Special issue article

Urban Analytics and City Science

EPB: Urban Analytics and City Science
2019, Vol. 46(8) 1439–1457

DOI: 10.1177/2399808319828734
journals.sagepub.com/home/epb
SAGE

The visual quality of streets: A human-centred continuous measurement based on machine learning algorithms and street view images

Yu Ye
Tongji University, China

Wei Zeng
Shenzhen Institutes of Advanced Technology, China

Qiaomu Shen
Hong Kong University of Science and Technology, Hong Kong

Xiaohu Zhang
Singapore-MIT Alliance for Research and Technology Centre (SMART), Singapore

Yi Lu
City University of Hong Kong, Hong Kong

Abstract

This study proposes a workable approach for quantitatively measuring the perceptual-based visual quality of streets, which has often relied on subjective impressions or feelings. With the help of recently emerged street view images and machine learning algorithms, an evaluation model has been trained to assess the perceived visual quality with accuracy similar to that of experienced urban designers, to provide full coverage and detailed results for a citywide area. The town centre of Shanghai was selected for the site. Around 140,000 screenshots from Baidu Street View were processed and a machine learning algorithm, SegNet, was applied to intelligently extract the pixels representing key elements affecting the visual quality of streets, including the building frontage, greenery, sky view, pedestrian space, motorisation, and diversity. A Java-based program was then produced to automatically collect the preferences of experienced urban

Corresponding author:
Wei Zeng, Shenzhen Institutes of Advanced Technology, Shenzhen, China.
Email: wei.zeng@siat.ac.cn

designers on representative sample images. Another machine learning algorithm, i.e. an artificial neural network, was used to train an evaluation model to achieve a citywide, high-resolution evaluation of the visual quality of the streets. Further validation through different approaches shows this evaluation model obtains a satisfactory accuracy. The results from the artificial neural network also help to explore the high or low effects of various key elements on visual quality. In short, this study contributes to the development of human-centred planning and design by providing continuous measurements of an 'unmeasurable' quality across large-scale areas. Meanwhile, insights on the perceptual-based visual quality and detailed mapping of various key elements in streets can assist in more efficient street renewal by providing accurate design guidance.

Keywords

Street view images, machine learning, urban design, street, visual quality

Introduction

Streets are an essential component of urban public space closely related to urban residents (Madanipour, 1996). High-quality street space not only helps to generate urban vitality but it also contributes to positive social interactions, outdoor activities, and public health (Handy et al., 2002; Ye et al., 2018). In short, the perceptual-based spatial quality of urban streets has been regarded as an important public good. In this context, many design initiatives and urban renewal policies on streets have been proposed internationally, including *Urban Street Design Guide* (National Association of City Transportation Officials, 2013), and *Shanghai Street Design Guidelines* (Shanghai Planning Bureau, 2016), etc. As shown in these guides, the focus on streets is gradually evolving from one that is 'transportation-centred' towards one that is 'human-centred'. Many evaluation platforms, such as *Walk Score* and *Stated of Space Index*, were proposed to assist in accurate quality assessment and design interventions. The simultaneous presentation of design initiatives, urban policies, and evaluation platforms indicates an increasing global concern for street quality.

Nevertheless, the pursuit of high-quality street space is not a new issue. Many urbanists, from Jacobs (1961) to Montgomery (1998), have given classical discussions on various design elements affecting this quality based on their subjective experiences. With few exceptions, these classical urban design studies have not provided an objective exploration of design elements and the related perceptual attribute of visual quality. In recent years, quantitative-based approaches, such as the stated-preference survey and behaviour mapping, have been gradually introduced into this field. Some further understanding about the relationship between visual quality, an intangible perception of space, and the corresponding design elements has been achieved (Ewing and Handy, 2009; Gehl et al., 2006). Nevertheless, these studies are often conducted on manually collected, small-scale data sets using time-consuming approaches that cannot efficiently support further studies.

Current movements that focus on taking back the streets require us to address two issues that have been hard to achieve. The first is quantitatively measuring design elements in streetscapes and their effects on street quality. The second is to achieve a rapid evaluation of the perceptual-based visual quality with high spatial resolution to provide accurate guidance for urban planners, designers, and policy makers. Without insights on these two issues, urban design research cannot efficiently address the rapidly emerging demands for high-quality street space.

Recently, the availability of new data and new analytical techniques has brought new research potential to the handling of these two issues (see online Appendix I). Large data sets of panoramic images taken from a human-centred viewpoint, such as Google Street View and Baidu Maps, provide new data on streetscapes. Emerging analytical techniques have opened up new opportunities for mapping street design elements at high spatial resolution, which can then be transformed into a measure of urban design quality in streetscapes. For instance, automatic processing of street view images (SVIs) has been used in quantifying street greenery across neighbourhood scales (Li et al., 2015). The rise of machine learning algorithms has introduced new tools, including a series of deep convolutional networks for precisely extracting spatial features from SVIs. Meanwhile, some other machine learning algorithms, e.g. artificial neural networks (ANNs), can help to address the complex relationships amongst the various design elements and the visual quality to achieve an efficient evaluation of this intangible value. In short, new research potential is emerging to aid in measuring the 'unmeasurable'.

The conceptual framework

Visual quality of streets

The term 'quality' is a broad concept, and it is necessary to clarify that 'quality' as used herein is a perceptual value representing the visual experience affected by the streetscape. In other words, it is an intangible value highly connected to the surrounding physical features. It does not attempt to represent the full meaning of quality in urban design, which includes not only high-quality physical streetscapes, but also activities, human emotions, and a sense of community.

This concept can be traced back to Theil's (1961) 'space score' that reflects the extent to which design elements perceived by people in motion would lead to a change in their behaviour. It was also partly related to the 'sensuous qualities' or 'sense' of place by Kevin Lynch (Banerjee and Southworth, 1991). It mainly focuses on the visual impression amongst sense of place indictors discussed by Chen (2014) and Gokce and Chen (2018). When this concept is applied to streets, on the one hand, it reflects users' willingness to stay or use and enjoy the space in streets. On the other hand, it can also be measured as a series of key elements discussed in urban design theories (Gehl et al., 2006). Therefore, it is possible to evaluate the visual quality of streets by consistently measuring these design-related elements that can be extracted and measured from SVIs.

Identifying key design elements and their operational definitions

The key design elements used in this study are selected through a combined consideration of urban design theories and analytical techniques. Specifically, the first concern is how frequently these elements are mentioned in the existing studies discussing visual quality, and whether these elements can be easily described with operational definitions from the perspective of spatial features. The second concern is how easily they can be measured with existing machine learning algorithms.

We first review the constant efforts of urban designers to identify potential design elements. For instance, Jacobs (1961) mentions the importance of diversity. Trancik (1986) highlights the effects of enclosure, continuity, and building frontage. Similar discussions have been made by Gen and Pendola (2008) to discuss the effects between various elements like street scale, street width, building height and façade, and perceptual feelings.

Katz et al. (1994) adds the concern of being pedestrian friendly and having coherence. The street greenery, transparency (permeability), human-scale, legibility, and imageability, amongst others, are also often discussed (Montgomery, 1998). Following that review, an exclusion process was created. Some perceptual elements like coherence are excluded due to the lack of appropriate operational definitions. The remainder are all operational for urban designers as they point directly to specific design elements. We have also reviewed recent machine learning algorithms, including SegNet (Badrinarayanan et al., 2017), DeepLab (Chen et al., 2016), and YOLO (Redmon and Farhadi, 2018), to select key design elements, obtaining operational definitions within current techniques.

Finally, the following six key elements were selected: street greenery, sky views, building frontage, pedestrian space, motorisation, and diversity. The first five can be measured as the percentage of pixels representing greenery, sky, buildings, pedestrian paths and pedestrians, and motorways and cars, respectively. The last element, diversity, is measured as the percentage of pixels representing the rest of the design elements such as street lights, street furniture, etc. This operational definition of diversity is generated from a visual perceptual viewpoint, which is different from its traditional definition that mainly focuses on the functional mixture. These six elements can cover a majority of key urban design qualities as defined by Ewing and Handy (2009), especially enclosure, human-scale, and complexity (see online Appendix II). As for the rest of the imageability, as stated by Ewing and Handy (2009: 72), 'it is related to many other urban design qualities, e.g., enclosure, human scale, complexity, and is in some way the net effect of these qualities'. Therefore, the representativeness of imageability might be achieved by integrating all elements together. Transparency is operationally defined as the proportion of the first floor with windows, which is hard to directly measure due to the lack of algorithms that work with low-resolution SVIs. However, it is still partially related to the proportion of the building frontage as these operational definitions are all based on the visibility of the building frontage. In short, the relatively weak representativeness of imageability and transparency might not produce over-biased estimates in this study. It will be added in future studies with the help of technical improvements.

Related studies

Quantitatively measuring the perceptual-based visual quality of streets via small-scale data

Aside from the classical, subjective discussions, there are also quantitative studies using small-scale data to analyse visual quality. Specifically, there are two main quantitative approaches – received preference and stated preference. The former operates via mapping to observe behaviours on the street and abstract-related design elements, and then runs statistical analyses. Gehl et al.'s (2006) series of studies is a good example. The latter operates by collecting judgements from experts. Ewing and Handy's (2009) work shows a good example of this technique, as it employs an innovative method and quantitatively measures the effects of design elements on perceived qualities in urban design. Our study in this paper is similar to that of Ewing and Handy as it is also built on the modelling of expert scoring and selected key elements. Nevertheless, the proposed approach, which employs machine learning algorithms and large-scale SVIs, is more automatic and time-saving, especially for a citywide analysis.

Recent studies utilising machine learning algorithms and large-scale SVIs

Recently, a few interesting studies have emerged that have accompanied the fast development of machine learning algorithms in computer vision and the availability of SVIs. As an initial exploration, Doersch et al. (2012) attempted to answer, 'What makes Paris look like Paris?' through a machine learning-based clustering method and Google SVIs. Following this trend, some scholars have gone further to explore the perceptual and socio-economic performance hidden in SVIs (He et al., 2017; Quercia et al., 2014). Meanwhile, other scholars have paid attention to the measurement of physical and design elements in streets to assist in urban planning and design (Liu et al., 2017; Seiferling et al., 2017).

While we appreciate these exploratory works, our approach is somewhat different from the studies above that directly measured public perception from urban images. First, our approach aims to use the integration of urban designers' common sense to achieve an abstract of classical urban design theories on visual quality. It is a design-oriented study attempting to integrate classical findings and newly emerged techniques. Second, our approach focuses on the extraction of appropriate design elements from an urban design perspective. All these design elements are operational in urban design practice and guidance, which directly points to specific measures for improvement. Insights from the urban design viewpoint can directly contribute to addressing the rising demand for street quality.

Compared to studies sharing a similar focus on physical appearance and urban design qualities, our approach encompasses more complex considerations. Existing studies in this field primarily focus on only one or two design elements. For instance, a series of studies focus on measuring eye-level street greenery and their corresponding effects (Long and Liu, 2017; Seiferling et al., 2017). Yin and Wang (2016) have measured enclosure via machine learning algorithms and SVIs. Liu et al. (2017) have developed a machine learning-based method for evaluating the built environment quality via two issues: the construction and maintenance quality of a building's façade, and the continuity of the street wall. In short, a comprehensive and design-oriented consideration is still lacking, which is addressed in this study.

Methods

Analytical framework

This study contains four main phases, including data collection, extraction of features, evaluation effects, and finally, informing better design (Figure 1). First, SVIs and street networks of Shanghai were collected through the Baidu Maps API and Open Street Maps (OSMs), respectively. Subsequent image processing was performed in Python. Second, six key design elements affecting the quality of the street space were extracted from the SVIs through SegNet to achieve an objective and accurate pixel-wise image segmentation and measurement. The effects of the design elements were then measured. This phase contained three steps: (1) selecting sample images from the collected SVIs; (2) producing a Java-based program to collect the expert preferences through pairwise comparison; and (3) training an evaluation model though another machine learning algorithm, the ANN. Finally, this machine learning-based method is used to better inform urban design in the following three respects: (1) enabling a human-scale measurement of visual quality of streets at a citywide level, (2) achieving full coverage and consistent results on key design elements and their relative importance, and (3) assisting better planning management and streetscape design.

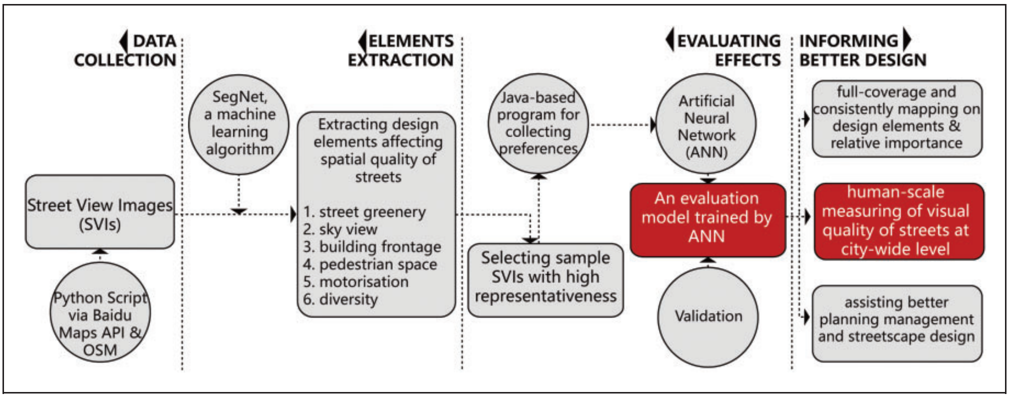

Figure 1. Analytical framework. API: Application Programming Interface; OSM: Open Street Map.

Study area and the collection of SVIs

Our analysis was conducted in the middle ring road area of Shanghai (in an area of 317.2 square kilometres), which is the town centre of this megacity (Figure 2). To achieve a comprehensive representation of the spatial features across the site, SVIs were collected every 40 metres. The 69,137 sample sites were generated along 13,672 streets with a total length of 2,611,079 metres. The SVI data collection was performed in April and May 2017. With the help of timestamps provided by the Baidu Maps, the SVIs initially taken in winter were deleted or replaced by alternatives to control for the effect of the season changing. China was less concerned with the street quality before the 2000s, as the country's urbanisation process was largely characterised as rapid and low quality in urban places. Recent critical thinking on the low quality of urban spaces has led to street renewal movements, especially for urban streets located in city centres. In this context, a quantitative assessment of the visual quality of streets becomes important.

The SVIs were requested in an HTTP URL form using the Baidu Maps API (Baidu, 2015). By defining the URL parameters sent through a standard HTTP request using the Baidu Maps API, users could obtain a static image from any direction and viewing angle, for any point where SVIs are available. A common shortcoming for existing SVI-related studies is that most of them use simplified headings measured with a reference direction towards the North Pole, e.g. 0° (true north), in the collection process. However, most street headings have certain angles with true north. Thus, these SVIs cannot reflect the street views that people perceive. To achieve more accurate data collection, the street headings were initially calculated through topological features computed from street networks to ensure that the front and rear views were always parallel with the street heading (see online Appendix III). The pitches (up/down angles of the camera relative to the horizon) passed to the Baidu Maps API were set to 0, to approximate a local resident's perception (Li et al., 2015). This pitch view helps to reflect the streetscape in the way that most people experience it according to priming theory in environmental psychology (Bargh et al., 1996).

Machine learning-based extraction of design elements

This work employed SegNet, an advanced deep convolutional neural network architecture that maps each image pixel into semantics, to extract the semantic information from SVIs,

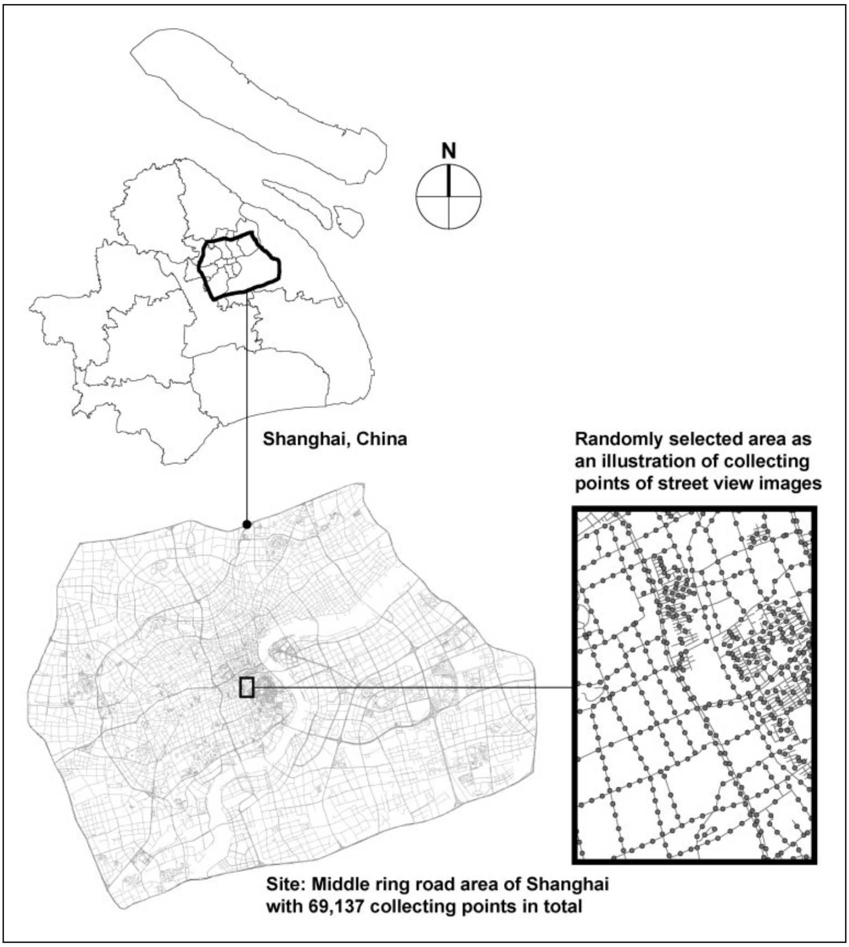

Figure 2. The site and collection points across the street network.

such as for the sky and buildings. Compared to other deep learning architectures, SegNet is efficient in terms of both memory and computational time. At the same time, it is capable of achieving good segmentation performance utilising low-resolution images. For street views, a global accuracy of 90.40% can be achieved for a total of 12 classes, and the accuracy is even higher for the classes of building, sky, car, and road (Badrinarayanan et al., 2017). We input all of the SVIs collected into SegNet and interpreted them into coloured categories via the SegNet decoder. As Figure 3 shows, the six key elements inside the images can be clearly extracted. The ratio of each element inside an image is based on the proportional results calculated by SegNet. The results for each sample site were subsequently calculated as the mean value of two images towards the front and rear.

Expert ratings

The evaluation of various key design elements was achieved by collecting the preferences of the experts (see online Appendix IV). First, 1500 SVIs were selected based on their

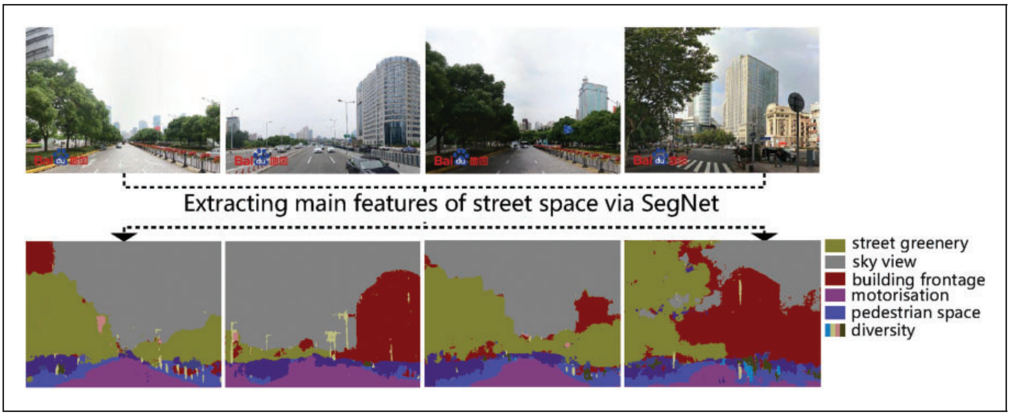

Figure 3. Applying a machine learning algorithm to extract key design elements.

representativeness in geometric distributions and image features. We divided the site into hundreds of grids, and the number of representative SVIs collected from each grid was then assigned according to the density of collected points. An automatic classification was then made via Python to ensure that the design elements of the selected SVIs were representative of their own grid. Second, a calibration of image brightness and chroma was made to exclude side effects. Then, a manual selection was made to select the first third with the highest representativeness to cover different typical streetscapes often appearing in Shanghai.

A Java-based program was then produced for collecting the expert panel's preferences on which SVI looks better via pairwise comparison. In other words, experts chose whether the left-side street view had higher quality than the right side or not. Ten urban design experts with professional degrees in architecture or urban planning were involved. All of them had enough experience living in Shanghai. In contrast to the previously renowned StreetScore study (Naik et al., 2014), we collected the preferences of experts rather than of the public for the following two reasons. First, the main focus of this study was to provide a workable approach for quick evaluations in urban planning and design practices. Compared with the public's evaluations, which are often biased by different individuals, the experts selected in our study had similar backgrounds allowing for more stable and similar results. It is possible to use a small-scale data set to achieve a satisfactory training of the evaluation model, which would allow this analytical approach to be easily applied in many other cities and to eventually assist in urban planning and design practices. In short, convenience is always important. It is easy to find limited numbers of experts in the urban planning and design committee of any city, but it is much more time consuming to collect millions of personal preferences from the public. Second, some views from the public might not be correct. For instance, many people in China may not have a clear understanding of what visual quality means with regard to streets. Some of them still believe that wide highways are a symbol of modernity and dislike relatively narrow but lively urban streets. It might be inappropriate to include all these viewpoints in the training data set.

Due to different personal perceptions, the labelled sample images were difficult to simply rank. Therefore, a pairwise comparison was applied. The scoring sheet method was not appropriate as it would be hard for participants to give a stable evaluation, which would have decreased the reliability.

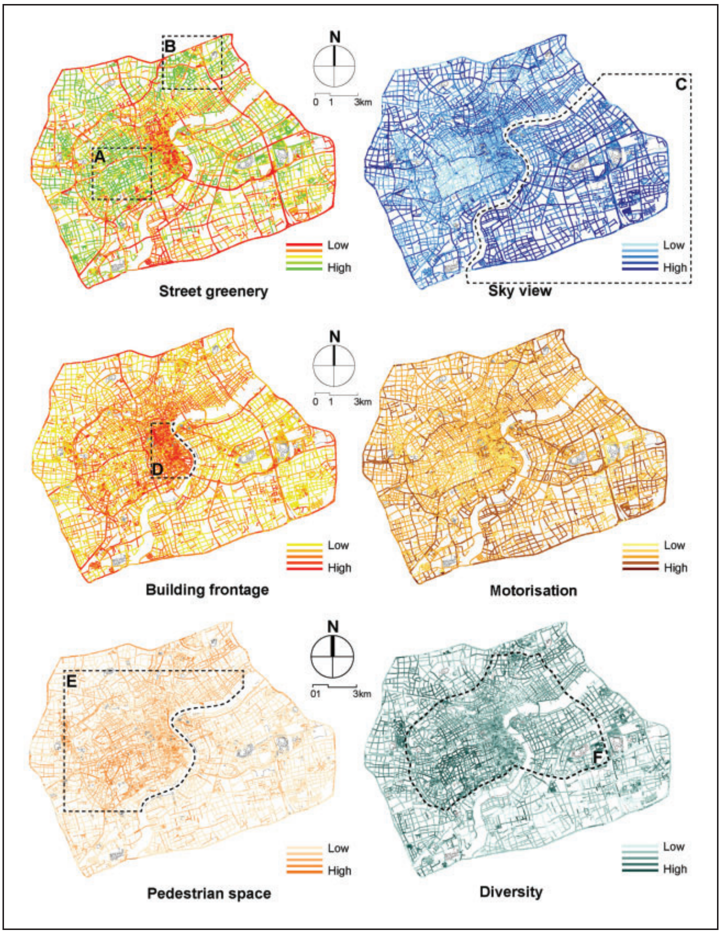

Figure 4. Key design elements on streetscapes. (a) Street greenery, (b) sky view, (c) building frontage, (d) motorisation, (e) pedestrian space, and (f) diversity.

Results

The measurement of key design elements

SegNet was applied to efficiently extract design elements from a large number of SVIs and measure their proportions (Figure 4). The co-presentation of a series of design elements that were difficult to directly measure within such a large scale brings many interesting findings. First, there are two main zones (A and B) that show highly visible street greenery. Zone A is a high-end historic community (Hengshan-Fuxing Road) that has maintained high quality greenery since the early 20th century. Zone B is the university town of Shanghai, which contains many green-filled campuses. From the perspective of sky views and motorisation, Zone C (the Pudong District) contains many roads that have high sky views and contains wide roads for transportation. In other words, this newly built area contains more transportation-oriented roads compared with the old town on the other side of the Huangpu River. Zone D, located in the Huangpu District, was historically called *xiazhijiao*, i.e. low-quality communities for poor people. Small alleys and high building density give high visibility of building frontages. It is also interesting to find that Zone E, the historical part of Shanghai, clearly has a higher proportion of pedestrian space compared to the other areas. Moreover, Zone F, the inner ring area, has higher diversity compared to the other areas.

Expert rating results and machine learning performance

After collecting expert ratings on the sample images, we converted the results from the pairwise comparison into scores. This transformation was achieved through the Elo (1978) rating system, as it is similar to the scenario of the relative skill calculation of players in competitor versus competitor games. The comparison results (*CR*s) were modelled as a set of triples T_i: $CR = \{T_1, T_2, T_3 \ldots\}$, where $T_1 = < lId,\ rId,\ result >$, lId, rId indicated the ID of the two images. Each sample image was initialised with an initial score, and its score was updated after each comparison. The whole updating process consists of two steps, prediction and updating. The *CR*s were entered into the rating system many times until the final results became stable. We then calculated the rating of all sample images by sorting them in descending order according to the score of each image (see online Appendix V). The ANN was then employed to investigate the relationship between design elements and image scores.

The ANN is a kind of computing algorithm inspired by biological neural networks. It is good at 'learning' due to its consideration of examples to perform specific tasks, e.g. using evaluated samples to rate the remaining ones (Schalkoff, 1997). An ANN consists of a series of connected nodes called *artificial neurons* that are mathematical functions conceived as a model of biological neurons. The artificial neuron, as an elementary element of the ANN, receives one or more inputs and sums them to produce an output. These artificial neurons are aggregated into different layers performing various kinds of transformations on inputs. The whole computing process usually starts from the first input layer, goes through one or many hidden layers, and finally reaches the last output layer.

Specifically, the ANN herein was trained using portions of design elements in an image as the explanatory variable and the scores as the responsive variable. Ten-fold cross-validation was used to determine the number of neurons in the hidden layers. Different combinations of hidden layers and the number of neurons were first tested with the 10-fold cross-validating approach. When the number of hidden layers increased, the result was observed with overfitting. We, therefore, narrowed our search using only hidden layers and tested its performance with a varying number of neurons. Further analysis showed that eight neurons produce a relatively good estimation, and this setting was used in the training of the evaluation model.

The normalised root-mean-squared error was calculated to test the predication performance, using the following equation

$$NRMSE = \frac{\sqrt{\sum_i (S_i - \hat{S}_i)^{2/n}}}{\sum_i S_i/n}$$

where S_i is the score of the ith image and $\hat{S}_i$ is the predicated score of the ANN.

An ANN was used as it fits well in complex non-linear regression and a comparison with other methods reports a higher accuracy. We compared the ANN method with the SVM and the decision tree. A grid search approach was adopted for hyperparameter optimisation of each method to find the optimised estimators. We then compared the optimised estimators by running each 100 times. The ANN had a slightly better result than the other methods with a mean NRMSE of 0.158 and standard deviation of 0.001, while the SVM reported a mean NRMSE of 0.160 and standard deviation of 0.0005, and the decision tree reported a mean R2 of 0.177 and standard deviation of 0.003. It should be noted that the *CR* is subject to hyperparameter optimisation, and the grid search may not necessarily have reached the best estimators for all methods. However, a detailed discussion of hyperparameter optimisation is beyond the scope of our study. We then applied it to the entire site to achieve a large-scale, high-resolution analysis.

As shown in Figure 5, the evaluated visual quality is illustrated with colours from red (low quality) to blue (high quality). Considering that the evaluation model was trained in a design-oriented way, urban streets in district and neighbourhood scales would more easily get higher ratings. Likewise, main roads due to traffic functions have more difficulty achieving positive evaluations on visual quality. Therefore, our current analysis on visual quality only focuses on urban streets. All the bridges, tunnels, trunks, expressways, and main roads

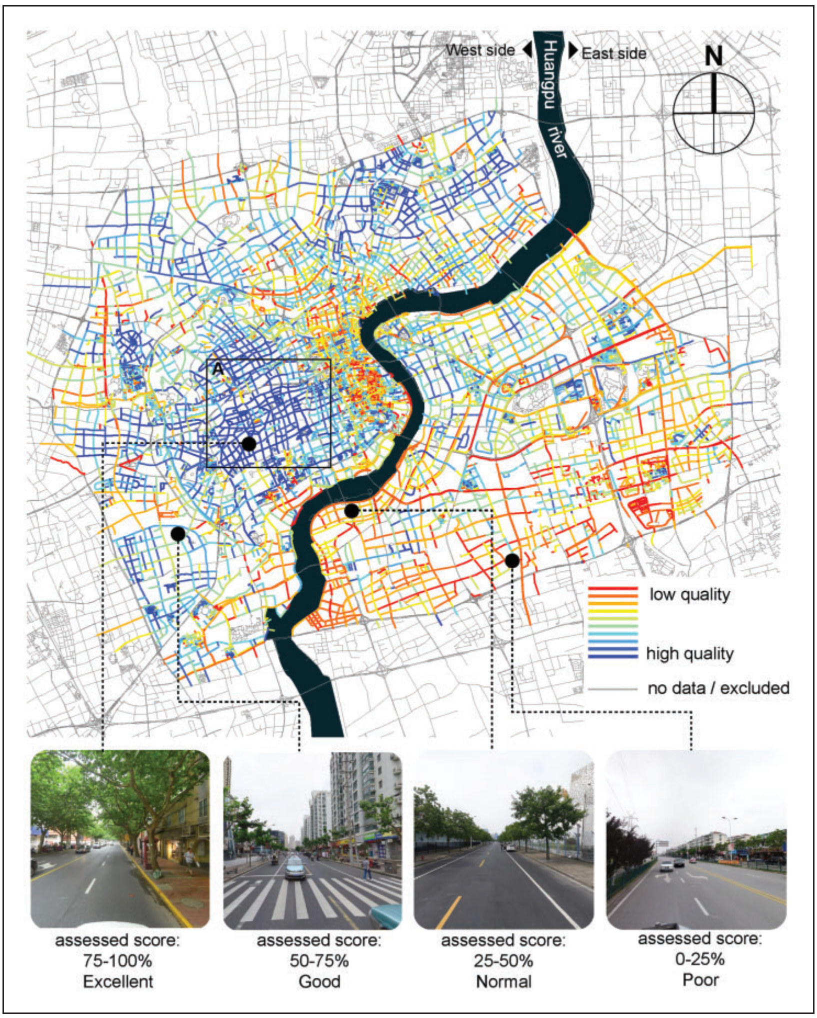

Figure 5. Measuring the visual quality of streets.

with more than four lanes are temporally excluded. Specific evaluating standards would be given to these transportation-oriented roads in our future studies.

In general, the overall results match classical urban design theories. Streets with high quality often have a human-scale streetscape, such as a well-integrated sense of enclosure, greenery, building frontage, and pedestrian space. In contrast, low-quality streetscapes tend to lack either these key elements or an appropriate combination of them. Moreover, the overall result from this evaluation matches the common sense of residents in Shanghai. The west side of the Huangpu River, which is the older part of the city, clearly has higher quality streets compared with its newly built counterpart on the east side. This is because the east side was developed over the past several decades under functionalism planning, which often lacked concern for human-scale visual quality. The Hengshan-Fuxing Road (Zone A) is assessed as high quality in our evaluation model, which matches the consensus of local residents.

Validating the evaluation results

The difficult part in the validation of the results is that the visual quality in nature is a normative value. The evaluated results produced by the ANN model are a priori of urban design; in other words, it is a generalisation of experienced urban designers' intuition on place-making. Therefore, it is hard to produce a totally definitive verification due to the absence of a standard reference. We herein applied an alternative approach to run validations. Specifically, we compared the evaluated results with the common sense of urban planning and design experts. Specifically, positive examples of streetscapes marked in *Shanghai Street Design Guidelines* (Shanghai Planning Bureau, 2016) were compared with the automatic evaluated results. Most good examples, 27 streets, were assessed as excellent (a score of 75–100%) in our analysis, and 7 streets were assessed as good (a score of 50–75%). The rest of the streets mentioned in the guidance are not included in the current analysis. In short, there is a good match between the evaluated results and the general understanding of urban designers.

Relative importance of the key design elements

Measuring the relative importance of design elements helps to provide additional information for the in-depth understanding of visual quality. Moreover, it also helps to inform efficient urban design interventions for street renewal projects. Garson's method (1991), later modified by Goh (1995), was adopted to identify the relative importance of each spatial element. It measures the relative importance of an explanatory variable from all the weights that connect it to the response variable through all the nodes in the hidden layers. The method involves partitioning the hidden-output connection weights of each neuron into components associated with each input neuron. The original algorithm uses the absolute magnitude from zero to one, whereas we preserved the signs so that the direction of the response can be determined.

Specifically, street greenery has the highest relative importance (0.81), followed by diversity (0.69), sky view (0.33), and pedestrian space (0.04). The relative importance of motorisation is negative (−0.19), which matches classical urban design theories. Considering that there is an occlusion relationship between street greenery and building frontage, the negative result of building frontage (−0.18) can be explained. The high positive effect of street greenery offsets the relatively low positive effect of the building frontage, which leads to a relatively negative weighting in the overall analysis.

Potential applications of planning management and streetscape design

This quantitative measurement of visual quality may better assist in urban planning and design from many directions. Here we show two potential applications for urbanists. Explorations will be extended in further studies. First, the result inside a large area that also has a high spatial resolution helps to achieve a quantified illustration of visual quality amongst different planning units, which would assist fine-scale planning management. Specifically, we can easily compare visual quality on streets amongst different districts and compute the average quality per street amongst different districts (Figure 6). The average quality is calculated as follows

$$Quality_{average} = \sum_{i=1}^{n} (Quality_i \times Length_i) / \sum_{i=1}^{n} Length_i$$

where $Quality_i$ and $Length_i$ represent the visual quality and length of the street i, and n is the number of streets inside each district.

The Pudong District has only 13.1% of streets with high-quality scores (i.e. within the highest quarter). In turn, there are 32.9% of streets with low-quality scores (within the lowest quarter) in this district. It also has the lowest average quality (score = 939.39). The Xuhui, Yangpu, and Changning districts perform much better, as all have nearly 40% high-quality streets. These three districts are also the best three in average quality, with average scores of 1150.84, 1110.81, and 1107.25, respectively. This approach helps to provide a measurable result for illustrating the visual quality of streets amongst various planning

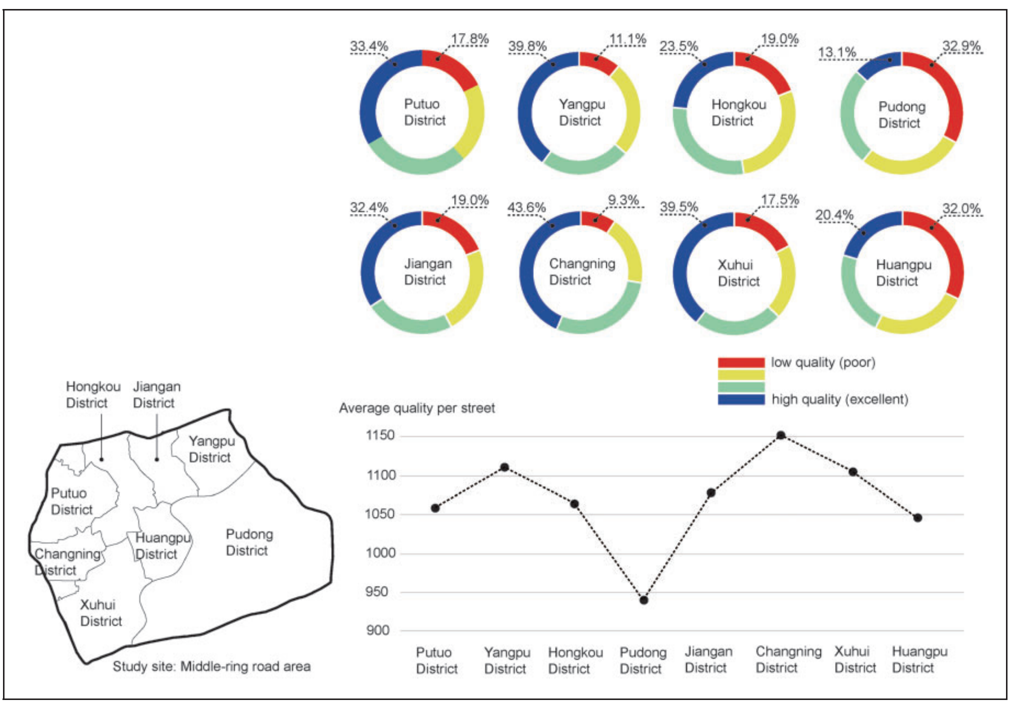

Figure 6. Comparing visual quality of streets amongst different districts.

units, which could contribute to data-informed zoning and an appropriate design strategy from a scientific perspective.

Second, the integration of evaluated visual quality and the internal feature of streets would help to classify the streets as different types for precise streetscape design interventions (Figure 7). The red colour represents the combination of low visual quality, i.e. the lowest one-third, with streets classified as 'residential street' in OSMs. Therefore, design interventions on these streets frequently used by local residents would have high priority. Meanwhile, these residential streets obtaining high visual quality (blue colour) should be well kept in the following years. As the two directions illustrated above, this quantitative measurement of visual quality could be used as a benchmark for further explorations on planning and design practices, which may provide more contributions to the streetscape design for certain purposes.

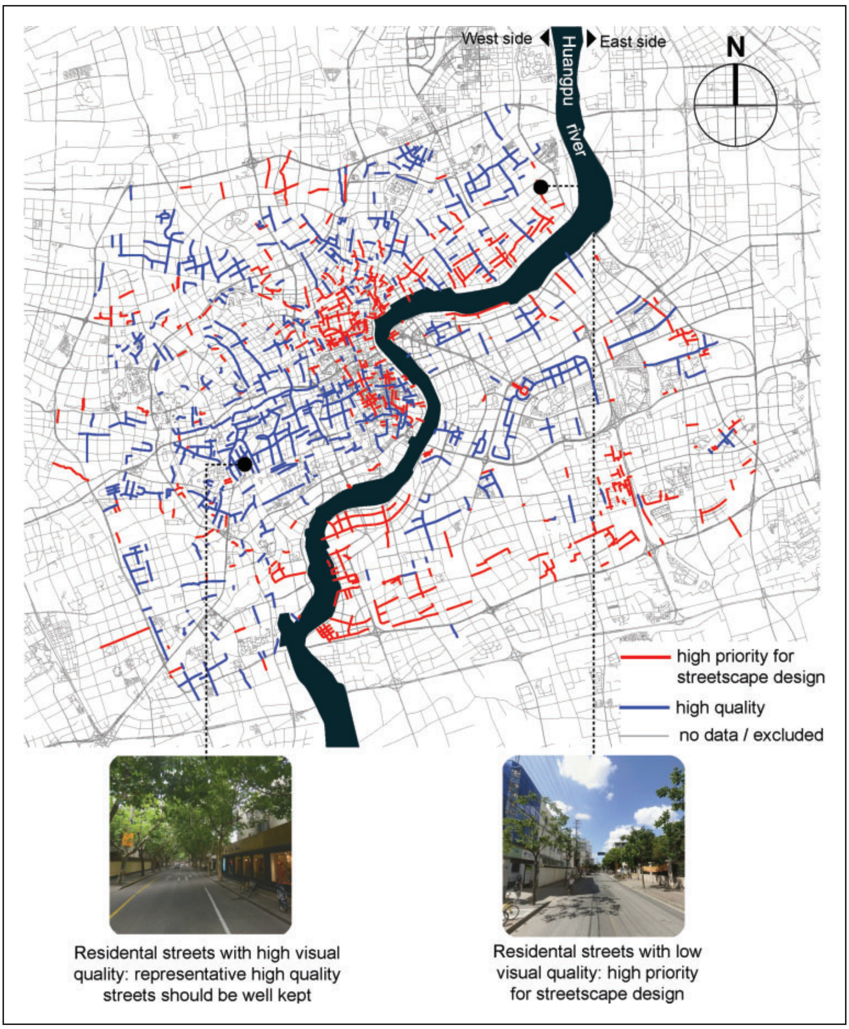

Figure 7. Identifying streets with priority for streetscape design.

Discussion

Measuring the unmeasurable with machine learning algorithms and SVIs

In contrast to recent studies utilising machine learning algorithms and SVIs, this paper focused on the urban design field to develop an automatic evaluation approach measuring the perceptual-based visual quality of streets. Using this method, measuring the visual quality of streets is no longer a challenging task relying most often on the subjective feelings of urban designers. This improvement can be a direct help in guiding human-oriented city zoning and urban design. The detailed measurements of design elements and the related visual quality could assist in selecting problematic streets, proposing appropriate zoning rules, and suggesting micro-scale design strategies according to the high or low values of different spatial elements and their relative importance on visual quality. For instance, for a neighbourhood street with enough pedestrian space and appropriate diversity, encouraging street greenery should be considered as a priority strategy in promoting high visual quality in streets. Moreover, this study also provides a workable approach for computing the relative importance of design-related elements that are difficult to clarify. This could serve as supplemental material for urban design theories. In short, these findings contribute to an objective measurement of previously subjective and experience-oriented issues.

Another contribution of this study is that it achieves a co-presentation of human-scale measurement with citywide analysis. Human-scale data have been incorporated into large-scale citywide analyses with difficulty. In turn, analyses focusing on human-scale data primarily operated on a small scale, such as block or neighbourhood scales, with concerns for time and costs. This shortcoming can now be overcome. The initial exploration in this study indicates that a large but also clear picture can be presented with the transition into a new science of cities, which integrates scientific thinking, design, and computer techniques (Townsend, 2015).

Potential applicability in the future

This study has the potential to be extended in both research and practical directions. For researchers, the measurement of visual quality provides a human-scale variable that can be analysed together with a series of behavioural and socio-economic variables. The co-evolution between visual quality, an urban design concern, with other social-economic features, such as property price, health, and pedestrian and cycling behaviours, may gain considerable interest from urbanists. In addition to this research potential, this workable approach has the potential to assist urban planning and design from many perspectives. For instance, the regular updating of SVIs helps to achieve a consistent and long-term monitoring of this human-scale urban design quality. This on-time mapping could identify whether an area is improving or deteriorating, and then allow responsive policy-setting. In addition, the combined analysis of street quality and behaviour data, such as pedestrian and cycling records, would help to identify streets that are low quality but are frequently used by the public.

Limitations and future improvements

There are also several limitations to the methodology described in our paper. First, this analytical approach has its limitations in fully representing 'quality' and it is not able to completely replace the traditional qualitative methods of spatial analysis. The evaluated result in this study is mainly a perceptual-based visual quality without consideration of sound, smell, noise, etc. Many well-discussed issues in traditional methods, e.g. human emotions and sense of community, are also not included. At present, it may work as a good supplement to existing methods by providing a fast evaluation. Further endeavours are still needed to extend this capacity. Second, another key element mentioned in many urban design theories, i.e. transparency, is hard to measure directly due to technical restrictions. We are developing appropriate algorithms that would accurately measure this issue. Third, the size of the expert-labelled data set is not very large, which means we might not completely utilise the capacity of machine learning algorithms. Extra sample data could be collected in the future to achieve higher evaluation accuracy. In addition, the gap between expert scoring and public experience is worth further exploration. The present study operated from an expert-oriented perspective, but some differences might exist between preferences of the public and that of experts. Large-scale data collection from local residents could be a good path forward in the future. Finally, further explorations should be made to classify streets by integrating their own functional and typological concerns with evaluated visual quality. This combined analysis may provide a good benchmark for many meaningful discussions as this new analytical approach could show higher validity based on a specific and scientific perspective. Studies in this direction may provide contributions to the streetscape design with various purposes.

Conclusion

This study developed a machine learning method to quantitatively measure the visual quality of streets – a key urban design concern that is hard to measure objectively. We chose six design elements inside streetscapes as the starting point. By combining eye-level SVIs, deep convolutional networks, and ANNs, an evaluation model has been trained to achieve a satisfactory performance. We expect this study to stimulate human-oriented and quality-of-place-oriented explorations in the era of new urban data and new analytical techniques. It is a good response to the increasing scholarly interest in introducing systemic and scientific thinking into the previously qualitative and intuition-based field of urban design (Ye et al., 2017).

Acknowledgements

The authors want to thank the editor and two anonymous referees for their valuable comments and guidance. Haoming Tang and Hanting Xie provided help for this study as well.

Declaration of conflicting interests

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The authors are grateful for the financial support from National Natural Science Foundation of China (51708410, 51878428, 61802388) and Shanghai Pujiang Program (17PJC107).

ORCID iD

Yi Lu http://orcid.org/0000-0001-7614-6661

Supplemental material

Supplemental material for this article is available online.

References

Badrinarayanan V, Kendall A and Cipolla R (2017) Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39(2): 2481–2495.

Baidu (2015) Baidu maps API. Available at: http://developer.baidu.com/map/reference/ (accessed 22 August 2017).

Banerjee T and Southworth M (1991) *City Sense and City Design. Writings and Projects of Kevin Lynch*. Cambridge: MIT Press.

Bargh A, Chen M and Burrows L (1996) Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology* 71(2): 230–244.

Chen F (2014) Urban morphology and citizens' life. In: Michalos A (eds) Encyclopedia of Quality of Life and Well-Being Research. Dordrecht: Springer, pp.6850–6855.

Chen LC, Papandreou G, Kokkinos I, et al. (2016) DeepLab: Semantic image segmentation with deep convolutional nets, Atrous convolution, and fully connected CRFs. arXiv preprint arXiv:1606.*00915*.

Doersch C, Singh S, Gupta A, et al. (2012) What makes Paris look like Paris? *ACM Transactions on Graphics* 31(4). DOI: 10.1145/2185520.2185597.

Elo AE (1978) *The Rating of Chessplayers, Past and Present*. New York: Arco Publishing Company.

Ewing R and Handy S (2009) Measuring the unmeasurable: Urban design qualities related to walkability. *Journal of Urban Design* 14(1): 65–84.

Garson GD (1991) Interpreting neural network connection weights. *Artificial Intelligence Expert* 6(4): 46–51.

Gehl J, Kaefer LJ and Reigstad S (2006) Close encounters with buildings. *Urban Design International* 11(1): 29–47.

Gen S and Pendola R (2008) Does "main street" promote sense of community? A comparison of San Francisco Neighbourhoods. *Environment and Behavior* 40(4): 545–574.

Goh ATC (1995) Back-propagation neural networks for modeling complex systems. *Artificial Intelligence in Engineering* 9(3): 143–151.

Gokce D and Chen F (2018) Sense of place in the changing process of house form: Case studies from Ankara, Turkey. *Environment and Planning B: Urban Analytics and City Science* 45(4): 772–796.

Handy SL, Boarnet MG, Ewing R, et al. (2002) How the built environment affects physical activity: Views from urban planning. *American Journal of Preventive Medicine* 23(2): 64–73.

He L, Páez A and Liu D (2017) Built environment and violent crime: An environmental audit approach using Google street view. *Computers, Environment and Urban Systems* 66: 83–95.

Jacobs J (1961) *The Life and Death of Great American Cities*. New York: Random House.

Katz P, Scully VJ and Bressi TW (1994) *The New Urbanism: Toward an Architecture of Community*. New York: McGraw-Hill.

Li X, Zhang C, Li W, et al. (2015) Assessing street-level urban greenery using Google street view and a modified green view index. *Urban Forestry and Urban Greening* 14(3): 675–685.

Liu L, Silva EA, Wu C, et al. (2017) A machine learning-based method for the large-scale evaluation of the qualities of the urban environment. *Computers, Environment and Urban Systems* 65: 113–125.

Long Y and Liu L (2017) How green are the streets? An analysis for central areas of Chinese cities using Tencent street view. *PLoS One* 12(2): e0171110.

Madanipour A (1996) Design of Urban Space: An Inquiry into a Socio-Spatial Process. Chichester: John Wiley & Son Ltd.

Montgomery J (1998) Making a city: Urbanity, vitality and urban design. *Journal of Urban Design* 3(1): 93–116.

Naik N, Philipoom J, Raskar R, et al. (2014) Streetscore – Predicting the perceived safety of one million streetscapes. *In: 2014 IEEE conference on computer vision and pattern recognition workshops*, Columbus, USA, 23–28 June 2014, pp.793–799. Washington, DC: IEEE Computer Society, USA.

National Association of City Transportation Officials (2013) *Urban Street Design Guide*. Washington, DC: Island Press.

Quercia D, O'Hare NK and Cramer H (2014) Aesthetic capital: What makes London look beautiful, quiet, and happy? In: Proceedings of the 17th ACM conference on computer supported cooperative work & social computing, Baltimore, MD, USA, 15–19 February 2014, pp.945–955. New York: ACM.

Redmon J and Farhadi A (2018) YOLOv3: An incremental improvement. arXiv preprint arXiv:1804.02767.

Schalkoff RJ (1997) *Artificial Neural Networks*. New York: McGraw-Hill.

Seiferling I, Naik N, Ratti C, et al. (2017) Green streets – Quantifying and mapping urban trees with street-level imagery and computer vision. *Landscape and Urban Planning* 165: 93–101.

Shanghai Planning Bureau (2016) Shanghai street design guidelines. Available at: http://www.shgtj.gov.cn/xxbs/shij/201610/t20161024_697272.html (accessed 29 December 2016).

Theil P (1961) A sequence-experience notation for architectural and urban spaces. *Town Planning Review* 32: 33–52.

Townsend A (2015) Cities of data: Examining the new urban science. *Public Culture* 27(2): 201–212.

Trancik R (1986) Finding Lost Space: Theories of Urban Design. New York: Van Nostrand Reinhold.

Ye Y, Li D and Liu X (2018) How block density and typology affect urban vitality: An exploratory analysis in Shenzhen, China. *Urban Geography* 39(4): 631–652.

Ye Y, Yeh A, Zhuang Y, et al. (2017) "Form syntax" as a contribution to geodesign: A morphological tool for urbanity-making in urban design. *Urban Design International* 22(1): 73–90.

Yin L and Wang Z (2016) Measuring visual enclosure for street walkability: Using machine learning algorithms and Google street view imagery. *Applied Geography* 76: 147–153.

Yu Ye is an assistant professor in the Department of Architecture, College of Architecture and Urban Planning at Tongji University, China. He gained his PhD degree from The University of Hong Kong and he had worked as a Post-doctoral researcher in Future Cities Lab, ETH Zurich. His research mainly focuses on data-informed urban design utilising multi-sourced urban data and newly emerged analytical techniques, which has been published on Urban Geography, Urban Design International, Environment and Planning B, etc.

Wei Zeng is a computer scientist and visualisation specialist with research interests on visual analytics, information visualisation, human–computer interaction, and urban data modelling. He is currently working as an Associate Professor at Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences. He received his BE and PhD degree from the School of Computer Engineering at Nanyang Technological University, Singapore.

Qiaomu Shen is a PhD student at VisLab, Department of Computer Science and Engineering, Hong Kong University of Science and Technology. He got his bachelor and master degree at College of Computer Science, Sichuan University. His research interests include Information Visualisation, Graph Visualisation, and Urban visualisation.

Xiaohu Zhang is a postdoctoral associate at Singapore-MIT Alliance for Research and Technology. Before that, he worked as a post-doctoral researcher at the University of

Hong Kong. He received his PhD degree in 2014 from the University of Hong Kong as well. His research interests include Geographic Information System and Urban Mobility.

Yi Lu is an assistant professor at Department of Architecture and Civil Engineering in City University of Hong Kong. He received his master degree from National University of Singapore, and his PhD from Georgia Institute of Technology, USA. Lu is interested in the behavioural impact of built environment, such as how site design of an urban smart development project influence walking, how nursing unit design influence staff behaviour. He has published in leading journals such as *Environment and Behavior*, *Environment and Planning B*.