



Article Data Clustering in Urban Computational Modeling by Integrated Geometry and Imagery Features for Probabilistic Navigation

Chenyi Cai ^{1,2}, Mohamed Zaghloul ² and Biao Li ^{1,*}

- ¹ School of Architecture, Southeast University, 2 Sipailou, Nanjing 210096, China
- ² Department of Architecture, Swiss Federal Institute of Technology Zurich (ETHZ), Stefano-Franscini-Platz 1, 8093 Zürich, Switzerland
- * Correspondence: jz.generator@gmail.com

Abstract: Cities are considered complex and open environments with multidimensional aspects including urban forms, urban imagery, and urban energy performance. So, a platform that supports the dialogue between the user and the machine is crucial in urban computational modeling (UCM). In this paper, we present a novel urban computational modeling framework, which integrates urban geometry and urban visual appearance aspects. The framework applies unsupervised machine learning, self-organizing map (SOM), and information retrieval techniques. We propose the instrument to help designers navigate among references from the built environment. The framework incorporates geometric and imagery aspects by encoding urban spatial and visual appearance characteristics with Isovist and semantic segmentation for integrated geometry and imagery features (IGIF). A ray SOM and a mask SOM are trained with the IGIF, using building footprints and street view images of Nanjing as a dataset. By interlinking the two SOMs, the program retrieves urban plots which have similar spatial traits or visual appearance, or both. The program provides urban designers with a navigatable explorer space with references from the built environment to inspire design ideas and learn from them. Our proposed framework helps architects and urban designers with both design inspiration and decision making by bringing human intelligence into UCM. Future research directions using and extending the framework are also discussed.

Keywords: urban computational modeling; data clustering; feature extraction; IGIF; self-organizing map; probabilistic navigation

1. Introduction

Computational modeling is necessary for scientific investigation because it impacts how we translate real-world phenomena to a digital representation. The foundation of computational modeling technologies changes with growing diversity. Various modeling methods increase the capacity for applications in the urban study and urban design domain. Urban computational modeling (UCM) has evolved from analytical to computational network methods, which is followed by methods using data streams [1]. Cities are considered complex and open environments with multidimensional aspects including urban forms, urban imagery, and urban energy performance [2]. So, a platform that supports the dialogue between the user and the machine is crucial in urban modeling investigation compared to executing analytical urban programs to calculate answers. In previous studies, researchers applied computational modeling methods to interactively generate 3D city models and used street view images (SVI) for user behavior studies and urban analytics [3]. However, more studies that include those two aspects are still required to investigate generic urban modeling techniques for encoding complex phenomena and multidimensional urban study.

Existing urban computational studies do not account for the integrated modeling of explicit knowledge and implicit knowledge, and there is a lack of efficient modeling



Citation: Cai, C.; Zaghloul, M.; Li, B. Data Clustering in Urban Computational Modeling by Integrated Geometry and Imagery Features for Probabilistic Navigation. *Appl. Sci.* **2022**, *12*, 12704. https:// doi.org/10.3390/app122412704

Academic Editors: Piotr A. Kowalski and Małgorzata Charytanowicz

Received: 20 October 2022 Accepted: 23 November 2022 Published: 11 December 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). approaches to include the geometrical and imagery dataset. A comprehensive urban computational modeling approach should allow a high efficiency, flexible interface with different input, process with heterogeneous data, and the provision of an instrument for navigation among the explorer spaces formed by existing projects. There is a lack of applications with data streams in urban design for

- Integrating geometric and imagery aspects of computational modeling for urban design;
- Extracting condensed but informative feature vectors of urban cases for efficient machine learning implementation;
- Integration and interrelation of multi-source data for case retrieval from the built environment;
- Facilitating personal preference and intuition in case retrieval to bring humans in the loop.

In this paper, we propose a method to navigate within cities using integrated geometric and imagery features (IGIF) and machine learning. We use the map data from the city of Nanjing, China as a dataset. We encode the geometrical-based space features and imagery features from a collection of observers' viewpoints, using Isovist and SVIs, respectively. The self-organizing map (SOM) and k-nearest neighbor (KNN) technology communicate the geometrical and imagery aspects. Our work demonstrates that IGIF helps probabilistic navigation among complex city phenomena. It supports urban design decision making and opens a novel perspective for a generic urban modeling method. The novel contributions of this paper are as follows:

- 1. A feature extraction method for geometric and imagery aspects of urban space using IGIF which is condensed but informative, supporting efficient clustering;
- 2. Implementing a program that allows navigation among the cases in the built environment, creating an explorer space using SOM which can help design decision making;
- 3. A framework designed to facilitate new modes of UCM with data streams, allowing integration and interplay among heterogeneous data.

2. Literature Review

2.1. Urban Computational Modeling Context

Urban computational modeling (UCM) approaches have been growing diversely since the 1960s, evolving from analytical to systematic modeling approaches [4]. In the last decade, spatio-temporal urban data streams (such as street view imagery, OpenStreetMap, and satellite imagery) are exponentially increasing, spawning many data-driven applications [5]. In the early modeling phase, urban models essentially simulate city functions which translate theory into a form that is testable and computable. Those modeling methods require the idealization of reality, so they only approximate reality to a certain extent. To address more aspects of urban problems, a multi-model idealization is often used to integrate heterogeneous simulation into a multidimensional computational model [6]. However, the increasing number of properties and their relationships make the model expand in an exponential manner [1]. In addition, it is hard to represent all the related conditions and implicit knowledge (preference, perception, etc.) by explicit properties. Consequently, the conventional approaches are not adequate to address the urban complexity and dimensionality.

With the growing diversity of modeling approaches with machine learning and accessibility of digital data, it is apparent that we should use data streams to embrace a shift in the modeling paradigm in urban studies. Data streams allow for analyzing and modeling the complicated interactions between spatial, temporal, emotional, and social aspects in a dynamic urban system [7]. Therefore, it offers opportunities to integrate explicit urban indicators with implicit urban perceptions, to advance the knowledge and understanding of urban dynamics. With this in mind, we aim to construct models to deal with complex urban problems by proposing a flexible framework that helps integrate dimensions of different urban aspects. Achieving the aforementioned goals requires using proper techniques from the multidisciplinary field of UCM. The subset of UCM techniques discussed here relates to rule-based and data-driven computation approaches. The following sections discuss some existing techniques useful for urban form and urban perception studies from the perspective of representation and data streams.

2.2. Data Streams and Feature Extraction in UCM

Models are regarded as representations of reality [4]. In other words, models using cellular automata, agent-based systems, and shape grammar are popular and powerful in many applications: for instance, land use, mobility, spatial planning, and so on. They require a set of properties to define the inquiry objects. An example of replicating a specific urban form is the construction of a multi-agent system computing properties and their interrelationships that represent the village morphology, roadwork pattern, and building structure form [8]. Integer programming is used for generating optimized urban and building layouts that meet the sunlight requirement, where the buildings are encoded as grid-based infill modules, by translating the spatial planning problem under the limitation of sunlight requirements to a packing problem in a fixed area [9]. Stream-line-based shape grammar for site splitting is used for generating street and parcel layouts where the spatial planning problem is represented by a subdivision problem [10,11].

With the rapidly growing accessibility of digital data, researchers in various fields widely use urban remote sensing and geospatial data processed by machine learning techniques. Cartographic data and street view imagery are ingrained as important urban data sources [3]. Analytical cartography is essential for urban planning and morphology to illustrate spatial patterns. The urban form visualizations are comprehensive information artifacts that reflect urban complexity [12]. Therefore, researchers have made efforts to build collaborative spatial information systems for exploring urban fabric patterns and spatial order to inform urban planning and urban design [13,14]. Apart from geometry aspects, visual appearance is closely related to urban perception. It is relatively implicit knowledge but important for understanding the perception of spaces, such as greenery [15], understanding symmetries of urban blocks and so on [16]. SVI is used in a significant number of urban perception studies because it enables characterizing urban space from a human perspective.

One of the main challenges in UCM with data streams lies in integrating heterogeneous data from multiple sources [17]. Therefore, in a technical view, the challenge is to find a generic feature extraction mechanism for data in different formats (geometric data, pictorial data, etc.). Feature engineering and clustering approaches have the potential for going over urban dimensionality to multi-source integration.

Feature extraction is an essential step for extracting samples' features that are useful for clustering. Multiple quantitative approaches are implemented for extracting urban features. Researchers use pre-proposed indicators (e.g., semantic indicators, geometrical indicators) to index urban spatial characteristics for the follow-up studies [18]. The indicators are decided depending on the specific task and should be verified by experiments. So, it takes effort to decide on the indicators. Classic statistical indicators for urban morphology include the "plot size", the plot "edge length", "Size/Edge2", "GSI" and "FSI". Using compressed feature vectors from neural networks provides a flexible way to comprehensively represent urban fabrics that excludes the artificial selection of morphological indicators [19]. The visual field, isovist, refers to visibility that represents the size of a space that people visually perceive. Isovist is popular in analyzing visibility in space, taking human-centric perception to describe spatial characteristics [20,21].

In a convolution neural network (CNN), the feature mapping of the images extracts the input's underlying features by convolution kernels, promoting deep learning algorithms in recent decades [22]. Neural networks are applied to solve problems such as the prediction of energy performance [23], the pattern recognition of 2D images [24], as well as typological form-finding on 3D models [25]. Therefore, apart from the compressed geometry, feature

extraction by neural networks provides the possibilities for extracting urban features under a generic mechanism.

2.3. Geometric and Imagery Urban Data Clustering Implementations

Clustering in urban studies not only brings efficiency but also the potential of connecting heterogeneous samples [26]. The clustering of geometry data such as city form and urban parcel form is carried out to study the relationship between urban form and air pollution [27], heat land [28], urban vibrancy [29], and so on. Researchers have been using SVI to measure urban perceptual attributes such as vibrancy, comfort, and attitude toward the living environment [30]. Quantification studies on SVI for assessing the characteristics of the neighborhood's visual appearance are considered important for promoting people's physical activities and improving residents' sense of comfort [31]. Semantic segmentation techniques are applied for extracting urban feature composition and understanding streetlevel morphology [32]. Apart from the studies and investigation of urban knowledge, the data clustering technique provides urban designers an instrument to create an explorer space to investigate. The self-organizing map (SOM) is a well-known method that has very rich literature with a diverse set of applications [33]. SOM can perform very well in data clustering and classification; thus, SOM is a generic and flexible computational technique that can be used for different purposes depending on the perspective. Personal preference characterization in urban spaces is implemented to predict likable places for a specific observer using geotagged satellite and perspective images from diverse urban environments [34]. Clustering with the self-organizing map links building skin performances with building geometries in a highly effective way, subsequently helping architects in design-performance interaction at the conceptual design level [35]. Design space for responsive and fast conceptual design is demonstrated in the structural field [36,37].

2.4. Key Issues in the Literature

The gaps we see in the existing literature are listed as follows:

- 1. There is a lack of study on encoding urban plots' geometry as feature vectors that carry the spatial characteristics information for further urban form-related studies.
- 2. There is also a lack of frameworks that integrate data-clustering techniques to provide a responsive conceptual design process according to both the master plan and human perception aspects that are heterogeneous data sources.
- 3. These gaps result in the lack of implementations in the related domain that involve multiple urban aspects for holistic design decision making.

The following sections propose several novel contributions to achieve the goal of developing a framework of a responsive platform for probabilistic navigation according to users' queries by urban geometry and perception. These contributions include a methodology for encoding plot geometric features by isovists from the points along the plot boundary, data retrieval based on the self-organizing maps that are trained with these features, and a framework that ties these components together for interplayed navigation.

3. Material and Methods

3.1. Dataset Construction

We chose Nanjing for our case study. Map data including buildings, road networks, and satellite images are available online from mapping services such as OpenStreetMap. However, Google street view image is not available in mainland China due to business restrictions [38]. Therefore, we collected geometric data in shapefiles, street view images from Baidu Street View (BSV), and satellite images from Baidu satellite view. The geometric data include building footprints and area of interest (AOI), which is the boundary of each plot (Figure 1). We take each plot and the buildings in the plot as a sample according to the AOI. The attributes of a building include height and area. Figure 2 shows 200 out of 5803 plots, where the color transparency of the buildings refers to the value of height.



Figure 1. Visualization of map data including the area of interest, building footprints, and road networks.



Figure 2. Examples of plot slices including AOI and buildings.

Static street view images and satellite images can be requested from BSV and Baidu satellites via its public API. One can retrieve an image with the HTTP request where the FOV, heading, pitch, and panorama ID, which are expressed in degrees, need to be supplied. We set the FOV value as 90. A developer key is required. An image can be retrieved by simply pasting a request URL into a web browser or by sending API requests in a batch using any programming language. The heading value is related to the angle of looking at the panorama from the observing point. We set the observing points along the AOI boundary. We retrieve 4 SVIs from 4 heading values which are determined by the orientation of the boundary line where the point is located at. We set 2 angles to look at the corresponding plot and the other 2 to look at the plot on the other side of the street because we want to see more of the buildings along the boundary. We take the AOI boundary box location as a reference for requesting satellite image tiles, which is followed by merging



the tiles into one image. Figure 3 shows examples of satellite images, master plan images, and SVIs.

Location:118.83,32.3415

Figure 3. Examples of the samples in the dataset, include satellite images, master plans, and SVIs.

3.2. Feature Extraction

In this study, our interests lie in the spatial and visual experience along the plot, and SVIs are not accessible in private plots. Therefore, we focused on the encoding of spatial characteristics of the plots along the street to interrelate the geometric features with the visual features from the SVIs. Taking into consideration the spatial feature at one observing point in the street and the spatial sequences of walking along the street, we distribute observing points along the AOI boundary at each corner and every 60 m at the edge. Visibility is an important factor to describe a space because it shows people's perception of the space when standing at a specific point. Therefore, the isovist, which is usually used for evaluating visibility, can be applied for encoding the spatial features, the numbers of which can be the input of clustering. For each observing point, 32 rays are distributed and 32 ray distances (RD) from the point to the first obstacles are stored as the geometric feature vectors. Building footprints within a 100 m radius circle around the target plot are included for the measurements. Figure 4 shows an example of observers, rays and spatial radial chart of a plot.



Figure 4. Allocation of the observers and the corresponding rays on a plot.

The elements in an SVI are indicators of visual appearance and perception, so the semantic segmentation technique is applied to the SVIs. We implemented the semantic segmentation for the SVIs by a pre-trained model "Ademxapp Model A1 Trained on Cityscapes Data" on the Mathematica platform. Nineteen elements including road, sidewalk, building,

wall, fence, pole, traffic light, traffic sign, vegetation, terrain, sky, person, rider, car, truck, bus, train, motorcycle, and bicycle are labeled in each image. The images are resized to 300×300 pixels. Each pixel is labeled as an index for the element after the pixel-based training. Based on the labeling, we can visualize the mask of segments for each SVI and evaluate the percentage of each element in the image. The elements' percentage values after normalization are used as imagery feature vectors. In another word, the feature vectors for SVIs are reduced to only nineteen dimensions. Figure 5 shows scatterplots for the percentage of each element in the images. Figure 6 shows the colored visualization of each element in the images.



Figure 5. The scatterplots for visualizing the percentage of the segments from all the samples.



Figure 6. Semantic segmentation for the SVIs, coloring according to the segments.

3.3. Data Clustering Using the Self-Organizing Map

SOM is a generic methodology, which has been applied in many classical modeling tasks such as clustering, visualization of high-dimensional space, classification, prediction, and many other tasks. Compared with the classic clustering algorithm such as K-means, the nature of SOM makes it less affected by noises, more flexible in case retrieval and has better continuous visualization as a spectrum in a 2D space. K-mans is a linear method, and

a K value should be defined as the number of clusters that are discrete. SOM is optimized based on a 2D network, where each node has a weight value and a coordinate to keep the topology of the network while optimizing. The cases can be retrieved according to one closest node, nodes within a certain radius, or even the furthest nodes because the nodes are topologically connected. Therefore, flexible case retrieval for the search engine can be applied according to various criteria set by users.

SOM is a general-purpose nonlinear data transformation method that offers solutions for data clustering and 2D visualization as it creates continuous visual patterns on top of high-dimensional data [1]. The node weight value is a high-dimensional number whose dimension is the same as the input feature vectors and is updated by reducing the distance between the node and input data. For calculating the distance between the node and the sample, we use Euclidean distance because it is the most commonly used method, which is the most common metric to calculate the distance between node weights and sample vector. After several iterations, each node is associated with a set of input objects that are closer to this node than the other nodes. This method serves as the clustering process. During the optimization of SOM, the network would be stretched to cover the high-dimensional data space as much as possible. Therefore, the coordinates of the nodes help visualize the SOM in a 2D plane, serving as a dimension-reduction technique that transforms objects with high-dimensional features to a low-dimensional plane while the neighboring relationship remains.

Two simultaneous processes explain the SOM algorithm. The training data set can be considered as $X = x_1, ..., x_N$, as a set of N points in an n-dimensional space. An SOM can be considered as a grid with K nodes, with a set of indices y_j , each attached with a high-dimensional weight vector, w_j . During the training process, an index y_j can be assigned for each data xi. The index is also called the best-matching unit (BMU). $s(w_j, x)$ is a similarity function which is calculated by the inverse of the distance between the input sample feature vectors x and the weight vectors of the SOM node j, w_j .

$$BMU(x) = \arg\max_{i}(f(y_{i}|x))$$
(1)

$$f(y_j|x) = \frac{\exp s(w_j, x)}{\sum_{i=1}^k \exp s(w_j, x)}$$
(2)

The weight vectors of the indices, w_j , start with an initial value and will be adapted to become similar to their assigned input data and further to be similar in weight vectors to their topological neighborhoods in the grid. The process is called competition and adaption mechanisms [39]. The adaption process can be stopped after several iterations if the global error is less than a threshold. In this experiment, we use the 51-dimensional IGIF which joins the geometric feature vectors and imagery feature vectors. We trained two SOMs both with a 30 × 30 grid. A ray SOM and a mask SOM are trained, respectively, with 32-dimensional geometric features and 19-dimensional imagery features.

We take quantization error and topological error into consideration for model evaluation. The quantization error is the average distance between each data vector and its BMU showing the global mapping quality of the SOM. The topographic error is the proportion of all data vectors for which the first and second BMUs are not adjacent units, showing the distortion of topology when the SOM is overfitted. Therefore, we trained 40 ray SOM models and 40 mask SOM models with the size of 30×30 by 400 iterations. Figure 7 shows the plot of the quantization error and topological error of the ray SOM and mask SOM models, respectively. A rule of thumb is a model that has a topographic error very near zero and has the lowest quantization error because it is important to hold the topographic error very low in order to make the components from the network smooth. Therefore, we choose model 29 from the ray SOMs and model 16 from mask SOMs for further case studies.



Figure 7. The scatter plots showing quantization error and topographic error of ray SOM models (**left**) and mask SOM models (**right**).

The final output of an SOM can be visualized in different ways. One is to represent the pattern among different dimensions, rendering each dimension of the weight vectors in parallel in a new coordinate system. Another common way is to visualize one of the input data that is assigned in the node, since a set of input data is assigned to each node. The assigned data's feature vectors are similar to the weight vectors of the node. We visualize the ray SOM and mask SOM in a 2D plane where each node is shown by a sample that is closest to the node (Figure 8).



Figure 8. Results of ray SOM (left) and mask SOM (right) with a demonstration sample for each BMU.

4. Case Studies of the Framework

In Figure 8, clusters of narrow space, open space, semi-open space, cross space, T-shape space, and so on can be intuitively seen in the visualization of the ray SOM. The mask SOM shows an obvious gradient of SVI patterns. For instance, from right up to left bottom, it shows the gradient from mainly building-covered images to wide open space mainly with the sky. The right bottom part shows SVIs with a large proportion of vegetation. With the help of the SOMs, one can search for similar urban plots by spatial radar or SVI input. We show some case studies of queries for the SOMs.

We test the SOMs in two different ways, including querying a single SOM for similar patterns, and an interplayed query of two SOMs for cases similar in both features. As

shown in Figure 9, the user interface is capable of two kinds of input, including location and SVI, regarding the geometry feature and imagery feature, respectively. The retrieval loop can go from A to B to A, which means retrieving locations with similar spatial traits by inputing a location or the ray distances. It can also go from A to B to C, which means retrieving street view images of the locations with similar spatial traits. Similarly, if the retrieval loop goes from C to D to C, it means finding similar street view images. If it goes from C to D to C, it means finding locations that are similar in visual elements. Moreover, the framework supports the full loop retrieval from A to B to C to D to A. Users can find locations that are similar in both spatial traits and visual appearance by inputing geometric features and subsequently selecting locations that have close imagery features from the results or vice versa. We show some case retrievals in the following paragraphs.



Figure 9. The retrieval loop with 4 ends including 2 user interface ends and 2 database ends.

Firstly, we query for locations in Nanjing that have similar spatial traits as the input. Three examples are set as input, including narrow space, half-open space, quarter-open space, and so on. Eight retrieved locations of each are shown in Figure 10. From the results, the plots that have the spatial characteristics are successfully retrieved by RD, and the locations of the observer are displayed.

Apart from retrieval according to one SOM, a recursive search is essential, especially for design inspiration. Therefore, the interplay between two SOMs enables users to search among the possibilities by sharpening the target in the input–feedback–input loop. We show the case studies of retrieving across the SOMs. Firstly, we search for locations and street view images with a viewpoint similar with the input radar or location. Then, we take one point with the SVIs that we are interested in according to the retrieved locations. Furthermore, we query again for the cases that have similar visual appearance from the previous output. Therefore, we query for case packages with master plans, street view images and satellite images according to the spatial ray and user's feedback. Figure 11 shows two case studies in this query way.



Figure 10. Locations in similar spatial traits are retrieved according to geometric features.



Figure 11. Query for locations and SVIs that have similar spatial characteristics with the input marked with a red box (step 1) and query for case packages according to the selected SVIs (marked with a red box) based on the step 1 output (step 2).

5. Discussion and Further Research

We have presented a framework for the collection, processing, storage, clustering, and interconnecting of spatial and visual characteristics to enable probabilistic navigation in the space to support decision making using urban cases in Nanjing. This section further justifies the system, starting with a discussion of the objectives mentioned in the introduction. This is followed by discussing the framework as a basis for future research in navigation among the explorer space for urban design.

The IGIF provides an efficient and comprehensive encoding and feature extraction approach for urban plots. Compared with pixel-based feature extraction where the samples are

represented by 2048-dimensional features [40], IGIF has more condensed 51-dimensional features. Therefore, the following data clustering process has higher efficiency. Although the IGIF has fewer feature dimensions, it is still informative. It takes into consideration master planning and urban perception to bring implicit knowledge into urban computational modeling. It performs well in finding observing points with similar spatial characteristics by geometric vectors extracted by the obstacle-detecting method and in finding points with a similar visual appearance by imagery vectors that are composed of semantic segments percentages.

The data clustering by SOM provides spectra of the urban plots corresponding to the different aspects, which further provides a platform for urban designers to navigate based on probabilistic. An SOM is spreading the network in the data space to reach the data as much as possible, forming a spectrum for case retrieval. Data with similar spatial or visual characteristics are put in the same or close clusters. Then, urban designers can perform query and feedback loops instantly because of the efficiency to navigate among the cases in the built environment and learn from the existing cases. In addition, the interplay between the SOMs connects data from different sources. Therefore, in our framework, SOM does not only cluster but also generates, predicts, and visualizes from one dimensionality to another according to the probabilities.

From the urban design perspective, the framework provides flexible computational modeling by encoding implicit knowledge with the generic mechanism, interconnecting heterogeneous data, and creating an explorer space to bring together machine and human intelligence. It supports a mindset shift from generating to retrieving. Moreover, it proposes that the machine should not only be used for optimization but rather as an instrument that leverages the active role of the designer.

The framework opens possibilities for UCM in the data stream context. In fact, a methodology similar to that used on developing CFD emulation would be a valuable extension [35]. In other words, the framework is flexible to be further extended to implement urban performance, urban perception, urban fabric aspects, and so on. A demonstration of the conceptual design process based on a specific task and personal preference using the program would be also beneficial. There are also possible applications in practical studies. For instance, the core program could serve as a back end for a web-based city search engine for place finding such as real estate or urban renovation, with flexible selection criteria depending on specific tasks. An immersive experience for place selection, on-site visiting, and decision making can also be applied by integrating extended reality techniques.

6. Conclusions

In this paper, we propose a framework of the UCM approach with data streams, which integrates implicit knowledge of urban form and appearance using data-clustering techniques, SOM. We use a dataset from the Baidu map including building footprints and SVIs of Nanjing. Additionally, we propose a novel feature extraction method for encoding urban spatial and visual appearance characteristics using Isovist and semantic segmentation for the condensed feature vectors IGIF. It is shown to be more effective and flexible than pixel-based feature extraction.

We trained a ray SOM and a mask SOM with the IGIF. By interlinking the two SOMs, the program retrieves urban plots which have similar space and visual appearance, providing urban designers with references to the built environment. The performance of image retrieval is shown to be robust because of an advanced deep convolutionary neural network. The IGIF also allows the search engine to retrieve plots similar in both aspects.

We recommend applying artificial intelligence to inspiring urban design ideas by navigating the explorer space formed by existing projects and learning from them. Our proposed framework helps architects and urban designers with both design inspiration and decision making by bringing human intelligence into UCM. We also discuss the potential of this framework in multi-source case retrieval, urban performance, and data science. Similar future frameworks provide a new and flexible approach to obtaining targeted analytics and discussion for stakeholders. Future research will look at an extension of the dataset and urban visual segments knowledge discovery. The integration datasets of building and urban performance are to be included. The framework can also be further applied with mixed reality techniques for augmented design and decision making.

This research tries to contribute to the common challenge that arises when using urban computational modeling approaches, which is, on one hand, the paradox between diversity and completeness of the design decisions, and on the other hand, the difficulty for humans to interact with a large amount of data. This research could serve as a basis for further research that enhances the interaction and integration of machine and human intelligence.

Author Contributions: Conceptualization, C.C.; methodology, C.C. and M.Z.; software, C.C.; validation, C.C., M.Z. and B.L.; formal analysis, C.C.; investigation, C.C.; resources, C.C.; writing—original draft preparation, C.C.; writing—review and editing, C.C., M.Z. and B.L.; funding acquisition, B.L. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by the National Natural Science Foundation of China (NSFC) project (No.51978139) and 2021 Jiangsu Province Special Fund for Sustainability in Building Development.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data was obtained from Baidu map and are available with Baidu API.

Conflicts of Interest: The authors declare no conflict of interest.

References

- 1. Moosavi, V. Pre-Specific Modeling: Computational Machines in Coexistence with Urban Data Streams. Ph.D. Thesis, ETH Zurich, Zurich, Switzerland, 2015.
- 2. Kropf, K. Aspects of urban form. Urban Morphol. 2009, 13, 105. [CrossRef]
- 3. Biljecki, F.; Ito, K. Street view imagery in urban analytics and GIS: A review. Landsc. Urban Plan. 2021, 215, 104217. [CrossRef]
- 4. Batty, M. Urban Modelling; Cambridge University Press: Cambridge, UK, 1976.
- Yap, W.; Janssen, P.; Biljecki, F. Free and open source urbanism: Software for urban planning practice. *Comput. Environ. Urban* Syst. 2022, 96, 101825. [CrossRef]
- Chadzynski, A.; Li, S.; Grisiute, A.; Farazi, F.; Lindberg, C.; Mosbach, S.; Herthogs, P.; Kraft, M. Semantic 3D City Agents—An intelligent automation for dynamic geospatial knowledge graphs. *Energy AI* 2022, *8*, 100137. [CrossRef]
- Cheng, J.; Gould, N.; Han, L.; Jin, C. Big Data for urban studies: Opportunities and challenges: A comparative perspective. In Proceedings of the 2016 Intl IEEE Conferences on Ubiquitous Intelligence & Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People, and Smart World Congress (UIC/ATC/ScalCom/CBDCom/IoP/SmartWorld), Toulouse, France, 18–21 July 2016; pp. 1229–1234.
- Li, B.; Guo, Z.F.; Ji, Y.Z. Modeling and Realizing Generative Design—A Case Study of the Assignment of Ji Village. J. Archit. 2015, 560, 94–98.
- 9. Hua, H.; Hovestadt, L.; Tang, P.; Li, B. Integer programming for urban design. Eur. J. Oper. Res. 2019, 274, 1125–1137. [CrossRef]
- Chen, G.; Esch, G.; Wonka, P.; Müller, P.; Zhang, E. Interactive procedural street modeling. In Proceedings of the ACM SIGGRAPH 2008 Papers, Los Angeles, CA, USA, 11–15 August 2008; pp. 1–10.
- 11. Yang, Y.L.; Wang, J.; Vouga, E.; Wonka, P. Urban pattern: Layout design by hierarchical domain splitting. *ACM Trans. Graph.* (*TOG*) **2013**, *32*, 1–12. [CrossRef]
- 12. Boeing, G.D. Methods and Measures for Analyzing Complex Street Networks and Urban Form; University of California: Berkeley, CA, USA, 2017.
- 13. Boeing, G. Spatial information and the legibility of urban form: Big data in urban morphology. *Int. J. Inf. Manag.* **2021**, *56*, 102013. [CrossRef]
- 14. Boeing, G. Exploring Urban Form through OpenStreetMap Data: A Visual Introduction. In *Urban Experience and Design*; Routledge: London, UK, 2020; pp. 167–184.
- 15. Lu, Y. The association of urban greenness and walking behavior: Using google street view and deep learning techniques to estimate residents' exposure to urban greenness. *Int. J. Environ. Res. Public Health* **2018**, *15*, 1576. [CrossRef]
- 16. Samiei, S.; Rasti, P.; Daniel, H.; Belin, E.; Richard, P.; Rousseau, D. Toward a computer vision perspective on the visual impact of vegetation in symmetries of urban environments. *Symmetry* **2018**, *10*, 666. [CrossRef]
- Yang, F.; Hua, Y.; Li, X.; Yang, Z.; Yu, X.; Fei, T. A survey on multisource heterogeneous urban sensor access and data management technologies. *Meas. Sens.* 2022, 19, 100061. [CrossRef]

- Chen, H.C.; Han, Q.; de Vries, B. Urban morphology indicator analyzes for urban energy modeling. *Sustain. Cities Soc.* 2020, 52, 101863. [CrossRef]
- Van Nes, A.; Berghauser Pont, M.; Mashhoodi, B. Combination of Space syntax with spacematrix and the mixed use index: The Rotterdam South test case. In Proceedings of the 8th International Space Syntax Symposium, Santiago, Chile, 3–6 January 2012; PUC: Santiago, Chili, 2012.
- 20. Benedikt, M.L. To take hold of space: Isovists and isovist fields. Environ. Plan. B Plan. Des. 1979, 6, 47-65. [CrossRef]
- Giseop, K.; Ayoung, K.; Youngchul, K. A new 3D space syntax metric based on 3D isovist capture in urban space using remote sensing technology. *Comput. Environ. Urban Syst.* 2019, 74, 74–87.
- 22. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. *Commun. ACM* 2017, 60, 84–90. [CrossRef]
- 23. Mohammadiziazi, R.; Bilec, M.M. Application of machine learning for predicting building energy use at different temporal and spatial resolution under climate change in USA. *Buildings* **2020**, *10*, 139. [CrossRef]
- 24. Cai, C.; Li, B. Training deep convolution network with synthetic data for architectural morphological prototype classification. *Front. Archit. Res.* **2021**, *10*, 304–316. [CrossRef]
- 25. De Miguel, J. Deep Form Finding-Using Variational Autoencoders for Deep Form Finding of Structural Typologies; Blucher: São Paulo, Brazil, 2019.
- Nogueira, K.; Penatti, O.A.; Dos Santos, J.A. Towards better exploiting convolutional neural networks for remote sensing scene classification. *Pattern Recognit.* 2017, 61, 539–556. [CrossRef]
- Liu, M.; Wei, D.; Chen, H. Consistency of the relationship between air pollution and the urban form: Evidence from the COVID-19 natural experiment. *Sustain. Cities Soc.* 2022, *83*, 103972. [CrossRef]
- 28. Liu, H.; Huang, B.; Zhan, Q.; Gao, S.; Li, R.; Fan, Z. The influence of urban form on surface urban heat island and its planning implications: Evidence from 1288 urban clusters in China. *Sustain. Cities Soc.* **2021**, *71*, 102987. [CrossRef]
- Li, F.; Yao, N.; Liu, D.; Liu, W.; Sun, Y.; Cheng, W.; Li, X.; Wang, X.; Zhao, Y. Explore the recreational service of large urban parks and its influential factors in city clusters—Experiments from 11 cities in the Beijing-Tianjin-Hebei region. *J. Clean. Prod.* 2021, 314, 128261. [CrossRef]
- 30. Wei, J.; Yue, W.; Li, M.; Gao, J. Mapping human perception of urban landscape from street-view images: A deep-learning approach. *Int. J. Appl. Earth Obs. Geoinf.* 2022, 112, 102886. [CrossRef]
- Wu, C.; Peng, N.; Ma, X.; Li, S.; Rao, J. Assessing multiscale visual appearance characteristics of neighbourhoods using geographically weighted principal component analysis in Shenzhen, China. *Comput. Environ. Urban Syst.* 2020, 84, 101547. [CrossRef]
- Middel, A.; Lukasczyk, J.; Zakrzewski, S.; Arnold, M.; Maciejewski, R. Urban form and composition of street canyons: A human-centric big data and deep learning approach. *Landsc. Urban Plan.* 2019, 183, 122–132. [CrossRef]
- 33. Kohonen, T. *Self-Organizing Maps*; Springer Science & Business Media: Berlin, Germany, 2012; Volume 30.
- Alvarez Marin, D. Atlas of Indexical Cities: Articulating Personal City Models on Generic Infrastructural Ground. Ph.D. Thesis, ETH Zurich, Zurich, Switzerland, 2020.
- Zaghloul, M. Machine-Learning aided Architectural Design-Synthesize Fast CFD by Machine-Learning. Ph.D. Thesis, ETH Zurich, Zurich, Switzerland, 2017.
- Mueller, C.T. Computational Exploration of the Structural Design Space. Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 2014.
- 37. Saldana Ochoa, K.; Ohlbrock, P.O.; D'Acunto, P.; Moosavi, V. Beyond typologies, beyond optimization: Exploring novel structural forms at the interface of human and machine intelligence. *Int. J. Archit. Comput.* **2021**, *19*, 466–490. [CrossRef]
- 38. Liang, J.; Gong, J.; Sun, J.; Zhou, J.; Li, W.; Li, Y.; Liu, J.; Shen, S. Automatic sky view factor estimation from street view photographs—A big data approach. *Remote Sens.* **2017**, *9*, 411. [CrossRef]
- 39. Kohonen, T. The basic SOM. In Self-Organizing Maps; Springer: Berlin, Germany, 1995; pp. 77–130.
- 40. Cai, C.; Guo, Z.; Zhang, B.; Wang, X.; Li, B.; Tang, P. Urban morphological feature extraction and multi-dimensional similarity analysis based on deep learning approaches. *Sustainability* **2021**, *13*, 6859. [CrossRef]