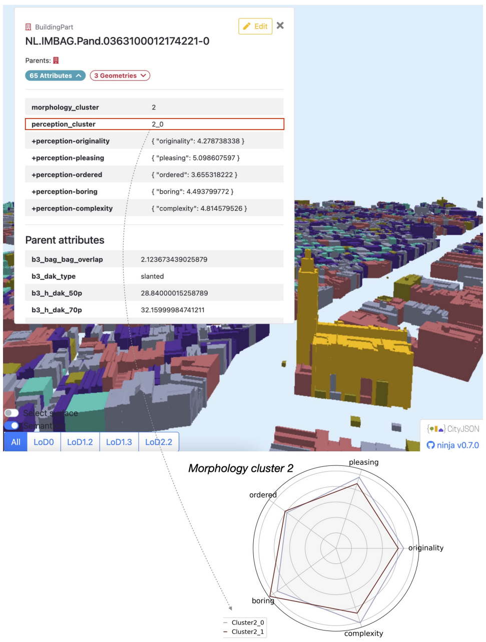

Figure 7. Visualising and styling buildings based on clusters of morphology and perception. The top plot illustrates cluster labels associated with 3D buildings. The bottom plot is a radar chart, demonstrating a comparative analysis between perception clusters.

3.3 Use case: an attribute-based query

In addition to clustering buildings with the extended semantics, we propose another use case related to identify a particular building based on its attributes. By integrating human perception into buildings, we can compile informative profiles for each building, which may be of use to different application domains. For example, each building in the semantically rich 3DBAG dataset includes an attribute that documents its construction year. Combining it with the newly added perception features, stakeholders can easily query for a building of interest (e.g. a building which is built before 1950 and is highly pleasing). It can be conducted in various tools, e.g. using desktop GIS software such as QGIS (Vitalis et al., 2020a), as illustrated in Figure 8. Therefore, this use case suggests a rather general use of our new data. When a study requires a collection of buildings with specific information (e.g. focusing on a group of historical buildings with complex architectural design), practitioners and researchers can perform a query based on building attributes derived from our work. Meanwhile, it allows the general public to explore and understand buildings of their interest. As such, the extended 3D dataset we introduce in this paper helps to advance descriptive information of buildings, which can further support city decision-making, enhance public knowledge, as well as facilitate the adoption of 3D building models among various groups.

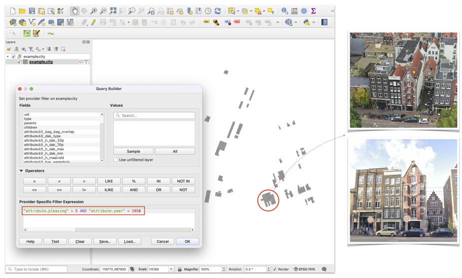

Figure 8. An attribute-based query for identifying a more complex and less boring building in QGIS. Source of the imagery: Google Maps 3D mode and Google Street View.

4. Discussion

4.1 Contributions

As the concept of humans as sensors is explored across numerous domains (Jayathissa et al., 2020, Liu et al., 2015, Dang et al., 2020, Wang et al., 2020), it becomes evident that integrating human sensing data has the potential to represent socio-economic activities in the city. Urban digital twins and 3D city models as growing tools boost the use of diverse data and techniques in solving urban problems. However, the human perspective is not fully discussed in such a techno-optimistic domain. Therefore, we explore the incorporation of human sensing data in 3D city models and investigate the usability of the new data in urban and geospatial analysis, shedding light on 3D building data enrichment with human-related phenomena. Given the role of buildings in both virtual 3D models and physical environments, we seek to understand how people perceive urban buildings and include this human-generated data in 3D buildings, which are the backbone of urban digital twins. As such, a 3D building can be characterised not only by physical attributes but also by dynamic perceptions. The concept of integrating human perception in 3D building data contributes to a variety of domains.

First, human perception of buildings is categorised into 5 aspects using street view imagery. Taking advantage of crowdsourcing, street view imagery enables people to generate their perceptions by interpreting buildings across multiple dimensions, e.g. architectural aesthetics, building façade, and the dialogue between buildings and the urban surroundings. Such diverse observations offer insights for interpreting space and places, underpinning urban research towards a human-centric direction. By integrating human perception in 3D models, we introduce a set of new perceptual attributes for each building, describing its subjective features. That is, beyond building physics, we establish a new concept that incorporates personal and diverse experience of buildings, which may lead to an expansion of adopting urban digital twins and 3D city models for socio-economic topic in various disciplines.

Second, this concept contributes to data standardisation and interoperability in 3D city models and urban digital twins. To ensure acceptance and enhance the usability of our new data, we standardise it by storing it according to a standard and validating the generated data and its schema. To do so, we use CityJSON and create an Extension that accommodates human perception data associated with each building, storing it as new attributes within the 3D building models. Validating the new data ensures consistency between the added perception attributes and existing building information, thus maintaining both syntactical and semantic accuracy. Further, the standardised and validated new 3D dataset ensures data quality for future interoperability among stakeholders and software systems, thereby enabling broader usability of our new concept and supporting scalability for further applications. The new concept paves a way for introducing custom data into a human-centric urban digital twins, such as human-defined building style, satisfying different uses and specific needs in future.

Third, the use case of clustering buildings brings a contribution to studying the urban form, a long-running line of research that has heavily relied on geospatial data. When studying the urban form, current research primarily focuses on building geometry. Nevertheless, place understanding in the urban context varies from people and is associated with activities in the surroundings as well. We make a pioneering attempt to interpret urban form through a human-centric lens. Using morphological metrics and perception values, we categories urban buildings into 8 clusters. That is, by adding perceptions, urban form and buildings can be represented from a novel perspective. This use case has the potential to inspire future research to take into account how people perceive and interpret cities.

4.2 Limitations

There are some limitations of this work, at the proof of concept stage, which may require attention and offer opportunities for future work. First, in this study, a number of 49 participants are recruited to describe their feelings towards buildings in Amsterdam. Such a small sampling size may introduce bias when collecting perception information, considering an unbalanced diversity and limited perspectives. As illustrated in Figure 5, the building density does not deliver explicit differences among each category of perception features and scores, which may be related to the number of participants. To obtain more reliable results, we plan to increase the number of participants and add more cities to eliminate noisy information. Second, in this initial effort to integrate human perception of buildings, we use 5 indicators (i.e. original, pleasing, ordered, boring and complex) to perceive urban buildings from a sentimental perspective. However, the perception of buildings should be comprehensive and multi-dimensional. For example, individual experience with buildings can lead to varied understandings of building style, influenced by personal background and context-based interpretations. Therefore, in a future study, we will take into account building style (e.g. modern and historical) as another attribute to characterise a building of interest, serving a broader scope of research. Third, we use Amsterdam as a case study given the high quality of 3DBAG dataset in the Netherlands. Though this country-wide 3D dataset is invaluable to support our data integration and information storage, some challenges may be encountered when transferring our method to other cities. That is, various instances of 3D city models and urban digital twins may impact integrating human perception, in terms of their quality and availability (e.g. open access, data format, and modelling rules). The workflow we propose is generic, nevertheless reproducibility in other areas requires a consideration of local 3D data to avoid potential hindrances (e.g. invalid schema, geometric inconsistencies, or inaccurate semantic and thematic information) and use of approaches to refine the data before being used for this purpose (Wysocki et al., 2021).

5. Conclusion

To enhance a human-centric perspective in the landscape of 3D city models and urban digital twins, we introduce an idea and method to integrate the human perception of buildings as a new set of perceptual attributes in 3D building models. Using Amsterdam as an example, we gather data on how people perceive buildings by developing a computer vision model trained on survey responses. Meanwhile, the 3DBAG dataset in the Netherlands is adopted as the example for integration and extension. By incorporating such perception information in 3D buildings, we upgrade the current 3D city model in Amsterdam augmented with a human angle. A CityJSON Extension is then created to robustly store the perception data, and it is validated according to its standard and schema, ensuring the reliability for multiple uses, such as data exchange and maintenance. The validity of the work has been affirmed with the visualisation of the semantically enriched dataset in 3D. Further, we apply the generated data (i.e. perception attributes associated with each building in 3D) combined with 3D building metrics to unlock a new use case — clustering buildings based on morphology and perception in Amsterdam. As such, we characterise buildings from contrasting dimensions, offering a combined objective and subjective understanding of urban buildings.

Incorporating human perceptions unveils a social dimension in 3D building datasets, which have traditionally been confined to physical and objective attributes such as year of construction, floor area, building function, roof type, and energy consumption. Such a latent aspect, can advance 3D city models and urban digital twins and lead to a broader adoption in socio-economic topics, facilitating their usability to address urban challenges and serve a variety of disciplines. Further, the perception of buildings uncover a potential to accommodate a variety of human sensing data in 3D models in future, serving specific needs by following this generic workflow of standardising and validating customised new data. We deem that this work not only contributes to enriching subjective information in 3D buildings, but also demonstrates the practicality of shifting the techno-optimism to a more social focus in 3D city models and urban digital twins. For future work, we plan to explore new use cases of such integration and the incorporation of other types of human-centric information.

Acknowledgments

We thank the members of the NUS Urban Analytics Lab and Stelios Vitalis and Anna Labetski (TU Delft) for the discussions. This research is part of the projects (1) Multi-scale Digital Twins for the Urban Environment: From Heartbeats to Cities, which is supported by the Singapore Ministry of Education Academic Research Fund Tier 1; (2) Large-scale 3D Geospatial Data for Urban Analytics, which is supported by the National University of Singapore under the Start Up Grant R-295-000-171-133; (3) From Models to Pavements: Advancing Urban Digital Twins with a Multi-Dimensional Human-Centric Approach for a Smart Walkability Analysis, which is supported by the Humanities Social Sciences Seed Fund at the National University of Singapore (A-8001957-00-00). The 2nd author acknowledges the NUS Graduate Research Scholarship granted by the National University of Singapore (NUS).

References

Agugiaro, G., Benner, J., Cipriano, P., Nouvel, R., 2018. The Energy Application Domain Extension for CityGML: enhancing interoperability for urban energy simulations. *Open Geospatial Data, Software and Standards*, 3(1), 139.

Arribas-Bel, D., Fleischmann, M., 2022. Understanding (urban) spaces through form and function. *Habitat International*, 128, 102641.

Arslan, H. D., Yıldırım, K., 2023. Perceptual evaluation of stadium façades. *Alexandria Engineering Journal*, 66, 391–404.

Biljecki, F., Chow, Y. S., 2022. Global building morphology indicators. *Computers, Environment and Urban Systems*, 95, 101809.

Biljecki, F., Kumar, K., Nagel, C., 2018. CityGML Application Domain Extension (ADE): overview of developments. *Open Geospatial Data, Software and Standards*, 3(1), 13.

Caprari, G., Castelli, G., Montuori, M., Camardelli, M., Malvezzi, R., 2022. Digital twin for urban planning in the green deal era: A state of the art and future perspectives. *Sustainability*, 14(10), 6263.

Charitonidou, M., 2022. Urban scale digital twins in data-driven society: Challenging digital universalism in urban planning decision-making. *International Journal of Architectural Computing*, 20(2), 238–253.

Dang, L. M., Min, K., Wang, H., Piran, M. J., Lee, C. H., Moon, H., 2020. Sensor-based and vision-based human activity recognition: A comprehensive survey. *Pattern Recognition*, 108, 107561.

Dembski, F., Wössner, U., Letzgus, M., Ruddat, M., Yamu, C., 2020. Urban digital twins for smart cities and citizens: The case study of Herrenberg, Germany. *Sustainability*, 12(6), 2307.

Dukai, B., Peters, R., Vitalis, S., van Liempt, J., Stoter, J., 2021. Quality assessment of a nationwide data set containing automatically reconstructed 3D building models. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLVI-4/W4-2021, 17–24.

Dukai, B., Peters, R., Wu, T., Commandeur, T., Ledoux, H., Baving, T., Post, M., van Altena, V., van Hinsbergh, W., Stoter, J., 2020. Generating, storing, updating, and disseminating a country-wide 3D model. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLIV-4/W1-2020, 27–32.

Ferré-Bigorra, J., Casals, M., Gangolells, M., 2022. The adoption of urban digital twins. *Cities*, 131, 103905.

Galesic, M., Bruine de Bruin, W., Dalege, J., Feld, S. L., Kreuter, F., Olsson, H., Prelec, D., Stein, D. L., van Der Does, T., 2021. Human social sensing is an untapped resource for computational social science. *Nature*, 595(7866), 214–222.

Ghomeishi, M., 2021. Aesthetic preferences of laypersons and its relationship with the conceptual properties on building façade design. *Journal of Asian Architecture and Building Engineering*, 20(1), 12–28.

Gifford, R., Hine, D. W., Muller-Clemm, W., Reynolds JR, D. J., Shaw, K. T., 2000. Decoding modern architecture: A lens model approach for understanding the aesthetic differences of architects and laypersons. *Environment and behavior*, 32(2), 163–187.

Goodchild, M. F., 2007. Citizens as sensors: the world of volunteered geography. *GeoJournal*, 69(4), 211–221.

Gröger, G., Plümer, L., 2012. CityGML–Interoperable semantic 3D city models. *ISPRS Journal of Photogrammetry and Remote Sensing*, 71, 12–33.

Grütter, J. K., 2020. *Basics of Perception in Architecture*. Springer.

Hämäläinen, M., 2021. Urban development with dynamic digital twins in Helsinki city. *IET Smart Cities*, 3(4), 201–210.

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.

Heath, T., Smith, S. G., Lim, B., 2000. Tall buildings and the urban skyline: The effect of visual complexity on preferences. *Environment and behavior*, 32(4), 541–556.

Helbich, M., Jochem, A., Mücke, W., Höfle, B., 2013. Boosting the predictive accuracy of urban hedonic house price models through airborne laser scanning. *Computers, environment and urban systems*, 39, 81–92.

Huang, B., Zhou, Y., Li, Z., Song, Y., Cai, J., Tu, W., 2020. Evaluating and characterizing urban vibrancy using spatial big data: Shanghai as a case study. *Environment and Planning B: Urban Analytics and City Science*, 47(9), 1543–1559.

Jayathissa, P., Quintana, M., Abdelrahman, M., Miller, C., 2020. Humans-as-a-Sensor for Buildings—Intensive Longitudinal Indoor Comfort Models. *Buildings*, 10(10).

Kang, Y., Zhang, F., Gao, S., Lin, H., Liu, Y., 2020. A review of urban physical environment sensing using street view imagery in public health studies. *Annals of GIS*, 26(3), 261–275.

Katal, A., Mortezazadeh, M., Wang, L. L., Yu, H., 2022. Urban building energy and microclimate modeling–From 3D city generation to dynamic simulations. *Energy*, 251, 123817.

Ketzler, B., Naserentin, V., Latino, F., Zangelidis, C., Thuvander, L., Logg, A., 2020. Digital twins for cities: A state of the art review. *Built Environment*, 46(4), 547–573.

Kolbe, T. H., Gröger, G., Plümer, L., 2008. CityGML–3D city models and their potential for emergency response. *Geospatial information technology for emergency response*, 257.

Kotseruba, I., Gonzalez, O. J. A., Tsotsos, J. K., 2016. A review of 40 years of cognitive architecture research: Focus on perception, attention, learning and applications. *arXiv preprint arXiv:1610.08602*, 1–74.

Labetski, A., Kumar, K., Ledoux, H., Stoter, J., 2018. A Metadata ADE for CityGML. *Open Geospatial Data, Software and Standards*, 3(1), 42.

Labetski, A., Vitalis, S., Biljecki, F., Arroyo Ohori, K., Stoter, J., 2023. 3D building metrics for urban morphology. *International Journal of Geographical Information Science*, 37(1), 36–67.

Ledoux, H., Arroyo Ohori, K., Kumar, K., Dukai, B., Labetski, A., Vitalis, S., 2019. CityJSON: A compact and easy-to-use encoding of the CityGML data model. *Open Geospatial Data, Software and Standards*, 4(1), 1–12.

Lei, B., Janssen, P., Stoter, J., Biljecki, F., 2023a. Challenges of urban digital twins: A systematic review and a Delphi expert survey. *Automation in Construction*, 147, 104716.

Lei, B., Su, Y., Biljecki, F., 2023b. Humans as sensors in urban digital twins. *International 3D GeoInfo Conference*, Springer, 693–706.

Liang, X., Zhao, T., Biljecki, F., 2023. Revealing spatio-temporal evolution of urban visual environments with street view imagery. *Landscape and Urban Planning*, 237, 104802.

Liu, P., Zhao, T., Luo, J., Lei, B., Frei, M., Miller, C., Biljecki, F., 2023. Towards Human-centric Digital Twins: Leveraging Computer Vision and Graph Models to Predict Outdoor Comfort. *Sustainable Cities and Society*, 93, 104480.

Liu, X., Kang, C., Gong, L., Liu, Y., 2016. Incorporating spatial interaction patterns in classifying and understanding urban land use. *International Journal of Geographical Information Science*, 30(2), 334–350.

Liu, Y., Liu, X., Gao, S., Gong, L., Kang, C., Zhi, Y., Chi, G., Shi, L., 2015. Social sensing: A new approach to understanding our socioeconomic environments. *Annals of the Association of American Geographers*, 105(3), 512–530.

Lu, Z.-L., Sperling, G., 1995. The functional architecture of human visual motion perception. *Vision research*, 35(19), 2697–2722.

Ma, X., Ma, C., Wu, C., Xi, Y., Yang, R., Peng, N., Zhang, C., Ren, F., 2021. Measuring human perceptions of streetscapes to better inform urban renewal: A perspective of scene semantic parsing. *Cities*, 110, 103086.

Nochta, T., Wan, L., Schooling, J. M., Parlikad, A. K., 2021. A socio-technical perspective on urban analytics: The case of city-scale digital twins. *Journal of Urban Technology*, 28(1-2), 263–287.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V. et al., 2011. Scikit-learn: Machine learning in Python. *the Journal of machine Learning research*, 12, 2825–2830.

Peters, R., Dukai, B., Vitalis, S., Liempt, J. v., Stoter, J., 2022. Automated 3D Reconstruction of LoD2 and LoD1 Models for All 10 Million Buildings of the Netherlands. *Photogrammetric Engineering & Remote Sensing*, 88(3), 165–170.

QGIS.org, 2023. QGIS Geographic Information System. Open Source Geospatial Foundation Project.

Rahmadian, E., Feitosa, D., Virantina, Y., 2023. Digital twins, big data governance, and sustainable tourism. *Ethics and Information Technology*, 25(4), 1–22.

Redelmeier, D. A., Dickinson, V. M., 2011. Determining whether a patient is feeling better: pitfalls from the science of human perception. *Journal of general internal medicine*, 26, 900–906.

Stoter, J., Arroyo Ohori, K., Noardo, F., 2021. Digital twins: A comprehensive solution or hopeful vision? *GIM Int. Worldwide Mag. Geomat*, 2021.

Vitalis, S., Arroyo Ohori, K., Stoter, J., 2020a. CityJSON in QGIS: Development of an open-source plugin. *Transactions in GIS*, 24(5), 1147–1164.

Vitalis, S., Labetski, A., Boersma, F., Dahle, F., Li, X., Arroyo Ohori, K., Ledoux, H., Stoter, J., 2020b. Cityjson+ Web= Ninja. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 6, 167–173.

Wang, C., Wei, S., Du, S., Zhuang, D., Li, Y., Shi, X., Jin, X., Zhou, X., 2021. A systematic method to develop three dimensional geometry models of buildings for urban building energy modeling. *Sustainable Cities and Society*, 71, 102998.

Wang, J., Lu, C., Zhang, K., 2020. Textile-based strain sensor for human motion detection. *Energy & Environmental Materials*, 3(1), 80–100.

Wei, J., Yue, W., Li, M., Gao, J., 2022. Mapping human perception of urban landscape from street-view images: A deep-learning approach. *International Journal of Applied Earth Observation and Geoinformation*, 112, 102886.

Wysocki, O., Schwab, B., Hoegner, L., Kolbe, T. H., Stilla, U., 2021. Plastic surgery for 3D city models: A pipeline for automatic geometry refinement and semantic enrichment. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, V-4-2021, 17–24.

Yan, Y., Feng, C.-C., Huang, W., Fan, H., Wang, Y.-C., Zipf, A., 2020. Volunteered geographic information research in the first decade: a narrative review of selected journal articles in GIScience. *International Journal of Geographical Information Science*, 34(9), 1765–1791.

Ye, X., Du, J., Han, Y., Newman, G., Retchless, D., Zou, L., Ham, Y., Cai, Z., 2023. Developing human-centered urban digital twins for community infrastructure resilience: A research agenda. *Journal of Planning Literature*, 38(2), 187–199.

Zhang, F., Zhou, B., Liu, L., Liu, Y., Fung, H. H., Lin, H., Ratti, C., 2018. Measuring human perceptions of a large-scale urban region using machine learning. *Landscape and Urban Planning*, 180, 148–160.

Zhao, H.-x., Magoulès, F., 2012. A review on the prediction of building energy consumption. *Renewable and Sustainable Energy Reviews*, 16(6), 3586–3592.