Landscape and Urban Planning 221 (2022) 104358

Contents lists available at ScienceDirect

Landscape and Urban Planning

journal homepage: www.elsevier.com/locate/landurbplan

Subjective or objective measures of street environment, which are more effective in explaining housing prices?

Waishan Qiu [a,*], Ziye Zhang [b], Xun Liu [c], Wenjing Li [d], Xiaojiang Li [e], Xiang Xu [f], Xiaokai Huang [g]

[a] Department of City and Regional Planning, Cornell University, Ithaca, NY 14850, USA
[b] Paul and Marcia Wythes Center on Contemporary China, Princeton University, Princeton, NJ 08544, USA
[c] School of Architecture, University of Virginia, Charlottesville, VA 22904, USA
[d] Center for Spatial Information Science, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa-shi, Chiba 277-8568, Japan
[e] Geography and Urban Studies, College of Liberal Arts, Temple University, Philadelphia, PA 19122, USA
[f] School of Architecture, South China University of Technology, Tianhe District, Guangzhou, 510641, China
[g] Graduate School of Design, Harvard University, Cambridge, MA 02138, USA

HIGHLIGHTS

- Objective measures collectively explain more price variances.
- Subjective measures individually show stronger magnitude.
- Less ambiguous perceptions exhibit more accurate prediction and larger variance.
- Less ubiquitous elements like person had stronger associations to housing prices.
- Prior studies using sky, tree and building views might result in biased estimation.

ARTICLE INFO

Keywords:
Street design qualities
Subjective measures
Human Perception
Property value
Street view imagery

ABSTRACT

Houses with better street design are found to relate to a price premium. Prior studies mainly present the street quality using objective indicators like tree counts and distance to parks with land use data, or most recently using the greenery view index extracted from street view imagery (SVI). We argue that objective indicators cannot completely describe people's sense of a place, as perception is a highly subjective process. We hypothesize that subjective measures using visual surveys could capture more subtle human perceptions, thus providing stronger predictive power to housing prices. However, the role of subjectively measured street design qualities is less known due to the lack of large-scale perception data. To test our hypothesis, we first collected designers' perceptions on five urban design qualities from pairwise SVIs rankings in Shanghai with an online visual survey. Unlike the mainstream of using generic image features, we followed urban design theory and used rule-based features, i.e., about thirty streetscape elements extracted from SVIs to train machine learning (ML) models to predict subjective perceptions. The predictive power of five qualities versus ten selected individual streetscapes on housing price were compared using the hedonic price model. Besides the standard ordinary least squares (OLS), spatial regression and geographical weighted regression (GWR) were also developed to account for the spatial dependence and heterogeneity effects. We found both subjectively measured design qualities and objective indicators outperformed housing structural attributes in explaining housing price. While the objective view indexes collectively explained more price variances, the five perceptions individually exhibited stronger strength. Third, less-studied perceptions like "human scale" showed stronger strength than commonly studied "safety" and "enclosure". Fourth, less-studied view indexes like "person" and "fence" outperformed ubiquitous features like trees and buildings. Lastly, prior studies might have resulted in biased estimations due to ignoring the multicollinear issues between the sky, tree and building views. Our study addressed the effectiveness of incorporating subjective perceptions at a micro level to infer housing prices. Correlations between subjective perceptions were strong while that of objective indicators were negligible, therefore subjective perceptions can

* Corresponding author at: Department of City and Regional Planning, Cornell University, Ithaca, USA.
E-mail addresses: wq43@cornell.edu, qiuwaishan@126.com (W. Qiu), ziyez@princeton.edu (Z. Zhang), xl4xw@virginia.edu (X. Liu), liwenjing@csis.u-tokyo.ac.jp (W. Li), lixiaojiang.gis@gmail.com (X. Li), arcity001@mail.scut.edu.cn (X. Xu), tjuhxk@gmail.com (X. Huang).

https://doi.org/10.1016/j.landurbplan.2022.104358
Received 8 June 2021; Received in revised form 8 January 2022; Accepted 15 January 2022
Available online 24 January 2022
0169-2046/

complement the objective indicators. The findings provide important reference to decision makers when selecting street quality indicators to infer urban design, city planning and community and housing development plans.

1. Introduction

1.1. The micro scale neighborhood environment and housing prices

It is not a novel idea that human perceptions, residents' behaviors, the physical environment, and the socioeconomic outcomes including housing costs are highly correlated (Glaeser, 2018; Kang, 2021; Naik, 2015). At the micro scale, the appearance of streets could affect human perceptions and behaviors. For instance, street scene with dense buildings and narrow streets (Harvey, 2014), or less sky view (Yin & Wang, 2016) are perceived safer. Such perception in turn would affect people's appreciation of the place (Ma, 2021), the walkability (Ewing & Handy, 2009), bikeability (Ito & Biljecki, 2021), mode choice (Cervero & Kockelman, 1997), route choice (Salazar Miranda, 2021), physical activity and physical health (Brownson, 2009; Jackson, 2003), mental health (Wolch, Byrne, & Newell, 2014), and the quality of life (Glaeser, 2018). At the macro level, aggregating these individual perceptions and behaviors, the quality of physical street environment interacts with the socioeconomic outcomes of the neighborhood and city. For example, neighborhood physical disorders, such as broken windows and abandoned housings were related to crimes (Wilson & Kelling, 2011); the visual safety of a neighborhood is positively reflecting the neighborhood income (Naik, 2015). Housing, as an important component of human society that is strongly associated with both physical and socioeconomic environments (Kang, 2020, 2021), its value is therefore affected by the surrounding street environment and the different human perceptions and social interactions tied to the micro-level street environment. That said, a better understanding of what makes the sense of place (Couclelis, 2009; Shaw & Sui, 2020) is crucial for understanding housing prices.

1.2. Subjective perceptions for HPM

Street design quality has been measured subjectively, objectively, or combined (Lin & Moudon, 2010). Subjective measures refer to perceptual ratings from survey and interview (Nyunt, 2015), while objective ones measure quantitative metrics from field data collection (Lin & Moudon, 2010), such as land use in Geographical Information System (GIS). A neighborhood is the physical, social, and economic environments in a nutshell. However, prior studies mainly use objective indicators (Chen, 2020; Pandit, Polyakov, & Sadler, 2014; Ye, 2019) to address street environment's impact on housing price with the HPM method (Rosen, 1974). The "sense of place" (Kang, 2021; Shaw & Sui, 2020), especially those perceptions attached to the micro-scale street design qualities (Ewing & Handy, 2009) have long been ignored due to lack of effective method and data at large scale.

Field audit is a common way to collect subjective street design quality. However, it must be conducted by trained researchers onsite, which is time-consuming and costly (Griew, 2013; Kelly, 2013; Queralt, 2021; Rundle, 2011), being difficult to implement at the urban scale (Salesses , Schechtner, & Hidalgo, 2013; Dubey et al., 2016). Recently, online visual source like Google SVI provide an alternative to the in-person street audit. As it is publicly available, the resources needed for large-scale assessment is dramatically reduced (Griew, 2013). SVI audit is reliable in measuring perceptions like aesthetic, disorder, and safety: it achieves high level of agreements with in-person audit; the result is reliable and consistent regarding the inter- and intra-rater agreements (Griew, 2013; Kelly, 2013; Queralt, 2021; Rundle, 2011). Thanks to crowdsourcing and artificial intelligence (AI) (e.g., CV, deep learning (DL) and machine learning (ML)), the integration of AI, crowdsourcing and SVI dataset has become high-throughput to collect large-scale urban perceptions (Naik, Philipoom, Raskar, & Hidalgo, 2014; Salesses et al., 2013; Zhang, 2018; Dubey et al., 2016) . Since our task is to map the urban scale perceptions on several urban design qualities, SVI audit is therefore an ideal approach to measure human perceptions associated with the micro-level street environment. Specifically, we summarize the following four knowledge gaps regarding street measures for housing price studies using SVI data.

1.3. Research gaps

First, how the subjective perceptions are related to housing prices is largely unknown. Prior studies mainly relied on objective indicators to proxy street qualities. For example, studies using GIS data found tree counts (Donovan & Butry, 2010), tree canopy size (Pandit et al., 2014), canopy coverage (Sander, Polasky, & Haight, 2010), and distance to parks (Kim & Carruthers, 2015) significantly affected housing prices. Most recently, using CV to extract streetscapes pixels from images as view indices, studies found tree (Ye, 2019), sky and building views (Chen, 2020) significantly affect housing prices. However, neither GIS data nor streetscape view indexes completely capture the subtle human experiences on the streets. Perception is a highly subjective process of attaining sensory information (Ewing & Handy, 2009). Subjective measures derived from perceptual questions could explain human behaviors more completely (Lynch, 1960). Additionally, although perceived safety, walkability, and enclosure have been measured by many prior studies, they were mostly to inform crime and walkability studies (Buonanno, Montolio, & Raya-vílchez, 2013; Naik, Philipoom, Raskar and Cesar Hidalgo, 2014; Yin & Wang, 2016; Zhou, 2019). Less has been discussed to infer housing prices. Other important perceptions in urban design, such as imageability, complexity and human scale (Ewing & Handy, 2009), have never been incorporated.

Second, few locally-collected data exist to describe the perceptions on the contemporary urban landscapes in cities from Mainland China. The only few studies incorporating human perceptions (Naik, 2015; Kang et al., 2020; Kang et al., 2021) were all built on Place Pulse data (Salesses et al., 2013; Dubey et al., 2016). While Place Pulse provide the largest and global-scale urban perception data, applying its data to other places is data-dependent (Yao, 2019). No SVI from mainland China was included. Only Hong Kong and Taiwan (Dubey et al., 2016) were sampled. Although respondents' demographics would not cause bias (Salesses et al., 2013), the street appearance from mainland China varies and could be very different. A locally collected perception dataset, for example, from potential homebuyers would be more appropriate to infer willingness to pay.

Third, urban planners have little agreement regarding what perceptions matter with social scientists. On the one hand, the Place Pulse focus on more personal and emotional perceptions, such as safety, lively, beautiful, boring, depressing (Dubey, et al., 2016), or safe, class, and uniqueness (Salesses et al., 2013). These dimensions are most relevant to residents' everyday lives, thus are favorable by social scientists. However, they provide less implications for designers and planners. On the other hand, urban design studies (Ewing & Handy, 2009; Ewing, 2006; Park, 2019) emphasize more "professional" perceptions like enclosure, human scale, imageability, complexity, and transparency. These perceptions are less accessible to an average pedestrian while being closely tied to zoning codes, thus can provide more actionable implications to urban planning (Ewing & Clemente, 2013; Qiu, 2021).

Fourth, even within the objective measures, a comprehensive perception has never been addressed enough. Only a few visual elements were used, such as trees, sky, and buildings (Chen, 2020; Fu, 2019; Ye, 2019). Other elements such (e.g., commercial signboards, street furniture) from urban design studies were ignored (Ewing & Handy, 2009). Moreover, only individual impact was tested, while their collective effects were ignored.

Lastly, to what extent subjective perceptions complement or conflict objective indicators in influencing property values is never known. Existing studies comparing both measures were concentrated in health (Lin & Moudon, 2010; Ma, 2021; Nyunt, 2015), and they indicated poor agreements (Lee & Moudon, 2006). Lin and Moudon (2010) found objective measures showed more significant relationships with pedestrian behaviors, while Nyunt (2015) indicated the two measures were complementing each other. There is value of using both measurements in housing studies is unclear, but it is possible that they can help to clarify and corroborate the meaning of the counterpart.

1.4. Research questions and contribution

Using SVI data, which measurement exhibits stronger strength in affecting housing prices? How do subjective perceptions complement or conflict with objective view indexes? We evaluated the impact of five important perceptions on property values in Shanghai, namely enclosure, human scale, complexity, imageability, and safety. Our contributions are three-fold. First, while prior studies merely focused on several objective view indexes, we enriched housing price studies by revealing the effectiveness of subjective perceptions. Second, we provide important reference on the pros and cons as well as their complementary effects between subjective and objective measures. Third, our analytical framework enriched traditional urban design measures with crowd-sourcing surveys, opensource SVI datasets and artificial intelligence (AI) applications. While Ewing and Handy (2009) manually counted streetscapes from street videos to explain perception scores, we applied CV segmentation models to efficiently extract and count features. While the few existing studies (Kang, 2020, 2021) based on Place Pulse is data-dependent thus would provide limited implications for China, we locally collected and measured five perceptual qualities' impact on housing price in Shanghai. Moreover, the perceptions we measured were identified by urban design literature thus provide more implications to urban planning and urban design. It is worth noting that our intention is not to make any causal statement. The goal is simply to use the correlation to justify the effort and value of incorporating extra micro-scale urban perception data. Our study provides important reference to economists and planners for urban perception indicator selection in explaining housing prices, enriching the locally trained subjective perception mapping for Chinese cities.

2. Literature Review

2.1. Hedonic Price Model and Spatial Dependence

The marginal implicit price of street environments can be inferred from the HPM method (Rosen, 1979) which deconstructs property value into component attributes including structural, neighborhood, and locational attributes. Recent years, environmental characteristics have also been added to infer the value of street and neighborhood appearance. Ordinary Least Squares (OLS) regression is the most used HPM (Huang, 2017) owing to its great interpretability. However, an OLS is valid only if its residuals are not spatially autocorrelated. The spatial autocorrelation is mainly evaluated by calculating Moran's I spatial statistic (Anselin, 1988a), whereas a value of 1 (-1) indicates that similar values are perfectly clustered (perfect anti-correlation). The spatial dependence often remains in housing price data's regression residuals, which could lead to biased parameter estimates and type I error (i.e., falsely rejecting the null hypothesis) (Dormann & C., 2007).

Both theory and empirical studies suggest including spatial lag and error terms to account for the spatial interaction (Anselin, 2003; Kim & Carruthers, 2015; Manski, 1993; Pandit et al., 2014). The intuition behind is that: from the supply side, nearby houses share more similarities in structure, accessibility, and physical environment; from the demand side, homebuyers could emulate one another's behavior, especially within submarkets; from the financing side, an underwriter also looks at nearby transactions to establish value reference before a loan is issued.

Another common problem with housing price data is spatial heterogeneity (Anselin, 1988c; Geniaux & Martinetti, 2018), a data pattern where continuous change and uneven distribution is presented. Geographically Weighted Regression (GWR) model is an appropriate alternative (Cleveland & Devlin, 1988; Fotheringham, Yang, & Kang, 2017). GWR provides a local regression of variables to calculate the coefficients. It not only identifies spatial heterogeneity in processes but also takes the advantage of the spatial dependence in data (Fotheringham et al., 2017). Therefore, it has been widely applied in HPM literature to explore the potential non-stationarity of relationships between housing prices and its explanatory variables (Huang, 2017; Kang, 2020, 2021). To improve the estimation of housing prices, spatial regression and GWR are advised while the OLS analysis sets the baseline.

2.2. Street Environment's Role in the HPM

Prior studies mainly presented street qualities by objectively measuring various characteristics such as building height, street width (Cervero & Kockelman, 1997), and streetscapes like sky, building, and mostly, the trees. Sander et al. (2010) found 10% more tree coverage would increase housing prices by 0.48% in Minnesota using hedonic price model (HPM). In Portland, street trees added $8870 to the sales price, while reduced listing duration by 1.7 days (Donovan & Butry, 2010). Besides trees, some other street features like street lighting (Willis, Powe, & Garrod, 2005), open space (Anderson & West, 2006), and street traffic flow (Larsen & Blair, 2014) were found significant association with housing prices based on the HPM approach. Meanwhile, several studies used objectively constructed indicators like Walk Score to investigate the impact of more comprehensive street qualities like walkability on property values (Kim & Kim, 2020; Pivo & Fisher, 2011).

2.3. SVI, CV and ML for Street Measures

Recently, with the wide application of SVI data, the efficiency and accuracy of human-centered street audit have been improved over the traditional field survey. SVI data is publicly available, capturing ground-level panorama view of urban streets for the city-wide region (Seiferling, 2017). Unlike on-site auditing, which is expensive and time consuming, SVI data is more efficient for large-scale objective data collection (Rundle, 2011). Griew (2013) validated that SVI audits achieve high agreement (75%-96.7%) with in-person auditing, and its inner-rater and intra-rater agreement are also high thus is a reliable method for assessing built environment characteristics. Moreover, modern CV segmentation models made it efficient and accurate to extract features as the view index to proxy street qualities (Dubey et al., 2016; Zhou, 2019; Qiu, 2021), count road traffic (Yin, Cheng, Wang, & Shao, 2015), and predict micro-level environments such as the sun glare (Li, 2019).

2.3.1. Objective Measures

Along this line, Li (2015) extracted street-level greenery as 'green view index' (GVI); Yin et al. (2015) counted pedestrians; Yin and Wang (2016) found sky view is as a significant indicator for "visual enclosure" to infer walkability. Most recently, studies emerged to construct equations recombining view indexes in proxy to perceptions. For example, Zhou (2019) measured psychological greenery, visual crowdedness, outdoor enclosure, and visual pavement using view indices extracted from SVIs to construct a visual walkability index. Ma (2021) recombined tree, sky, building, pavement, fence, road, and sign view indexes to proxy five perceptions (i.e., openness, greenness, enclosure, walkability,

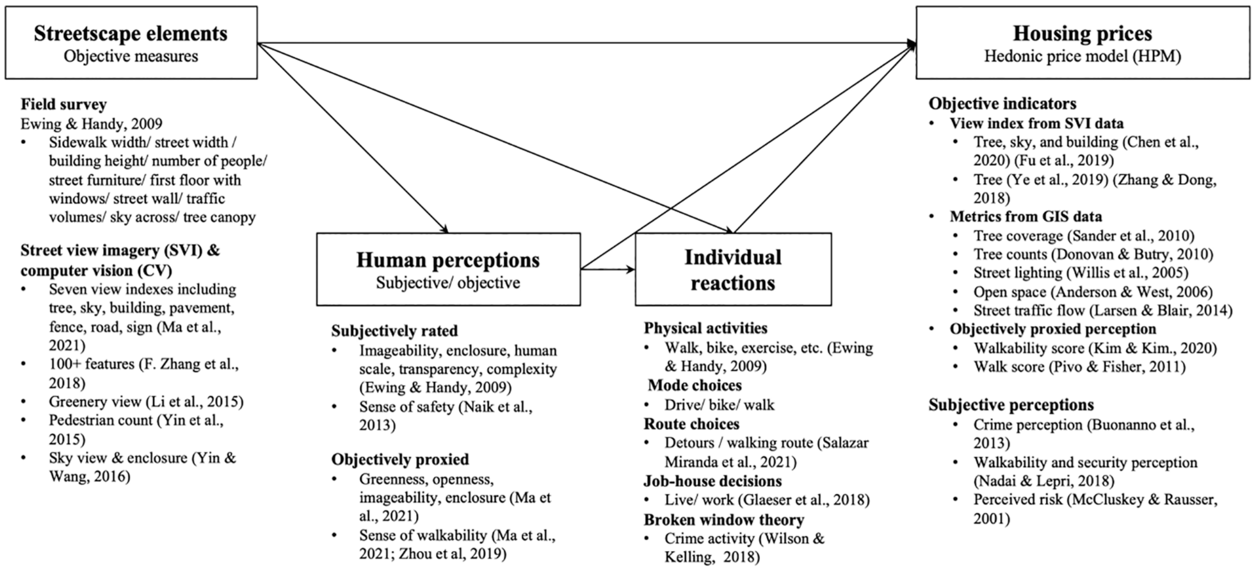


Fig. 1. Conceptual framework & key literatures. The design of this diagram is inspired by Ewing and Handy (2009).

and imageability) to inform urban renewal effects. Housing price studies also emerged to take objective indicators. Ye (2019) found greenery view obtained the second-highest coefficient in the HPM for Shanghai. Fu (2019) extracted greenery, sky and building views and found greenery and sky view increased housing prices in Beijing and Shanghai. Chen (2020) found a non-linear relationship between housing prices and GVI in Shanghai.

However, these new housing price studies were limited to objective measures. They only relied on view indices of individual features extracted from SVIs, which cannot fully represent residences' overall feelings on the environment (Ewing & Handy, 2009).

2.3.2. Subjective Measures

Subjective perceptions from surveys are believed to explain human behaviors more completely (Lin & Moudon, 2010) and have been reported to affect housing prices. Buonanno et al. (2013) found a standard deviation increase in perceived security is associated with a 0.57% increase in the district's average property value in Barcelona. McCluskey and Rausser (2001) found perceived risk around a hazardous waste site would lower property values. However, conventional field audits and surveys have been criticized for being inconsistent in the operation due to individual rater differences (Naik et al., 2014), time-consuming, expensive, and low-throughput (Zhang, 2018). Additionally, the results are difficult to interpret, providing less instructive implications to policymakers (Lin & Moudon, 2010). Nevertheless, objective indicators from the visual materials could be modelled to explain the subjective perceptions. Ewing and Handy (2009) reviewed 51 subjective perceptions and correlated five urban design qualities (i.e., imageability, visual enclosure, human scale, transparency, and complexity) to objectively-quantified streetscapes like the number of people and trees from field survey. The perceptual qualities were rated by an expert panel after watching street video clips.

Today, integrating crowdsourcing survey, SVI, CV and ML to uncover urban perceptions for large geospatial region is effective and reliable. It is worth noting the Place Pulse projects – the largest crowdsourced urban perception dataset for training scene understanding algorithms. Place Pulse 1.0 (Salesses et al., 2013) ranked 4100 SVIs from NYC, Boston Salzburg and Linz with inputs from over 7800 participants from 91 countries. In total, about 209 thousand pairwise votes were collected. Place Pulse 2.0 (Dubey, et al., 2016) collects ~111 thousand images from 56 cities and 1.2 million pairwise comparisons from 81,630 online volunteers along six perceptual attributes.

Built on the Place Pulse, studies explored the predictive power of different image features and compared the performance of different statistics models; the perception data were also correlated to various urban issues such as household income, neighborhood-level socio-demographic and economic changes, and homicides. With Place Pulse 1.0, Naik, Philipoom, Raskar, and Hidalgo (2014) explored the predictive power of generic image features on perceived safety and achieved an overall 0.57 R-Squared. Rossetti (2019) extracted both low- and high-level features from SVI as explanatory variables and applied discrete choice models. They found high-level features (e.g., view indices) increased both the model fit and the model interpretability. Using Place Pulse 2.0 data, Zhang (2018) predicted the six perceptions in Beijing and Shanghai. They also identified streetscapes that lead to different perceptions. Noticeably, two recent studies used the Place Pulse dataset to incorporate human perceptions measured from SVI to improve predicting housing prices (Kang, 2021) and appreciation rate (Kang, 2020). Although they addressed the role of subjective perceptions on housing prices, the perception data the Place Pulse is not appropriate to apply to that of Chinese urban landscapes. While one recent study locally collected perceptions in Wuhan (Yao, 2019), the perceptions were not derived from urban design qualities, and they were not used to infer housing price studies.

2.4. Conceptual Framework

To conclude, little has been done to construct local maps of subjective perceptions for the unique urban landscapes in China. The many important perceptions identified by classical urban design studies, such as imageability and complexity (Ewing & Handy, 2009), have never been incorporated into housing price models. It is also unknown whether subjective or objective measures of the street environment provide stronger strength of association to housing prices. Therefore, this study sets to locally collect subjective perceptions of the important street qualities informed from urban design literature using crowdsourcing and SVI data for Chinese large cities, taking Shanghai as an example. The hypothesis is that subjective perception may variate and complement the objective view indexes to better inform street environment's impact on housing prices (Fig. 1). The presence of physical features, such as sidewalks, tree canopy, buildings, and people, affect residents' perceived street design qualities, such as imageability. In turn, the perceived qualities affect residents' overall behaviors, including decisions to walk, to stay, and to live there, and consequently the

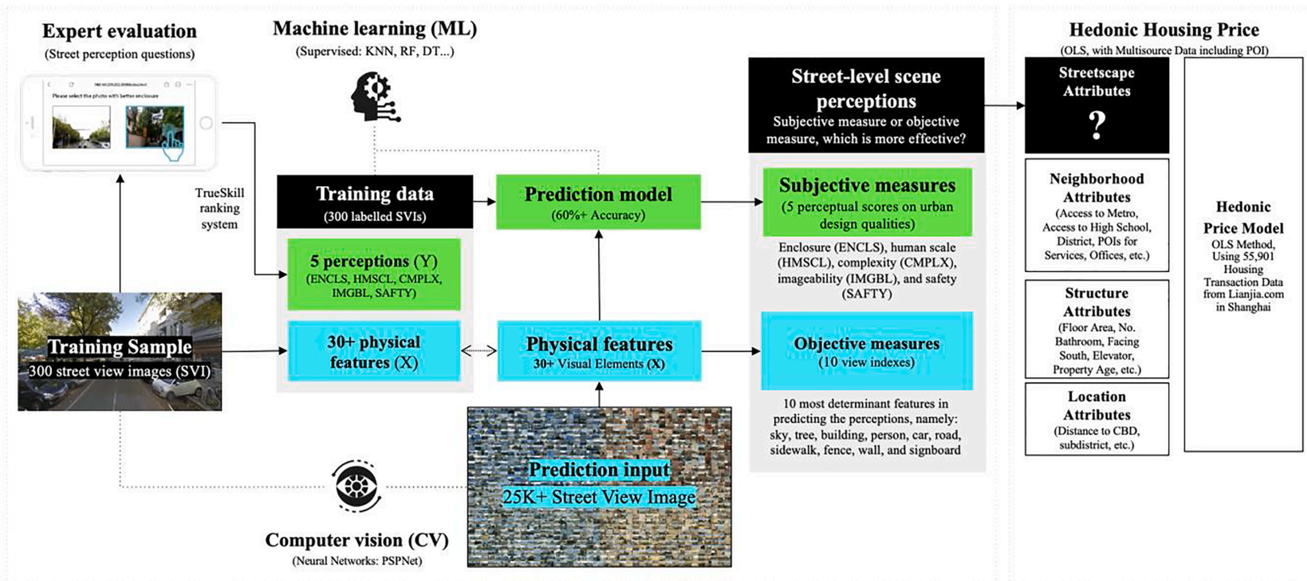


Fig. 2. Method and workflow.

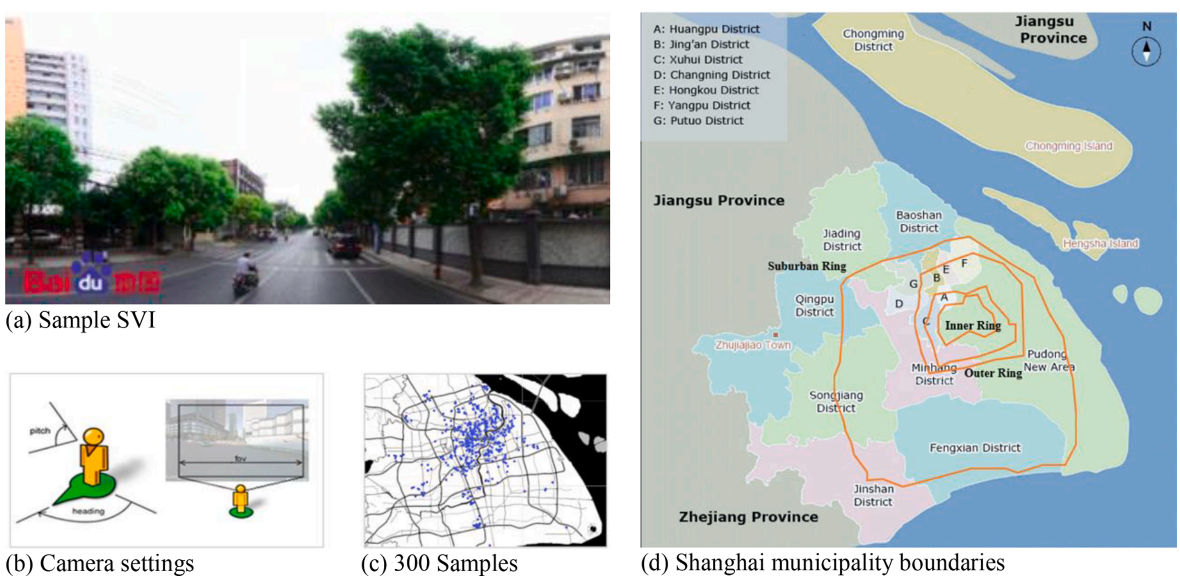


Fig. 3. Downloading Baidu SVIs (**a**) SVI example. (**b**) Camera settings diagram. (**c**) 300 SVI training samples. (**d**) Municipality boundaries.

housing prices.

3. Data and Methods

3.1. Analytical Framework & Study Area

First, we collected assessment of the five perceptual qualities on 300 randomly sampled SVIs from 43 participants in Shanghai using an online visual survey. Second, we extracted the pixel ratios of more than 30 physical features using a semantic parsing DL framework. Third, ML models were trained to predict the five perceptual scores using on the view indices extracted as explanatory variables. Fourth, we applied the best performance models to predict the subjective scores for all SVIs across Shanghai. Meanwhile, we calculate the Gini importance scores for the view indices to investigate the explanatory power of each streetscape feature. Last, we added both five subjective scores and the view indices of the ten most important physical features to the HPM and compared results from OLS, spatial regression and GWR. We quantified the strength of the subjective measures of streetscape perceptions on the neighboring housing prices (Fig. 2).

Accounting for spatial dependence and heterogeneity, we systemically compared the achieved standardized coefficients for five perceptual scores and view indices of the ten most important features, when other involved structural, location, and neighborhood attributes are controlled. Three questions are investigated: (1) To what extent would the subjectively measured eye-level streetscape perception affect housing price? (2) Would the subjective measured perceptions play a more important role than the objectively extracted view indices indicated by literature? (3) What are the divergent and coherent effects on housing prices?

Shanghai is China's financial center and has one of the most vibrant and expensive housing markets in China since the housing reform in 1998 (Chen, 2020). An empirical analysis of Shanghai can provide important references for relevant housing price studies with highly dense metropolitans (Fig. 3d). It will also inform policy implications on housing and urban design to facilitate more equitable urban

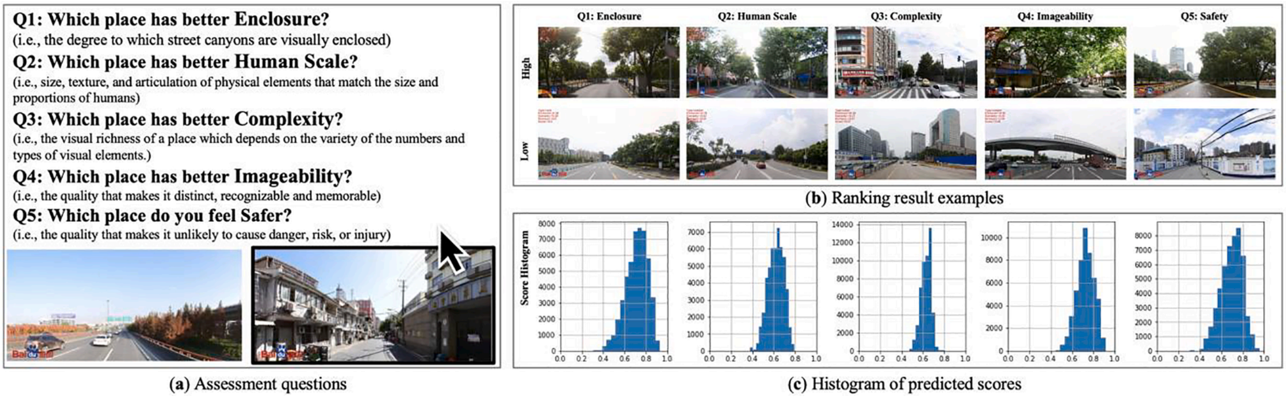


Fig. 4. Collecting perceptions (**a**) Online survey system. (**b**) High/low score examples. (**c**) Histogram of the score distribution.

environments.

3.1.1. Selection of Five Subjective Perceptions

Starting from urban design theory (Ewing & Handy, 2009), we include four design qualities, namely "enclosure", "human scale", "complexity", and "imageability". Additionally, the Place Pulse project urban scene understanding (Naik et al., 2014), and "safety", In total, these five qualities are selected to present subjective perceptions on street scenes. The definitions of these five perceptions have been relatively consistent, with the first four operationalized by Ewing and Handy (2009): Enclosure measures how a street is visually presented by vertical streetscapes like trees, walls and buildings (Yin & Wang, 2016); Human scale measures physical elements' texture, size, and articulation that match the scale of a human while correspond to walking speed, being correlated to visual elements such as pavement texture, facade details, street furniture, trees and plants; Complexity denotes the extent a street is visually rich, based on numbers as well as the variety of streetscapes such as street furniture, signage, greenery, building styles, and human activity; Imageability captures what makes a place distinct, recognizable, and memorable. Although 'transparency' has also been operationalized, it was deleted in the final models because we found its definition is relatively ambiguous, therefore, when converting rater preferences to ranking scores, this perception exhibit large variances making it difficult to converge. Additionally, perceived safety (from crime) was added, as it has been indicated to significantly affect resident behaviors (He, Páez, & Liu, 2017; Mehta, 2014; Naik et al., 2014; Buonanno et al., 2013).

3.2. Calculating Subjective Perception Scores

3.2.1. Downloading Baidu SVIs

SVI is appropriate to assess biking and walking environments. Ito and Biljecki (2021) and Ma (2021) both illustrated how a SVI angle sampled from street centerline could be useful to proxy pedestrian and biker's view and can be used to measure features usually quantified from Digital Evaluation Models. SVI audit is also reliable given its inter- and intra-rater consistency and its high agreement with in-person street audits (He et al., 2017; Rundle, 2011). Therefore, SVIs were downloaded from Baidu Street View Static API (https://api.map.baidu.com/lbsapi/) with constant camera settings (Fig. 3b). Baidu SVI provides a horizontal view of the street environment that is closer to the perception of pedestrians. Images are taken by street view cars along streets that snap 360-degree views at around 2.5 m high (Tian, 2021).

We sampled SVIs every 50 m (Chen, 2020; Ma, 2021; Yang, 2021) along the public street's centerlines within a 1 km radius for each property using their coordinates in QGIS. To best proxy the pedestrian view walking along the street (Fig. 3a), the tangent the street centerline was set as the camera 'heading'; 120 degrees the horizontal field of view (FOV); and 0 degree the 'pitch' which controls up/down angle of the camera. The image size was 640x360 pixels. In total, we obtained 25,276 SVIs. A 50 m interval was selected because Shanghai's average block size is 6.8 ha (NYU, UN Habitat and Lincoln Institute, 2016), with a 300–500 m length/width. Therefore, such an interval ensures 6–10 images represent a block's edge. A 1 km radius was selected because Chinese cities delineate a neighborhood by a 15 min walking distance, which is about 1 km (Zhou, 2019).

We also randomly selected 300 SVIs to use in the visual survey to collect perceptual preferences in the next step (Fig. 3c). To ensure the 300 images would cover the large variance of urban landscapes in Shanghai, we sampled them using the housing data points which spread throughout Shanghai city center, suburban, to the countryside and having more samples within the Outer-Ring Road. By doing so, we ensure that the 300 SVIs reflect the distribution of our housing samples across Shanghai. We also manually look at the 300 images to detect any blur, dark, or interior images (Ito & Biljecki, 2021; Salesses et al., 2013).

3.2.2. Collecting five subjective perceptions

Inspired by the Place Pulse 1.0 (Salesses et al., 2013) which integrated crowdsourced survey, CV, and ML, we built a crowdsourcing visual survey (http://140.143.239.153:3000/index.html#) where participants can choose a preferred photo from two random SVIs (Fig. 4a) in response to a question such as "Which place has better Enclosure?". A qualitative definition of Enclosure was provided. The vote did not allow a draw, and participants were not aware of the images' geolocations. Such a pairwise voting survey is more effective to reflect individual preference than a scoring system (Salesses et al., 2013).

The definitions of perceptions like imageability are ambiguous to pedestrians without urban design knowledge (Ewing & Handy, 2009). Therefore, all 43 voters were designers from a design workshop (DigitalFUTURES, 2020), including four undergraduates, 37 master students, and three faculties. 11 of them majored in Architecture, eight in City Planning and 14 in Landscape Architecture. Gender ratio is about 1.15:1 (male = 23, female = 20), and age distribution was between 20 and 35 (mean-age = 27.8, standard-deviation = 2.4). Compared to the expert panel from Ewing and Handy (2009) and Yao (2019) where only 10 and 20 urban designers/planners were recruited, the size and the diversity of our expert panel is larger. Moreover, all raters have experiences living in Shanghai and are either active homebuyers or future homebuyers.

The pairwise votes were translated to ranking scores using TrueSkill, a Bayesian skill rating system (Microsoft, 2005). The TrueSkill generates a ranked score for the winner and the loser by iteratively updating the scores after every two-player contest. In our study, the voting is the contest and the image selected is the winner. The scores were normalized into 0–1. These 300 SVIs with labeled scores became the training dataset for the ML models to predict five perceptions for all other unranked 25,276 SVIs. Fig. 4b depicted high and low score samples from the online survey. People tend to favor streetscapes with less sky exposure, more trees, and more pedestrian presence. Fig. 4(c) shows that scores were about normal distribution, a result expected from using the Bayesian system (Microsoft, 2005).

The 300-sample size was to balance prediction accuracy, survey reliability, and raters' workload. First, at least 75 to 100 samples or ten times the number of variables (Beleites, 2013) are necessary to train good ML models. Since about 30 streetscapes were extracted as explanatory variables, 300 samples would be sufficient. Second, on average, every SVI needed to be compared 12–36 times for the TrueSkill scores to converge (Herbrich et al., 2006). That said, to rank 300 SVIs, about 1800 – 5400 clicks (i.e., 300*12/2 – 300*36/2) are required. In our case, when 4426 pairwise ratings were collected from the 43 participants, the scores were stable. Each rater voted about 100 pairwise photos – a reasonable workload within three days. On average, each image was compared ~30 times (4426*2/300), much larger than Place-Pulse-1.0 (~16 comparisons/image) and Place-Pulse-2.0 (~3.4 comparisons/image). While we acknowledge that 300 was a rather small sample that largely limited the training accuracy, several recent studies also collected visual surveys of comparable sample sizes. For example, Verma, Jana, and Ramamritham (2020) recruited 73 participants to rate 200 images on six visual perceptions in Mumbai. Ito and Biljecki (2021) collected ratings on 400 SVIs on five perceptions related to bikeability for Tokyo and Singapore respectively. Additionally, urban design studies had much smaller sample size (Ewing & Handy, 2009; Park, 2019). For example, Ewing (2006) only have 48 video clips rated by 10 design and planning professionals.

In such a pairwise voting system, ranked scores represent the majority preferences, as each image was compared for about 30 times. Voters' sociodemographic differences would not cause serious bias, while the inter-user reproducibility and transitivity is also reported to be high (Salesses et al., 2013). When 22–32 votes per image are achieved, the TrueSkill system would produce a ranking with 75%+ transitivity (internal consistency) (Salesses et al., 2013). Therefore, like prior studies (Dubey, et al., 2016), the inter- and intra-rater reliability check was omitted in this study. However, we acknowledge this might cause problems to the final analysis, as the inter-rater reliability is not promised since we have very different SVI images and a much smaller sample size. This could be improved in the future by checking the inter-rater reliability.

3.2.3. Physical feature classification

View index is the pixel ratio of a streetscape object to the total pixels of an image (Fu, 2019), formally denoted by:

$VI_{obj} = \frac{\sum_{i=1}^{n} PIXEL_{obj}}{\sum_{i=1}^{m} PIXEL_{total}}, obj \in \{tree, building, sky, etc\}$ (1)where, VI_{obj} the view index, $\sum_{i=1}^{m} PIXEL_{total}$ the total pixels, and $\sum_{i=1}^{n} PIXEL_{obj}$ the pixels of streetscape feature *obj*.

A semantic segmentation framework – Pyramid Scene Parsing Network (PSPNet) (Zhao, 2016) was used to extract streetscapes pixels. PSPNet offers a higher accuracy rate in parsing complicated scenes with diverse visual elements (Gong, 2018), and has been used by many relevant urban studies (Chen, 2020; Fu, 2019; Verma et al., 2020). It predicts a class label for every pixel on the input image and avoids false segmentation by taking a pyramid pooling module to provide additional contextual information. Particularly, with street scenes collected from 50 cities in different seasons, the ADE20K dataset (Zhou, 2018) is tailored for semantic urban scene understanding. A high accuracy of 80.2% can be achieved in predicting 150 object classes of cityscapes with ADE20K. Therefore, we applied the pre-trained PSPNet based on ADE20K dataset.

3.2.4. Predicting subjective scores and analyzing feature importance

The training data is the 300 SVIs from the online visual survey, with their five ranked perceptual scores as labels, while the view indices extracted become the explanatory variables. We split them by 80% and 20% for training and testing purposes (Beleites, 2013). Tree-based models such as Random Forest (RF) (Breiman, 2001) are efficient to predict urban scene perceptions (Niu, Chen, & Yuan, 2020; Wang, 2019). With RF, Palczewska et al. (2014) demonstrated the correlation between urban function and urban perception; Yao (2019) fitted perception scores trained from expert panels ratings using ground streetscape views extracted from SVI. Since we have similar tasks regarding both prediction labels and explanatory variables, RF was selected to train for each perception separately.

Prediction results were compared by the balance performance judged by R-squared (R2), Mean Absolute Error (MAE), Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE). The R2 explains model fit, MAE is easy to interpret, RMSE punishes large errors, while MAPE can be used for cross-study comparison. The standard error is also reported with MAE to evaluate our prediction performance using Roy (2016) framework: an error of 10% of the training set range should be acceptable, more than 15% should be high, more than 20–25% is considered very high. Formally, a good prediction satisfies Eq.(2) while a bad prediction satisfies Eq. (3):

$$MAE \leq 0.1^{*}trainingsetrange, AND, MAE + 3^{*}std.dev. \leq 0.2^{*}trainingsetrange \tag{2}$$

$$MAE > 0.15^{*}trainingsetrange, OR, MAE + 3^{*}std.dev. > 0.25^{*}trainingsetrange \tag{3}$$

The best models were then applied to predict subjective scores for all 25,276 SVIs. We averaged scores from all SVIs within 1 km of a house (i.e., 15 min-walking-distance) to represent the surrounding neighborhoods' street qualities (Zhou, 2019).

Furthermore, Gini Importance (GI) was used to identify visual elements' importance in prediction (Chen, 2020). GI is calculated based on the impurity reduction of splits, which refers to the sum over the number of splits across all trees that include the feature, proportionally to the number of samples it splits (Nembrini, König, & Wright, 2018). GI score of each streetscape feature was calculated during a RF process using Scikit-learn in Python (Pedregosa, 2011).

3.2.5. Correlation analysis for perceptual scores

Subjective perceptions could have serious correlations, for example, Zhang (2018) found perceptions predicted from SVIs, such as "depressing-safe" and "beautiful-wealthy" were highly correlated using Pearson's correlation coefficient analysis. Similarly, the presence of various streetscapes could be correlated and bring multicollinearity issues. Therefore, the crossover Pearson's coefficient was reported to check the strength of the linear relationship between pairwise subjective perceptions, as well as the top ten objective view indexes. This will help to alleviate the multicollinearity issues when selecting explanatory variables for the HPM. Subjective perceptions and objective view indexes with strong correlations and high VIF will be removed from the HPM regression analysis.

3.2.6. Verification and cross-validation of the subjective scores

Using SVI to audit street environments is reliable, effective, and consistent (Griew, 2013; Kelly, 2013; Queralt, 2021; Rundle, 2011). The inter-rater and intra-rater consistency is ensured when 22–32 votes per image are achieved using a "two-player" survey design (Salesses et al., 2013; Dubey et al., 2015). With prior establishments, we did not conduct field audit to check the agreement between our prediction and field audit. However, this points to an important improvement to make. Nevertheless, we randomly selected four SVIs and plot their origin image, segmentation result, and score distribution to manually verify our results.

(a) By each apartment unit

count	53,445
mean	57,002
std	22,717
min	10,000
25%	41,353
50%	53,483
75%	68,698
max	472,111

(b) By each building

Count	7,430
mean	60,333
std	22,731
min	11,767
25%	45,314
50%	58,754
75%	72,864
max	250,000

(c) Histogram of housing price (up: by unit, bottom: by building)

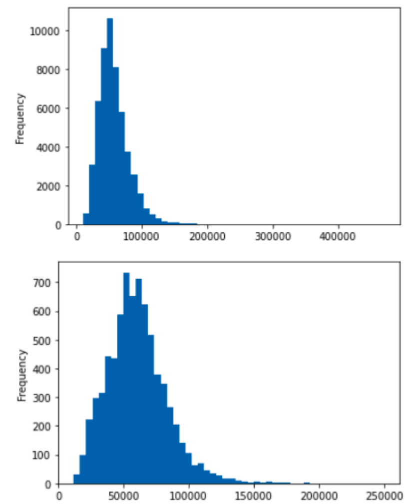


Fig. 5. Descriptive stats of price of **(a)** each apartment unit (b) each building.

3.3. Spatial Hedonic Price Models

HPM treats housing as a heterogeneous good, whose price is determined by three kinds of variables, namely the structural, location, and neighborhood attributes (Rosen, 1974). Structural attributes capture the house characteristics. Location attributes represent the property's location in the city. Neighborhood attributes describe the density and accessibility to important amenities and services. Although the subjective perceptual scores and the objective view indices belong to neighborhood attributes, to address the impact of street quality, we set them in a new attribute group (STRE). Formally, we extended the HPM as:

$PRICE = \alpha + \beta 1 STRU + \beta 2 LOCA + \beta 3 NEIG + \beta 4 STRE + \varepsilon$ (4)where α the constant, $\beta 1$ to $\beta 4$ the coefficients for structural (STRU), location (LOCA), neighborhood (NEIG), and streetscape (STRE) attributes, and ε is the error term.

3.3.1. Endogenous Variable – the Housing Price

Transactions in 2019 were collected from HomeLink/Lianjia.com (https://sh.lianjia.com/) – the largest real estate website in China (Fu, 2019). It provides price, structural and location information on each pre-owned apartment sold. Table 2 provides detailed description of the variables. In total 65,000 records were downloaded. We deleted records with missing property attributes, as well as those with unreal prices (e.g., no price, or per square-meter price ten times larger or smaller than the average). This left us with 53,445 records. Notably, several apartment units can locate within the same building having the same coordinate. By grouping the data using unique coordinates, we found 7430 unique buildings, that is, on average 7.1 units per building. The descriptive statistics of the price by apartment unit and the average price by building are depicted in Fig. 5.

Spatial unit of analysis. Having multiple housing points share the same location would raise a spatial overlap issue because when generating a spatial weight matrix, each point in the weight graph should have a unique location. Choosing buildings as the basic analysis unit seems to be more appropriate. However, we still chose the apartment as our spatial unit for analysis, because housing structural characteristics (e.g., floor height, whether facing south) could have large influences. These attributes cannot be averaged when aggregating apartments by building. To mitigate the spatial overlap issue, we split the 53,445 housing samples into 10 mutually exclusive random subsets (Kim & Carruthers, 2015), and discard the remaining 3455 samples. In other words, each subset contains 5000 unique apartments that do not have any spatial overlap. We run regressions on each of the subsets separately. This repeat-experiments technique (Kim & Carruthers, 2015) can demonstrate the robustness of the model if the results are rather stable. Notably, housing price was log-transformed in the regression models (Chen, 2020; Huang, 2017).

Autocorrelation in the housing price. Both the spatial autocorrelation and heterogeneity have been consistently reported on housing data (Anselin, 2003; Huang, 2017; Kim & Carruthers, 2015; Pandit et al., 2014). Therefore, Moran's I statistic is applied to the housing price data to check for spatial autocorrelation. The null hypothesis is that the variable is randomly distributed among the observations. We also use GEODA to generate a univariate local Moran's I cluster map to spatially visualize the clustering effects based on the location of the value and its spatial lag in the Moran scatter plot (GeoDa, 2020).

3.3.2. Independent Variables

Independent variables were selected based on literature and data availability (Table 1). First, structural attributes included continuous variables such as the number of bathrooms, year of construction, as well as categorical variables such as whether housing is south-facing, the interior decoration quality, and elevator availability. Categorical variables were transformed into dummies. Notably, the structural information contains the description of the interior decoration of the apartment units. Lianjia.com originally divides the decoration description into three categories: no decoration, simple decoration, and refined decoration. The decoration information is provided by either the apartment owners or their estate agents when listing the property. We combined no decoration and simple decoration into one category so that the DÉCOR variable can be described with one dummy.

Second, location attributes comprised the distance to (1) Central Business District (CBD, the Bund) and (2) nearest district center. Network distances were calculated based on road shapefile from Open Street Map (OSM) using QGIS. Fig. 6a shows a clear price drop when the distance to CBD increases (Chen, 2020; Rosen, 1974).

Third, neighborhood attributes measured the density or distance and accessibility to urban amenities and services, such as the number of restaurants, groceries, hospitals, and clinics POIs per km^2. Studies indicated that the vicinity of metro stations and high schools with good education quality can significantly increase the price (Wen, Xiao, & Zhang, 2017). Therefore, we geo-coded all metro stations as well as 68 high schools whose quality was recognized by Shanghai government as excellent. The road network distances from each property to the closest

Table 1
Descriptive statistics of all variables.

Variable	Description		Count	Mean	Std. Dev.	Min	Max	Data source
PRICE	¥/m^2, dependent variable		40,159	57,349	21,683	10,400	250,813	Lianjia.com
Structural attributes								
FLAREA	Total floor area (m^2)			85	43	15	588	
BEDRM	Number of bedrooms			2.1	0.8	1	8	
LIVRM	Number of living rooms			1.4	0.6	0	5	Web scraping from Lianjia.com
KITCH	Number of kitchens		40,159	1.0	0.2	0	5	
BATH	Number of bathrooms			1.2	0.5	0	7	
TTLFLR	Total floors of the building			11.0	7.9	1	62	
CSTRYR	Construction year of the building			1998	9.4	1912	2019	
	Description	**Values**	**Count**	**%**	**Avg. Price** (¥/m^2)	**Avg. Area** (m^2)		**Data source**
HGHT	On which floor in the building is the unit located?	0: Low/Mid	23,075	57.5%	59,021	90		
		1: High	17,084	42.5%	55,092	79		
TWR_SLB	The shape of the building	0: Slab	36,591	91.1%	56,346	85		
		1: Tower	3,568	8.9%	66,706	88		
STH_NTH	Is the unit south-facing?	0: Else	7,993	19.9%	56,110	94		
		1: South	32,166	80.1%	57,657	83		Web scraping from Lianjia.com
STRC	The structure of the building	0: Brick	18,003	44.9%	53,060	61		
		1: Steel	22,156	55.2%	60,819	105		
DECOR	Interior decoration quality	0: Simple	19,300	48.10%	53,680	74		
		1: Refined	20,859	51.90%	61,322	96		
ELEVTR	Is an elevator available?	0: No	24,106	60.0%	52,764	69		
		1: Yes	16,053	40.0%	64,235	110		
Locational attributes			**Count**	**Mean**	**Std. Dev.**	**Min**	**Max**	**Data source**
D2SCBD	Network distance (km) to district's geometric center		40,159	4.77	3.04	0.02	16.29	Shanghai road network shapefile (2018)
D2CBD	Network distance (km) to CBD (the Bund)		40,159	12.62	7.48	0.03	35.11	
	Description	**Values**	**Count**	**%**	**Avg. Price** (¥/m^2)	**Avg. Area** (m^2)		**Data source**
RING_X	Within which ring road is the unit located? X stands for the ring index.	1: Inner ring	9,290	23.1%	81,151	88		
		2: Middle ring	9,835	24.5%	63,057	79		
		3: Outer ring	8,742	21.8%	52,356	81		
		4: Outskirt ring	12,292	30.6%	38,345	92		
CTY_XX	Represents the district a property locates in. XX stands for the district name. *: In the final models, Baoshan district (CTY_BS) was the base group.	BS: Baoshan*	3,390	8.4%	44,159	81		
		CN: Changning	2,400	6.0%	70,051	83		
		FX: Fengxian	992	2.5%	24,524	95		
		HK: Hongkou	1,513	3.8%	66,210	80		
		HP: Huangpu	1,267	3.2%	92,725	103		Web scraping from Lianjia.com
		JA: Jin'an	964	2.4%	95,101	90		
		JD: Jiading	1,662	4.1%	37,527	87		
		MH: Minhang	4,806	12.0%	49,479	91		
		PD: Pudong	9,389	23.4%	57,590	87		
		PT: Putuo	2,941	7.3%	58,412	76		
		QP: Qingpu	678	1.7%	30,976	94		
		JS: Jinshan	2,201	5.5%	36,432	100		
		XH: Xuhui	3,060	7.6%	74,879	79		
		YP: Yangpu	3,091	7.7%	62,677	72		
		ZB: Zhabei	1,805	4.5%	63,647	79		
Neighborhood attributes			**Count**	**Mean**	**Std. Dev.**	**Min**	**Max**	**Data source**
DENSRV	Density of living amenities and services (1k/km^2)			0.115	0.187	0	3.5	Dazhongdianping.com
DENWRK	Density of office (1k/km^2)			9.52	22.45	0.00	573.5	
D2MTR	Distance to metro (km)		40,159	0.81	0.74	0.01	7.82	
A2MTR	Accessibility to metro			5.74	6.81	0.00	46.01	Calculated in QGIS
D2SCH	Distance to school (km)			2.72	2.33	0.02	11.94	
A2SCH	Accessibility to school			7.13	7.02	0.00	29.03	
Subjective streetscape attribute								
S1_ENCLS	Perceived enclosure			0.71	0.10	0.31	0.93	Predicted by ML models with view indices extracted from SVIs
S2_HMSCL	Perceived walkability			0.62	0.07	0.38	0.80	
S3_CMPLX	Perceived complexity		40,159	0.63	0.05	0.47	0.93	
S4_IMBLT	Perceived imageability			0.72	0.08	0.28	0.91	
S5_SAFTY	Perceived safety			0.70	0.10	0.31	0.96	
Objective streetscape attributes								
O1_SKY	Sky view index			0.25	0.08	0.01	0.48	
O2_TREE	Green view index			0.22	0.07	0.01	0.46	
O3_BLDG	Building view index			0.18	0.08	0.00	0.47	
O4_PRSN	Person view index			0.00	0.00	0.00	0.02	Derived scores by recombining selected physical feature view indices
O5_CAR	Car view index			0.02	0.01	0.00	0.11	
O6_ROAD	Road view index		40,159	0.10	0.03	0.01	0.19	
O7_SDWK	Sidewalk view index			0.03	0.01	0.00	0.13	
O8_FENC	Fence view index			0.01	0.01	0.00	0.09	
O9_WALL	Wall view index			0.02	0.03	0.00	0.42	
O10_SIGN	Signboard view index			0.00	0.00	0.00	0.05	

Table 2
Data summary of (a) view indices extracted from sample SVIs and (b) Gini importance.

(a) Descriptive summary				(b) Gini importance in predicting perceptions							
Sort	View Index	Mean	Std.	Sort	View Index	Sum	S1. Enclosure	S2. Human Scale	S3. Complexity	S4. Imageability	S5. Safety
1	Sky	39.7%	17.1%	1	Sky	1.05	0.40	0.23	0.10	0.09	0.22
2	Tree	21.7%	17.7%	2	Tree	0.60	0.07	0.12	0.09	0.14	0.18
3	Road	11.6%	6.4%	3	Building	0.44	0.06	0.08	0.12	0.09	0.04
4	Building	11.5%	13.8%	4	Person	0.39	0.05	0.06	0.15	0.10	0.01
5	Plant	2.1%	3.9%	5	Car	0.31	0.04	0.05	0.05	0.11	0.03
6	Wall	2.1%	5.4%	6	Road	0.26	0.05	0.05	0.04	0.07	0.10
7	Sidewalk	1.8%	2.6%	7	Sidewalk	0.24	0.04	0.05	0.06	0.04	0.05
8	Fence	1.7%	2.8%	8	Fence	0.22	0.03	0.04	0.06	0.05	0.03
9	Grass	1.5%	2.8%	9	Wall	0.18	0.02	0.04	0.04	0.05	0.06
10	Car	1.5%	2.6%	10	Signboard	0.18	0.03	0.03	0.05	0.05	0.02
11	Earth	1.1%	2.8%	11	Plant	0.17	0.05	0.03	0.03	0.03	0.08
12	Ceiling	0.6%	5.1%	12	Grass	0.14	0.03	0.04	0.03	0.05	0.04
13	Railing	0.3%	1.3%	13	Streetlight	0.13	0.02	0.03	0.04	0.03	0.00
14	Bridge	0.3%	2.6%	14	Minibike	0.08	0.02	0.03	0.02	0.01	0.00
15	Signboard	0.3%	0.9%	15	Earth	0.07	0.01	0.01	0.02	0.01	0.02
16	Water	0.3%	1.4%	16	Ceiling	0.07	0.01	0.03	0.01	0.01	0.00
17	Van	0.1%	0.7%	17	Railing	0.07	0.01	0.01	0.01	0.01	0.01
18	Person	0.1%	0.3%	18	Bicycle	0.07	0.01	0.02	0.03	0.01	0.01
19	Skyscraper	0.1%	0.8%	19	Bridge	0.07	0.03	0.02	0.01	0.01	0.00
20	Streetlight	0.1%	0.2%	20	Column	0.06	0.01	0.01	0.01	0.03	0.01
21	Column	0.1%	0.5%	21	Van	0.06	0.00	0.01	0.02	0.01	0.02
22	Minibike	0.1%	0.3%	22	Skyscraper	0.04	0.00	0.01	0.01	0.01	0.01
23	Bicycle	0.0%	0.3%	23	Ashcan	0.03	0.00	0.00	0.01	0.00	0.01
24	Awning	0.0%	0.3%	24	Awning	0.03	0.00	0.01	0.00	0.01	0.01
25	Ashcan	0.0%	0.1%	25	Windowpane	0.01	0.00	0.00	0.00	0.00	0.00
26	Windowpane	0.0%	0.3%	26	Mountain	0.01	0.00	0.00	0.00	0.00	0.00
27	Mountain	0.0%	0.2%	27	Chair	0.00	0.00	0.00	0.00	0.00	0.00
28	Fountain	0.0%	0.1%	28	Sculpture	0.00	0.00	0.00	0.00	0.00	0.00
29	Pier	0.0%	0.1%	29	Fountain	0.00	0.00	0.00	0.00	0.00	0.00
30	Chair	0.0%	0.0%	30	Booth	0.00	0.00	0.00	0.00	0.00	0.00
31	Booth	0.0%	0.1%	31	Water	0.00	0.00	0.00	0.00	0.00	0.02
32	Sculpture	0.0%	0.0%	32	Pier	0.00	0.00	0.00	0.00	0.00	0.00
33	Bulletin	board	0.0%	0.1%	33	Bulletin	board	0.00	0.00	0.00	0.00
34	Lamp	0.0%	0.0%	34	Lamp	0.00	0.00	0.00	0.00	0.00	0.00
35	Sofa	0.0%	0.0%	35	Sofa	0.00	0.00	0.00	0.00	0.00	0.00
36	Lake	0.0%	0.0%	36	Lake	0.00	0.00	0.00	0.00	0.00	0.00

metro station and high schools were calculated. Furthermore, we took the numbers of metro stations and high schools within 1 km and 5 km of a property as the accessibility measure. POI data was collected from Dazhongdianping.com in 2019. Locations of metro stations and schools were from AutoNavi's map in 2019 (Fig. 6b).

Lastly, we took the five perceptions as subjective streetscape attributes. For the objective indicators, the selection of features used in the final regression is first based on literature review in housing price studies and urban design studies with SVI and CV. Sky, tree, building view indexes have been tested by prior housing studies (Chen, 2020; Fu, 2019; Ye, 2019), while other features such as person, sidewalk, car, fence were indicated by walkability studies (Ma, 2021; Park, 2019). Additionally, these ten objective streetscapes all have modest to large existence in street views, as well as large Gini importance scores. We used these 10 view indices as objective attributes. Table 1 provides descriptive statistics for all variables.

3.3.3. Model Architecture

Implied by the large amount of literature, the spatial dependence and spatial non-stationary effect violate the basic assumptions of OLS. When both spatial effects present, OLS could results in bias estimation in the coefficients, and report false significance. Following the empirical studies in spatial HPM literature, we consider a Kelejian-Prucha model (Elhorst, 2010; Geniaux & Martinetti, 2018) and a GWR model (Cleveland & Devlin, 1988; Fotheringham et al., 2017).

(a) **A full spatial interaction model – the Manski Model.** When the spatial dependence is presented, there are three kinds of spatial interactions could exist, the lagged dependent variable (WY), autocorrelated error term (Wμ), and lagged exogenous variables (WX). When all three spatial effects present, it is the Manski model (Elhorst, 2010; Manski, 1993):

$Y = \rho WY + \alpha\iota_N + X\beta + WX\theta + \mu \text{where} \mu = \lambda W\mu + \varepsilon,$ (5)where Y (N*1) a vector of the dependent variable, X (N*K) a matrix of the explanatory variables, ρ (K*1) a vector of the spatial autoregressive coefficient, λ the spatial autocorrelation coefficient, ι_N (N*1) a vector of ones associated with the constant term parameter α, θ (K*1) a vector of fixed but unknow parameters, ε a vector of disturbance terms that are independently and identically distributed error terms. W (N*N) describes the spatial arrangement of samples, often defined by contiguity (e.g., Queen/Rook type), KNN, or Voronoi relationships depend on the spatial patterns within the data points. WY the endogenous interactions, WX the exogenous interactions, and Wμ the interaction effects among the disturbance terms of different spatial units.

(b) **OLS and Moran's I test in the residuals.** When the spatial interaction effects among Y, X, and disturbance terms are ignored (i.e., ρWY, WXθ, and λWμ are omitted), Eq. (5) is a non-spatial linear regression, a most common approach (Elhorst, 2010). We start with this specific-to-general approach using the OLS. Moran's I and the robust Lagrange multiplier (LM) tests (Anselin, 1988b; GeoDa, 2020) are conducted for all OLS residuals to test whether spatial interaction effects should be added using the PySal packages in Python (Rey & Anselin, 2010). The robust LM test would return p-value to imply whether spatial lag and error terms are presented. If so, they should be incorporated using Eq (3) appropriately.

(c) **The Kelejian-Prucha model**. In our case, the robust LM test implies both spatial lag and error effects exist. Therefore, we

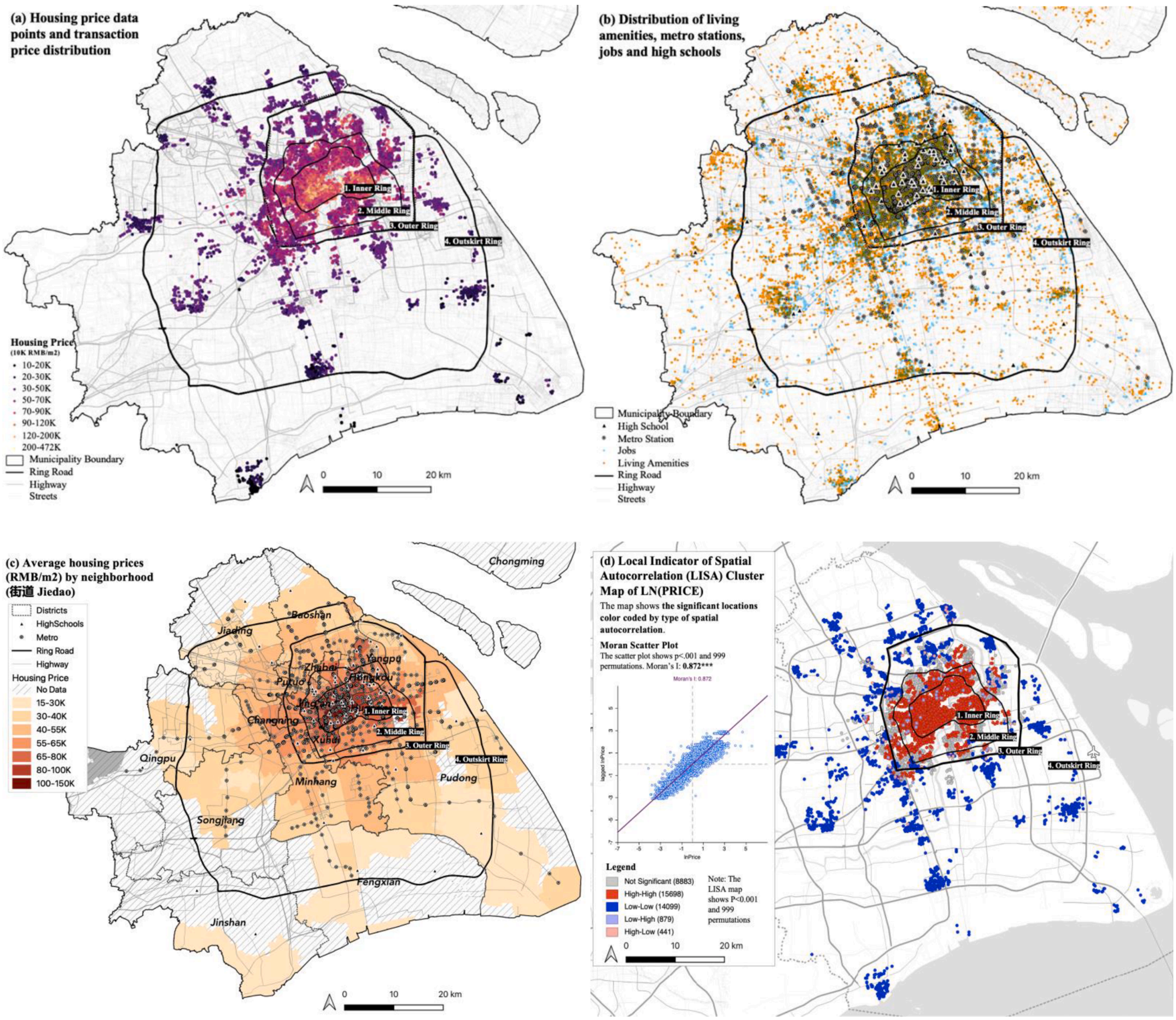


Fig. 6. Spatial distribution of (**a**) housing transaction price, (**b**) neighborhood attributes including amenities and service POIs, metro stations, jobs, and high schools, (**c**) average housing price (RMB/m2) by municipal neighborhood boundary, and (**d**) LISA cluster analysis (hotspot) of log price.

follow Kelejian & Prucha (1998) to use a model with both a spatially lagged dependent variable (WY) and a spatially autocorrelated error term (Wμ) (Elhorst, 2010), which is also known as a spatial autoregressive combined (SAC) model (Rey & Anselin, 2010).

(d) **GWR model.** Traditional regression assumes that the relationships being examined through the model's parameters are constant over space, a GWR relaxes this (Fotheringham et al., 2017). A semiparametric GWR (SGWR) extends GWR and allows local and global relationships to co-exist. We use SGWR from MGWR package in Python to investigate the spatial non-stationary issues.

(e) **Modeling selection.** First, we added each group of attributes into separate OLS models to understand their explanatory power. Second, a baseline model was constructed using structural, locational and neighborhood attributes. No streetscape measures were included. Insignificant variables were removed. Third, with the baseline, we added all five perceptual scores (Model 1) while all ten objective view indices (Model 2), respectively, to compare their different strengths. Variance Inflation Factor (VIF) was calculated to examine variables with correlation problems (VIF >10). The Gini Importance (GI) (Nembrini et al., 2018), which has been widely applied to identify relevant features based on the impurity reduction of splits, was conducted to rank the feature importance (Pedregosa, 2011). Less important variables with multicollinearity (VIF >10) were then removed (Kutner, Nachtsheim, & Neter, 2004).

To this end all the OLS models have been listed. Based on the Model 1 and Model 2, we add the spatial interaction terms of Y and error. Particular attention is focused on the change of coefficient sign, magnitude, and significance levels. The spatial weight matrix W is measured using KNN method and K = 85, meaning that the 50 closest nearby house values and their OLS residuals are included into the Kelejian-Prucha model estimation, such that spatial dependency is presented. In addition to the spatial autocorrelation on housing price and residuals, we add GWR to model the spatial heterogeneity within the relationships between the housing price and its determinants. The best bandwidth of the Gaussian kernel and Akaike Information Criterion (AIC) is estimated from the data using the MGWR package in Python. It is worth noting that the spatial unit for analysis is an individual

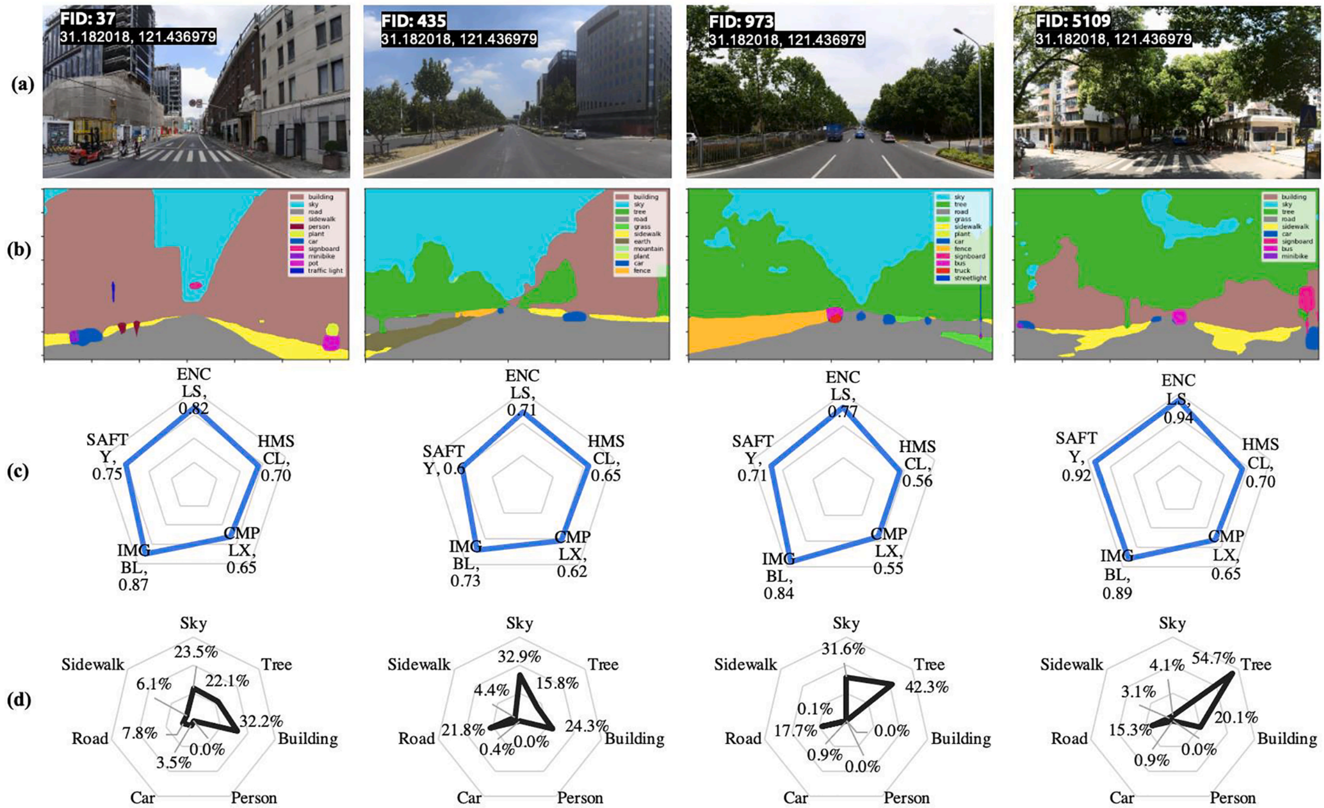


Fig. 7. Samples of (**a**) origin SVI, (**b**) semantic segmentation result, (**c**) predicted subjective perception scores, and (**d**) extracted objective view indices. The radar charts present scores and view indices from 0 to 1 inside out.

apartment. When generating the weight matrix to conduct spatial analysis, the potential spatial overlap issue should be avoided. Therefore, we run all models with a random subset comprising 5000 data points. The same analysis is also conducted to other nine subsets, and results are compared to validate our result is robust and not data dependent (Kim & Carruthers, 2015).

4. Results

4.1. Verifying View Indices and Perceptual Scores

Fig. 7 depicts the cross-validation result. First, segmentation results seemed accurate, with sky, tree, and building and road view indices varied. For example, we can see that the first three images presented similar sky views while the first two images present a similar portion of roads. These visual impressions were reflected correctly from Fig. 7(b and d) quantitatively.

Second, both average scores and score variances are larger with perceived ‘enclosure’ and ‘safety’ (Fig. 7c), being consistent with Fig. 4c. **‘Enclosure’ and ‘safety’ are less ambiguous concepts, indicating people tend to have a stronger preference for more straightforward design qualities.**

4.2. Objective View Indices

In total 36 visual elements appeared in sample SVIs, with large differences in their quantities. According to feature engineering (Jain & Chandrasekaran, 1982) and urban perception theory (Ewing & Handy, 2009; Zhang, 2018), not all elements are significant to perceptual scores. **On one hand, certain features with ubiquitous existence, such as sky, tree, and road, directly and significantly affect human perceptions like ‘enclosure’ according to definition** (Ewing & Handy, 2009). Therefore, all the ten most vast existed elements in Table 2a are indicated by literature significantly affecting perceptions, within which sky, tree and building view have been tested by prior housing price studies (Chen, 2020; Fu, 2019; Ye, 2019). **On the other hand, less ubiquitous elements like person and signboard, are revealed to be important affecting perceptions.** This finding is consistent with the literature (Ewing & Handy, 2009). Therefore, the ten most important view indices were fitted into the HPM and their strengths of associations with housing prices were compared with that of five subjective scores.

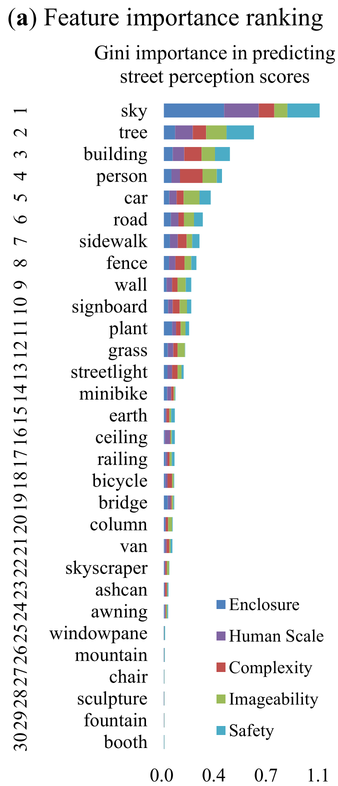
4.3. Measuring the Subjective Perceptions

4.3.1. Prediction Accuracy

The model specifications and performance on each perception score

Table 3
Performance of Random Forest (RF) predictions.

Perceptions	R2	MAE	RMSE	MAPE	Std. Dev.	Estimators (Bootstrap)	Min Sample Split (Min Leaf)	Max Feature (Max Depth)	Roy (2016) Criteria
Enclosure	0.61	0.09	0.11	0.33	0.06	180 (True)	5 (4)	Auto (80)	Moderate
Human Scale	0.51	0.12	0.14	0.38	0.07	60 (False)	10 (4)	Sqrt (60)	Bad
Complexity	0.49	0.13	0.16	0.41	0.08	130 (True)	5 (4)	Auto (80)	Bad
Imageability	0.47	0.15	0.19	0.43	0.10	180 (True)	5 (1)	Sqrt (80)	Bad
Safety	0.57	0.11	0.13	0.35	0.07	100 (True)	2 (2)	Sqrt (60)	Moderate



(b) Pearson correlations of five perceptual scores

	Enclosure	Human scale	Complexity	Imageability	Safety
Enclosure	1.00	0.76***	0.64***	0.76***	0.87***
Human scale	0.76***	1.00	0.82***	0.51***	0.54***
Complexity	0.64***	0.82***	1.00	0.34***	0.40***
Imageability	0.76***	0.51***	0.34***	1.00	0.82***
Safety	0.87***	0.54***	0.40***	0.82***	1.00***

Note: Pearson Coefficient (*** p<0.01).

(c) Pearson correlations of top 10 streetscape features

	Sky	Tree	Build.	Person	Car	Road	Sidewalk	Fence	Wall	Sign board
Sky	1.0	-0.5***	-0.6***	-0.3***	-0.3***	0.2***	-0.5***	0.1***	-0.1***	-0.2***
Tree	-0.5***	1.0	-0.3***	0.0***	0.2***	0.0***	0.3***	-0.2***	-0.2***	0.0***
Building	-0.6***	-0.3***	1.0	0.3***	0.2***	-0.3***	0.3***	-0.1***	0.0***	0.2***
Person	-0.3***	0.0***	0.3***	1.0	0.1***	-0.1***	0.1***	-0.1***	0.1***	0.2***
Car	-0.3***	0.2***	0.2***	0.1***	1.0	0.0***	0.0***	-0.2***	-0.2***	0.0***
Road	0.2***	0.0***	-0.3***	-0.1***	0.0***	1.0	-0.1***	-0.1***	-0.2***	-0.1***
Sidewalk	-0.5***	0.3***	0.3***	0.1***	0.0***	-0.1***	1.0	-0.1***	-0.1***	0.2***
Fence	0.1***	-0.2***	-0.1***	-0.1***	-0.2***	-0.1***	-0.1***	1.0	0.1***	0.1***
Wall	-0.1***	-0.2***	0.0***	0.1***	-0.2***	-0.2***	-0.1***	0.1***	1.0	0.0***
Signboard	-0.2***	0.0***	0.2***	0.2***	0.0***	-0.1***	0.2***	0.1***	0.0	1.0

Note: Pearson Coefficient (*** p<0.01).

Interpretation: 0.0 to ±0.3: negligible correlation; ±0.3 to ±0.5: low; ±0.5 to ±0.7: moderate; ±0.7 to ±0.9: high; and ±0.9 to ±1.0: very high

Fig. 8. Gini importance and correlation analysis **(a)** Important features in predicting five subjective perception scores. **(b)** Pearson correlation analysis of subjective perception scores. **(c)** Pearson correlation analysis of top 10 important visual elements extracted from SVIs.

are reported in Table 3. The MAE for predicting **'enclosure' and 'safety' were the smallest, revealing that these two concepts might be more straightforward with raters** (Zhang, 2018) therefore the smaller variance in rating led to higher accuracy in prediction. This is somehow consistent with Yao (2019) where beautiful and boring got the lowest prediction accuracy due to the largest variance in survey scores.

In general, achieved accuracy was reasonable given our small training sample size. First, with R^2 ranging between 0.47 and 0.61, three out of five perceptions explained more than half of the variance, indicating a significant improvement from Ewing and Handy (2009) and Park (2019) where the variance explained ranged from 0.21 to 0.37. Second, with MAEs varying between 0.09 and 0.15 for a 0–1 scale score system, the interpretation of qualities will not be offset if we convert 0–1 to a quartile system where the interval is 0.25. Third, using Roy (2016) framework, models of enclosure and safety were moderate, while the rest three were considered bad.

Although the Roy (2016) standard was not met, with cross-study comparison, the R2 of our model is higher than or close to that of Naik et al. (2014) and Yao (2021), and much better than that of Ito and Biljecki (2021) who also locally collected own perception training samples, while relatively lower than those obtained by Dubey et al. (2016), and Verma et al. (2020). The MAE is relatively higher than the one obtains by Yao (2021).

4.3.2. Variances in Feature Importance

Although sky, tree, building, person, and car ranked highest regarding sum Gini importance, there are large variances for each perception (Fig. 8a). For example, while the appearance of person pixel is most important in predicting complexity, it is less relevant with enclosure and human-scale prediction. While sky view prevails in explaining enclosure and human scale, its explanatory strength is weaker for imageability and complexity. **Surprisingly, several elements proved to be important in perceived safety** (Jacobs, 1992; Jansson, 2019; Zhang, 2018)**, such as streetlights, windowpanes, and street furniture were not ranked top ten. There less ubiquitous existence in SVI samples might be a reason, which indicates future study areas to systematically revisit the associations between streetscape features and human perception with a larger sample size of subjective data.**

4.3.3. Correlation Analysis

While Ewing and Handy (2009) found positive correlations between 'enclosure' and 'complexity', in our study, all perceptions exhibit non-negligible and positive correlations (Fig. 7b), especially for "enclosure-safety" (0.87), "human scale-complexity" (0.82), and "imageability-safety" (0.82). The strong correlations could be explained by the overlaps in their definitions. For example, enclosure and human scale both positively relate to street elements defining the vertical scenes like trees and buildings (Ewing & Handy, 2009). Notably, correlations between "complexity-imageability" and "complexity-safety" were low, indicating less coherence with their connotations. **The strong correlations suggest more theoretical efforts to differentiate definitions of highly correlated perceptions. Otherwise, the use of subjective measures would be restricted by the ambiguities in the definition.**

Most correlations between view index pairs were negligible to low, while "sky" has a moderate and negative correlation with "building", "tree", and "sidewalk". A VIF test in the HPM step also confirms the multicollinearity issue between the view index of sky and three ubiquitous existed elements. Therefore, in the final regression models, we removed "sky". This also suggests that **prior studies where sky, tree and building view indexes stand together** (Chen, 2020; Fu, 2019) **might result in biased estimations.**

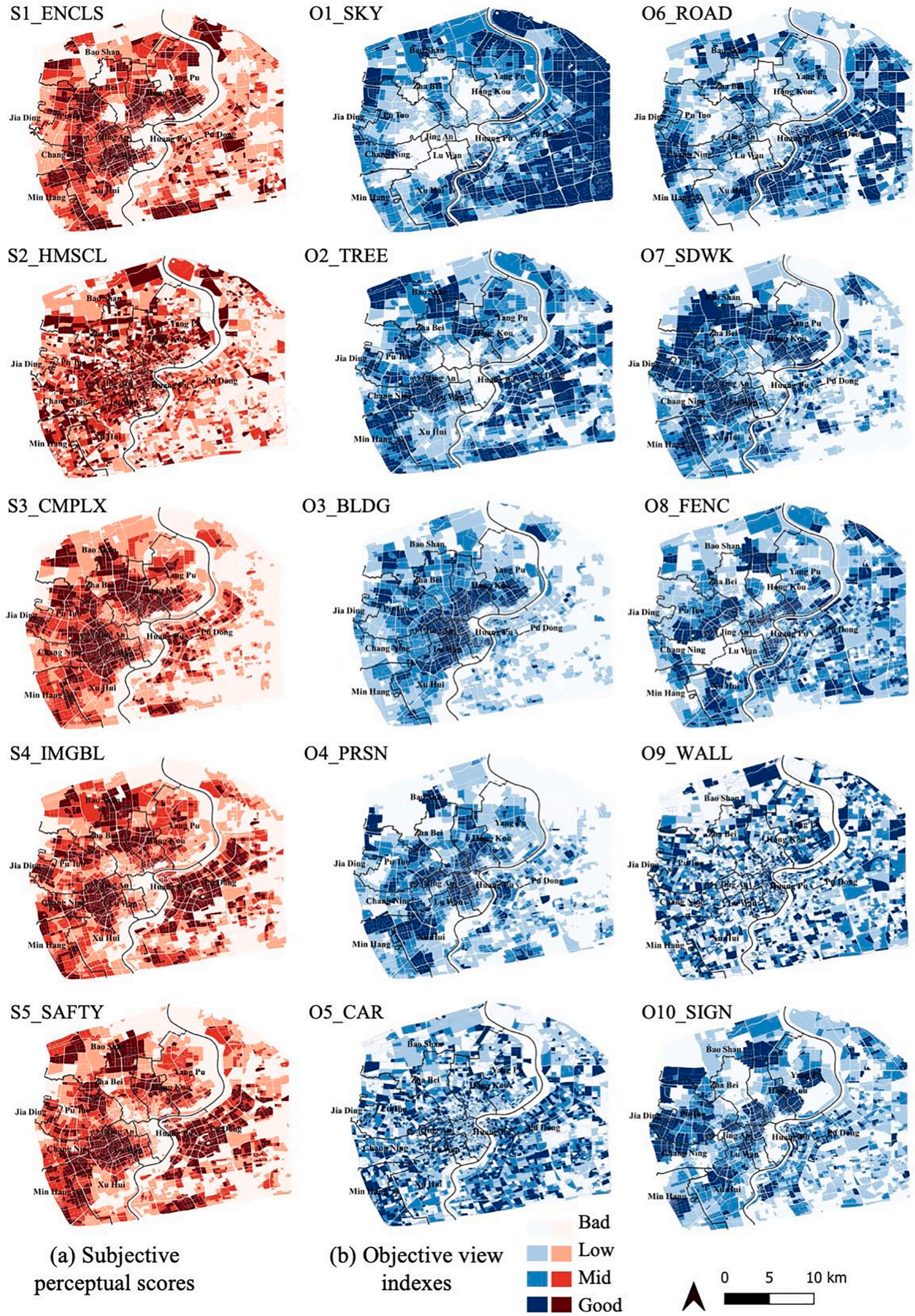


Fig. 9. Spatial distribution of (a) perception scores and (b) view indexes.

Table 4
Model performance by attribute groups.

OLS Diagnosis	Structural Attributes	Location Attributes	Neighborhood Attributes	Subjective Perceptual Scores	Objective View Indices
Adjusted R2	0.188***	0.678***	0.556***	0.275***	0.382***
Pr. (F-statistic)	0.000 ***	0.000 ***	0.000 ***	0.000 ***	0.000 ***
Moran's I on residuals (z-value)	0.65*** (175)	0.42***(121)	0.54***(147)	0.64***(174)	0.62***(173)
Robust LM (lag)	3082.77***	71.04***	722.97***	262.62***	996.49***
Robust LM (error)	4668.86***	1063.75***	3045.86***	2323.38***	8673.01***

Notes: p value *<0.1, ** $p < 0.05$, *** $p < 0.01$; z values of Moran's I test are shown in parentheses.

4.3.4. Spatial Heterogeneity of Street Design Qualities

Fig. 9 maps the spatial distribution of subjective and objective measures. In general, there is a clear spatial clustering tendency. For the subjective attributes, the west band of Huangpu River has higher scores than the east band (Pudong District), reflecting the fact that the core areas of the west band are well established in terms of the urban amenities, architecture style, urban density while Pudong is relatively new, and the development is more dispersed. For the score of enclosure, complexity, imaginability and safety, the distribution of higher subjective scores is concentrated in sub-city centers in Luwan, Zhabei, Jing'an, Hongkou and the CBD of Pudong district (i.e., Lujiazui). The higher scores of "human scale" distribute more dispersed, especially in Xuhui, Minhang and Pudong. Notably, compared to Minhang and Pudong, Xuhui has a much longer history in urban development, and has been the old commercial center of Shanghai. Xuhui's dispersed perception on the human scale requires further investigation.

For the objective attributes, the distribution of ratio of the sky, building and person generally follows the density of the urban development: city center has lower sky ratio, higher building ratio and more person; while suburban area and Pudong district has higher sky ratio, lower building ratio and less person. This is consistent with prior studies (Qiu et al., 2021, 2022; Ye, 2019). East bank of Huang Pu River has a higher ratio of road, while West Bank shows a higher ratio of sidewalks and signs. The distribution of cars, fences and walls does not show a clear pattern.

Additionally, the intercorrelations between subjective perceptions, and the relationships between perception and objective indexes hold in spatial. For example, the spatial distribution of enclosure, complexity, and safety scores hold in spatial, their high score and low score regions were highly overlapped. Further investigation can be conducted using bivariate Moran's I test, which we leave to future studies. Similarly, the pairwise maps of enclosure, sky and building reflect that enclosure is positively related to building view while negatively correlated to the sky view, a relationship we discussed in previous sections.

4.4. Spatial Hedonic Model Results

4.4.1. Clustering of Housing Price

The spatial distribution of housing price by original apartment units (Fig. 6(a)) and averaged price at the municipal neighborhood level (i.e., JuWeiHui) (Fig. 6(c)) indicated the price has a strong negative relationship with the distance to the city center. Further, the Moran's I statistic of the housing prices is 0.872 and significant at the 0.001 level, indicating the prices have strong and positive spatial autocorrelation. In other words, expensive apartments are highly likely to cluster, so as the cheap ones. Fig. 6(d) is the cluster map. The high and low-price clusters reflect the important role of distance to CBD and the Ring Road hierarchies in Shanghai. Expensive housing clusters are highly concentrated in the Inner Ring Road, while all low-price units are clustered outside the Outer Ring Road.

4.4.2. Strength of Association by Attribute Groups

Table 4 shows that five attribute groups' explanatory power using R^2 ranked as: location (0.678) > neighborhood (0.556) > objective streetscape attributes (0.322) > subjective streetscape scores (0.275) > structural attributes (0.188). All five attribute groups passed the F-statistic test ($p < 0.01$), **confirming the significant role that each group plays in affecting housing prices.**

Additionally, Moran's I confirm the strong spatial autocorrelation in OLS residuals. The results from the robust Lagrange multiplier test (Table 4) confirm that spatial lag and error processes both present (Anselin, 1988a), suggesting the use of spatial regression and GWR.

4.4.3. OLS Results

First, all structural, location and neighborhood attributes were fitted into an OLS model. Insignificant variables were removed. In addition, continuous variables with high VIF (>10) but lower Gini scores were removed. For example, distance to school/metro was correlated with accessibility measures. However, their Gini scores were smaller, therefore were removed. That said, the ease of reaching various services is more important than the vicinity to a specific service. The baseline model explains 78.3% price variance.

We then added five perceptual scores (Model 1) and ten view indices (Model 2) to the baseline model, respectively. For Model 2, sky view had VIFs larger than 10, was therefore removed. After processing, all three models and their variables were significant, with most VIFs below ten, indicating no evidence of strong multicollinearity.

Table 5 Part A reports the important regression diagnosis indicators such as R2 and the Moran's I on regression residuals for all models. First, Model-1 (R2 = 0.744) and Model-2 (R2 = 0.752) improved R^2 by 0.005 and 0.013 from the baseline (0.739), **indicating the effectiveness of incorporating subjective and objective street measures to improve the prediction ability.** Second, objective measures show slightly higher strength over subjective counterparts, which is consistent with health and walkability studies (Lin & Moudon, 2010). Third, results from Moran's I on OLS residuals and the robust LM test both confirmed the

Table 5A
Regression results and diagnosis. **Part A.** Regression performance and diagnosis for all OLS, SAC and GWR models.

	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Attributes	Baseline	Subjective	Objective	Subjective	Objective	Subjective	Objective
Method	OLS	OLS	OLS	SAC	SAC	GWR	GWR
Adjusted R^2 (Pseudo R^2)	0.739	0.744	0.750	(0.791)	(0.739)	0.897	0.905
Moran's I on Residual (p-value)	0.2034***	0.1857***	0.1678***	−0.0021 (0.1)	0.0007 (0.59)	/	/
Robust LM (lag)	283.859***	428.719***	4038.061***				
Robust LM (error)	16589.358***	13433.823***	14780.848***				

Note: p values are shown in parentheses; *<0.1, **<0.05, ***<0.01.

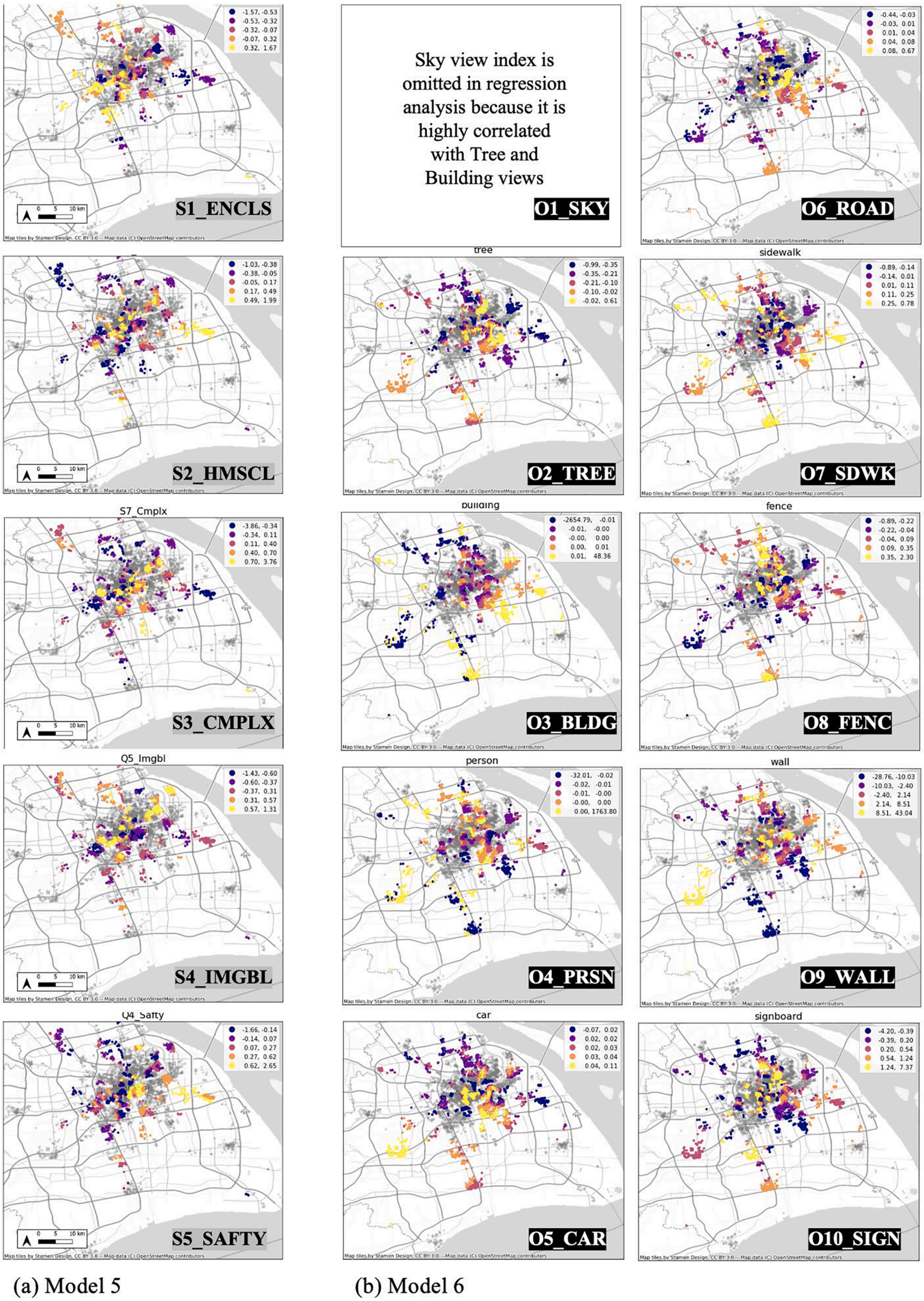
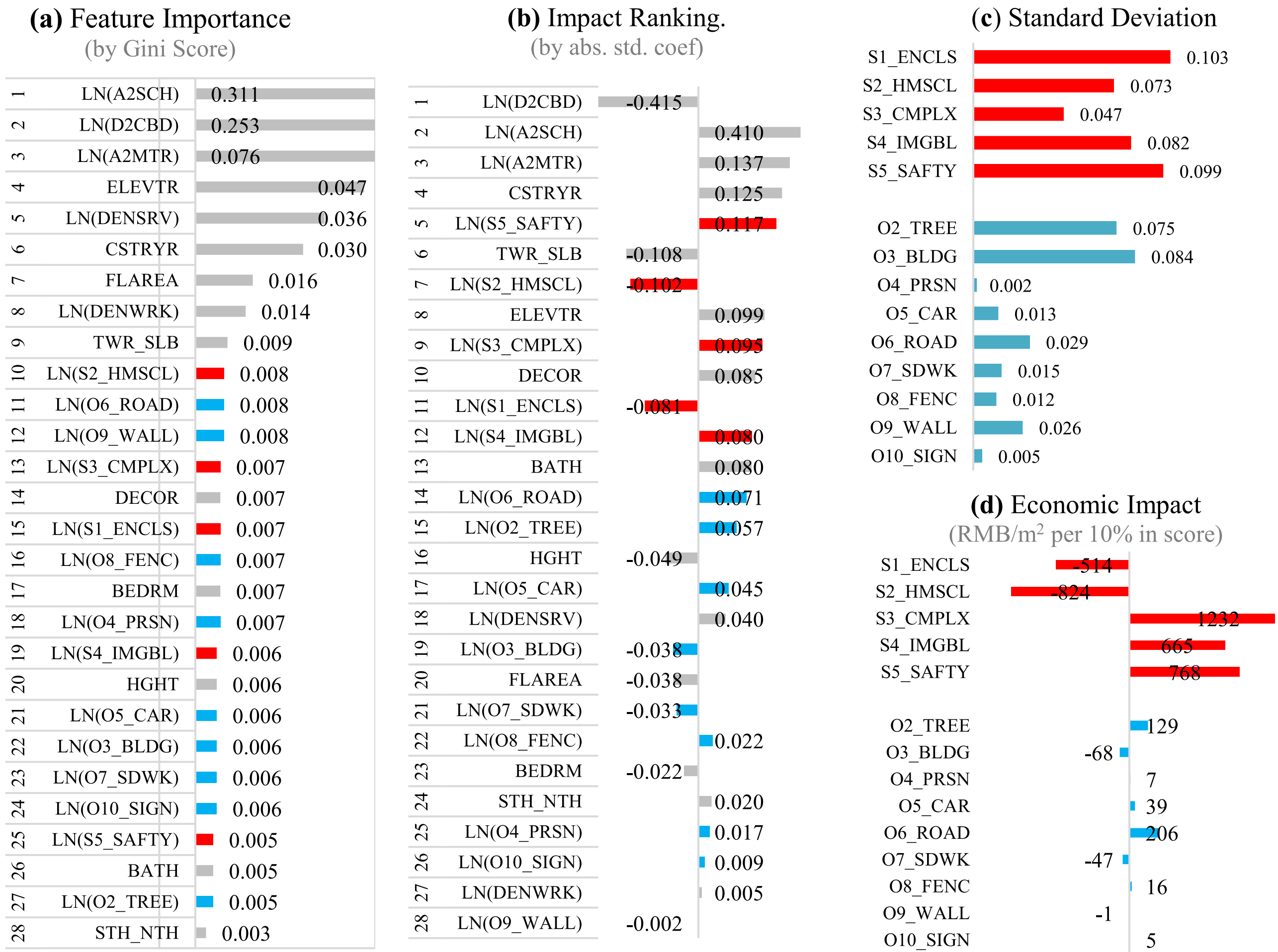


Fig. 10. Local coefficients of (**a**) Model 5 (Subjective Streetscape) and (**b**) Model 6 (Objective Streetscape).



Note: the colors are to differentiate subjective street variables (red), objective street variable (blue), and other variables (grey) in the regression analysis.

Fig. 11. Comparing variables' importance and effect size by (**a**) Gini importance score, (**b**) Absolute standardized coefficient, (**c**) Subjective and objective measures' standard deviation (without logarithm transformation), and (**d**) Impact on the selling price (RMB/m^2) if score increases by 10%.

spatial lag and error process present. Therefore, the parameter estimates, and significance of variables from OLS estimates are not robust. We move on to the results from spatial regression and GWR. The full regression result model is in Appendix Table A1.

4.4.4. *Comparing OLS and Spatial Regression Results*

First, when the spatial lag and spatial error interaction effects are considered (i.e., SAC type of model), Model-3 (R2 = 0.791) and Model-4 (R2 = 0.740) both yield higher goodness-of-fit compared to their OLS counterparts and both the lag and error terms are significant (Table 5 Part A), which is not surprising (Kim & Carruthers, 2015; Pandit et al., 2014). Second, the Moran's I on residuals become insignificant for both Model-3 and Model-4, indicating that by simultaneously modeling the spatial lag and error terms, SAC models effectively capture the autocorrelation effects. That said, the price of apartments in Shanghai are affected by their nearby neighbors. Third, for most variables, adding spatial interaction do not change the sign, the magnitude, and the significance in parameter estimates (Appendix Table A1), with few exceptions such as the trees, sidewalks, and signs, suggesting that OLS might result in mildly biased parameter estimation for these few features.

Regarding the streetscape attributes, all the subjective perceptions were consistently significant except for the imageability (Table A1), while all objective indicators were significant in the spatial regression. This indicated that OLS might falsely reject the null hypothesis, which is consistent with Part et al. (2019) where they found imageability changed to insignificant with walkability in the spatial model. The reason behind calls for a further investigation for future studies. In our study, the prediction accuracy for imageability is the lowest, therefore, the changing significance could be a result of poorer training accuracy, or the lower agreements on this perception within survey raters.

4.4.5. *GWR Results and Spatial Non-Stationarity*

For GWR models, the Gaussian kernel is used to distinguish between global and local effects. The best bandwidth is estimated at 189 nearest neighborhoods based on the data and is stable in different subsets. Fig. A1 (see Appendix) illustrates how the spatial distance of 189 nearest neighbors could variate with the data sample. In general, GWR models yield much higher adjusted R-Squared (Table 5) for the subjective (0.897) and objective (0.905) models, respectively, indicating that modeling non-stationarity would largely improve the predictive power. However, like previous studies, our GWR results are associated with extreme coefficients (Huang, 2017), and still present the dependence between spatial errors. Meanwhile, the significance, sign, and magnitude of variables from GWR results are more consistent with their OLS counterparts, except for signboard view index becomes insignificant (Appendix Table A1). This might be explained by the fact that the pixel of sign view is the least ubiquitous element identified on photos due to

Table 5B
Regression results and diagnosis. **Part B.** Selected regression model for comparison.

Section I. Regression Results	Model 0 Baseline		Model 1 Subjective		Model 2 Objective		Section II. Multilinearity and Feature Importance				Section III. Interpretation (Based on unstandardized coef.)		
Method	OLS		OLS		OLS								
Variables	Coef.	P > t	Coef.	P > t	Coef.	P > t	VIF	Gini Score	Std. Coef.	Ranking (Std. Coef.)	Delta X	% Price change	RMB/m2 change
CONSTANT	0.4041		0.4298		0.6830	***							
Structural Attributes													
FLAREA	−0.0001	***	−0.0001	**	−0.0001	**	5.5	0.016	−0.038	20	1 unit	−0.01	-¥8
BEDRM	−0.0040	*	−0.0037		−0.0047	**	2.9	0.007	−0.022	23	1 unit	−0.42	-¥240
BATH	0.0244	***	0.0238	***	0.0243	***	3.1	0.005	0.080	13	1 unit	2.40	¥1,378
CSTRYR	0.0022	***	0.0022	***	0.0021	***	2.3	0.030	0.125	4	1 unit	0.22	¥124
ELEVTR	0.0350	***	0.0348	***	0.0338	***	3.7	0.047	0.099	8	T/F	3.43	¥1,969
HGHT	−0.0166	***	−0.0172	***	−0.0177	***	1.1	0.006	−0.049	16	T/F	−1.75	-¥1,001
TWR_SLB	−0.0670	***	−0.0650	***	−0.0645	***	1.3	0.009	−0.108	6	T/F	−6.48	-¥3,715
STH_NTH	0.0084	***	0.0085	***	0.0085	***	1	0.003	0.020	24	T/F	0.85	¥488
DECOR	0.0295	***	0.0293	***	0.0289	***	1.1	0.007	0.085	10	T/F	2.91	¥1,666
Location Attributes													
LN(D2CBD)	−0.1143	***	−0.1102	***	−0.1136	***	3.7	0.253	−0.415	1	10%	−1.12	-¥642
LN(DENWRK)	0.0008	**	0.0006	**	0.0005	*	1.2	0.014	0.005	27	10%	0.01	¥3
LN(DENSRV)	0.0047	***	0.0048	***	0.0043	***	2	0.036	0.040	18	10%	0.05	¥26
LN(A2MTR)	0.0216	***	0.0234	***	0.0222	***	2.2	0.076	0.137	3	10%	0.23	¥131
LN(A2SCH)	0.0646	***	0.0657	***	0.0670	***	5	0.311	0.410	2	10%	0.66	¥380
Subjective Streetscape Attributes													
LN(S1_ENCLS)	/		−0.0896	***	/		9.3	0.007	−0.081	11	10%	−0.90	-¥514
LN(S2_HMSCL)	/		−0.1438	***	/		5	0.008	−0.102	7	10%	−1.44	-¥824
LN(S3_CMPLX)	/		0.2147	***	/		3.8	0.007	0.095	9	10%	2.15	¥1,232
LN(S4_IMGBL)	/		0.1160	***	/		3.5	0.006	0.080	12	10%	1.16	¥665
LN(S5_SAFTY)	/		0.1339	***	/		6.6	0.005	0.117	5	10%	1.34	¥768
Objective Streetscape Attributes													
LN(O1_SKY)	/		/		/		/	/	/	/	/	/	/
LN(O2_TREE)	/		/		0.0225	***	2.1	0.005	0.057	15	10%	0.23	¥129
LN(O3_BLDG)	/		/		−0.0118	***	3.4	0.006	−0.038	19	10%	−0.12	-¥68
LN(O4_PRSN)	/		/		0.0012	***	1.3	0.007	0.017	25	10%	0.01	¥7
LN(O5_CAR)	/		/		0.0068	***	1.3	0.006	0.045	17	10%	0.07	¥39
LN(O6_ROAD)	/		/		0.0360	***	2	0.008	0.071	14	10%	0.36	¥206
LN(O7_SDWK)	/		/		−0.0083	***	1.5	0.006	−0.033	21	10%	−0.08	-¥47
LN(O8_FENC)	/		/		0.0028	***	1.3	0.007	0.022	22	10%	0.03	¥16
LN(O9_WALL)	/		/		−0.0002	***	1.4	0.008	−0.002	28	10%	0.00	-¥1
LN(O10_SIGN)	/		/		0.0009	***	1.2	0.006	0.009	26	10%	0.01	¥5

Note: p values are shown in parentheses; ***, **, and * indicate significance level of 1%, 5% and 10%, respectively.

their small sizes and less prevalence in city spaces.

Second, the local adjusted R-Squared ranges from 0.41 to 0.88, and 0.43–0.90 for subjective (Model-5) and objective (Model-6) models respectively throughout Shanghai (Appendix Fig. A2). Notably, areas with higher R2 in both models are consistent: the regions between the inner and mid-ring roads, as well as the regions outside the outer-ring roads. This can be explained by the fact that these regions have either high or low-price clusters (Fig. 6d). Our finding is consistent with Huang (2017) where housing prices in downtown areas such as Jing'an and Huangpu or in the suburban areas such as Songjiang and Jiading are well explained.

Third, the local coefficients of subjective and objective measures are shown in Fig. 10. Like the pattern of local R^2 (Fig. A2), we see the areas in downtown or suburban have distinct magnitude or even signs of the local coefficient. For example, enclosure is positively related to housing prices in West bund while the effect is negative in east bund (Pudong). Safety sees a reverse pattern compared to that of enclosure. These findings suggest that the non-stationary effect of subjective street perceptions, and objective streetscape elements are both strong. Therefore, their implications on urban designers and housing policy should be carefully discussed and tailored from region to region for Shanghai.

5. Discussion

Since our focus is not the differences between OLS and spatial models, given the fact that the sign, magnitude, and significance of OLS coefficients are mostly consistent with those of SAC and GWR models (Appendix Table A1), we would discuss the implications of findings based on OLS results for simplicity (Section B of Table 5). The interpretation of both SAC and GWR models would just go beyond the scope and length of this paper, as SAC models capture both direct and indirect effects, while GWR needs extra cares for local effects, both of which would need specific interpretation.

5.1. Effects of Subjective and Objective Streetscape Attributes

For the streetscape attributes, subjective and objective measures indicate a large variance in strength. First, ranked by Gini importance (Fig. 11a), objective measures overall show stronger explanatory power than the subjective counterpart. Road, wall, fence, and person views were among the 20 most important variables, with 'road' ranked the highest within all subjective and objective streetscape attributes. This is consistent with prior discussion on Table 4 and Table 5 where OLS models of objective streetscape attributes collectively outperform subjective ones. **This finding indicates that besides sky, tree, and building, many other visual elements affect perceptions on street quality, ultimately having impacts on housing prices.**

Second, Table 5 Part B provides unstandardized interpretation of economic impacts, where a 10% score increase in the streetscape variables leads to divergent price change, largely due to the large variances in streetscape variables' standard deviations. For example, for less ubiquitous pixels like person and signboard, their magnitude is about ten times smaller (Fig. 11c). Therefore, Fig. 11(b) which provides the ranking by standardized coefficient is a better measure to compare the effectiveness of subjective and objective indicators. Ranked by "standardized coefficient" (Fig. 11b), however, subjective measures exhibit stronger overall strength. This is also consistent with Gini importance analysis (Fig. 11a) where 'human scale', 'complexity' and 'enclosure' were more important than many other objective indicators except for road and wall view. Safety ranked the highest within all streetscape indicators, with human scale, and complexity among the top-10 determinants, while most objective view indexes were out of top-15. **This finding suggests subjective perceptions on street design quality exhibit stronger strength individually than objective streetscape elements. Moreover, less addressed qualities such as safety, human scale and complexity are more important than the perceived enclosure** (Buonanno, Montolio and Raya-vílchez, 2013; Yin & Wang, 2016)**. This suggests that subjective perceptions have not been addressed adequately and more research efforts could be invested in this regard.**

Third, to intuitively illustrates the effect of street design quality, Fig. 11d provides the monetized impact if the score increases by 10%, using the average price of data as the base (57,349RMB/m^2). Only four street design variables showed negative signs: perceived "human scale" and "enclosure" from subjective attributes, and "building" and "sidewalk" views from objective indicators. It is reasonable because the four indicators all represent or contribute to scenes that opposite "openness" – a feature which is favorable by many homebuyers in China (Chen, 2020; Fu, 2019; Ma, 2021). More specifically, with 10% increase in the score, all five subjective measures, namely complexity (+1232RMB/m^2), human scale (-824RMB/m^2), safety (+768RMB/m^2), imageability (+665RMB/m^2), and enclosure (-514RMB/m^2) are associated with more than 500RMB/m^2 change in housing price. While for objective measures, only road (206RMB/m^2) and tree (129RMB/m^2) view showed similar scale of strength. That said, a 1% increase in treen view was related to 12.9RMB/m^2 higher price, an effect smaller than Chen (2020) where a 1% increase in green view index increased housing prices by 71 RMB/m^2. The smaller magnitude from this study is reasonable because prior studies only captured the effects of tree but ignored other important streetscape elements. **To conclude, housing prices changes attributed to subjective measures were larger in magnitude, indicating subjective scores might capture more information related to residence behaviors that cannot be completely explained by individual visual elements.**

5.2. Comparing Subjective and Objective Measures

Subjective perceptions individually exhibit stronger strength affecting housing prices than most view indices, largely due to its effectiveness in capturing factors related to psychological and social-demographical characteristics of street users. Meanwhile, when rating street scenes, viewers often notice more subtle elements, as well as the nonlinear relationships between elements. Therefore, perceptions capture more sensory information beyond individual indexes. The relative composition of visual elements, and subtle textures like earth, permeable pavements, paints all affect how people perceive a street environment. **Therefore, for comprehensive perceptions like imageability, subjective measures explain human behaviors more completely. However, due to the same reason, it will be more effective and accurate to predicting a straightforward concept based on visual elements. For example, 'enclosure' and 'safety' achieved higher accuracy than other more ambiguous concepts like imageability.**

Objective view indexes still provide advantages to complement subjective measures. **First, perceptions are multifaced. It will be difficult to exhaust all perceptions that might affect residence behaviors and decisions. For example,** Ewing and Handy (2009) tried to quantify 51 perceived qualities but only succeeded with five. **However, visual elements' presence in SVIs is relatively stable, therefore view indexes can be explicitly modeled.** We extracted 36 streetscape features, most of which also have appeared in various SVI studies across geospatial spaces (Chen, 2020; Ma, 2021; Zhang, 2018). Therefore, while we did not incorporate perceived "greenness" which has been an important dimension (Li, 2015; Ma, 2021), the tree view index became an ideal alternative to supplement the missing determinants from greenery. **Second, objective view indices indicated negligible or low correlations, while subjective perceptions exhibited serious multicollinearity. This is because the definition of** many perceptual qualities is far from mutually exclusive nor collectively exhaustive (Ewing & Handy, 2009). For example, **we identified strong correlations between "enclosure", "human scale", "imageability" and "safety".** Therefore, when more human perceptions must be incorporated, the correlation issues could be much stronger. More advanced statistic methods such as the principal component analysis should be applied to reduce multicollinearity, which in turn limits the application of subjective measures because technical requirements increase. It could be even more difficult to collect ten or twenty types of perceptual qualities consistently and accurately with randomly selected raters. **Therefore, certain view indexes or the recombination of them could complement subjective measures can reduce multicollinearity issues, while still fulfill missing information left behind by avoiding certain perceptual qualities.**

Therefore, when choosing between subjective and objective measurements, decisions should be made based on how familiar daily street users (raters) are with definitions of the subjective perceptions, and whether serious multicollinearity issues exist between the pairs of perceptions. Objective view indexes collectively might outperform the subjective perceptions, while they individually could complement subjective perceptions.

5.3. Effects of Conventional Attributes

First, within location attributes, the Gini score of distance to CBD (D2CBD) ranked the second within all variables, indicating its significant role in explaining housing prices, which is consistent with the literature (Kim & Carruthers, 2015). Housing prices decreased by 1.10% when the distance to CBD increased by 10% (Fig. 10b). We interpret it as a centrality variable that captures the level of services and opportunities, indicating living costs such as commuting. Second, all four neighborhood attributes were among the ten most influential factors, which is consistent with the literature (Wen et al., 2017; Zhang & Dong, 2018). Accessibility to high school (A2SCH), accessibility to metro (A2MTR), services density (DENSRV) and job density (DENWRK) positively affected prices, while the first three ranked the first, third and fifth by Gini importance. We found remarkable price premiums for school districts and the metro. With 10% more 'good' high schools and metro stations reachable within one- and five-kilometer radius, prices increased by 380 RMB/m^2 and 131 RMB/m^2, respectively. Third, the signs of structural attributes were consistent with the literature (Wen et al., 2017): units were sold 2.9% (or 1666 RMB/m^2), 0.85% (488 RMB/m^2), and 3.5% (1969 RMB/m^2) more expensive with the refined interior design, having south-facing room, and with elevators respectively.

5.4. Implications for Urban Planning

Our research provides insights into measuring subjective perceptions and objective view indexes using intelligence analytical tools to inform housing price studies. It has broader and practical applications in urban planning. It provides useful references to real estate developers, policymakers, researchers, and planners. First, our study revealed that the economic value of the micro-scale street environment and perceptions had been understated. Real estate developers should also cater for the various micro-environment in the public realm, such as the greenery, street furniture, and retails along the street interface where pedestrians and residents can perceive around their residential blocks within 15 min walking distance. Currently most developers pay attention to only the tree canopy or plantation coverage inside residential blocks. Secondly, our study helps policymakers to better make decisions on urban policy to adapt to various public facilities and citizen needs, which may contribute to sustainable urban planning and housing development. For example, **real estate developers obtain price benefits from the surrounding street environment, while the cities take care of designing, investing, and maintaining streetscapes.** A street environmental tax could be levied to compensate public budgets put in facilitating better street scenes (Ye, 2019; Zhang & Dong, 2018). The tax amount could be informed based on both subjective scores and objective view indexes. Thirdly, the result of this study supports the theories connecting the physical setting and subjective perceptions with the social-economic variables. Data-driven methods and advanced intelligence tools can help researchers understand human perceptions and emotions to better model the relationship between humans and the environment on a global level and a neighborhood scale. It also helps researchers to choose indicators when studying the underlying urban heterogeneity patterns. The subjective perceptions could be integrated into economics modeling and other urban agendas beyond settlement evaluation. Finally, our results reveal the relationship between micro-level view indexes and human perception. The subjective scores provide a new metric for street design guidelines (Ma, 2021). Urban designers and planners can examine the individual perception of the street environment, and better guide various applications such as sustainable public transportation infrastructure planning, urban micro renovation, lively and safe neighborhood design.

5.5. Limitations and Future Studies

There are several limitations regarding (1) the desirable causal inference and modeling procedure, (2) the reliability of the online visual survey, (3) the prediction accuracy and (4) the validation of perception scores, to be improved for future studies.

First, several important improvements are desirable regarding the data and method in the HPM procedure. In this study, our intention is not to make any causal statement, but simply to use the correlation to justify the effort and value of incorporating extra micro-scale urban perception data. No causal statement can be made yet because there could be a reverse causal relationship or a confounding factor with both housing prices and street environment that was omitted, such as a policy that beautifying the district and invest in new urban infrastructures, which is common in China. However, the causality relationships between housing price and many variables can provide much more convincing policy recommendations to decision makers as well as more profound empirical results that will largely enrich the literature (Kang, 2020). With the effectiveness of opensource SVI data, and web datamining, future studies can plan for a more serious panel or pseudo panel data to investigate the causal relationship. Moreover, as our reviewer pointed out, the property price in Shanghai has increased a lot since 2019. However, the streetscape might not change too much. We cannot validate whether our specific findings in this dataset still hold when the housing price data in another year, say 2021, is applied. This also points to an important area to that a future study should carefully design the model procedure to validate the findings are stable. We should be cautious about any biased estimations in HPM that overestimate/underestimate the monetized value of urban public space, which in turn result in wasting/lacking public investments.

Moreover, the non-linear relationship between different housing price bands and in different urban function regions or demographic neighborhoods can be better compared with separated submarket models or using the spatial regimes method (Kim & Kim, 2020). Only a single city during a single timeframe was studied. It will be desirable to apply the seemingly scalable and applicable method to cross-sectional and longitudinal housing prices and street view data. Both a multiple cities cross-study and multiple year studies can provide important implications on the spatial and temporal heterogeneity of street perception changes and urban development shifts. Additionally, the type I error (false significance) could exist when there are many independent variables in a multivariate analysis. This could be the case for our objective streetscape model. More advanced statistics processes such as the Principal Component Analysis (PCA) can be applied to reduce the dimensions (Kang, 2021).

Second, due to timeframe and resource limitations, the perceptual scores were collected from a small-size study group with raters all coming from very specific study groups – architects, planners, and designers. For future studies on homebuyers' preferences, ratings from potential homebuyers would be more desirable and might provide interesting and very different implications. This could be done by randomly selecting people who visit real estate offices. Moreover, due to the design of the online survey platform, we could not identify votes and voters, therefore neither the inter- nor intra-rater reliability analysis was feasible. Although prior studies have validated that the inter-rater variance or the sociodemographic bias from raters (Salesses et al., 2013; Dubey et al., 2016) were not significant, the capability of conducting reliability analysis is an important component to be considered when designing and developing the crowdsourcing platform.

Third, when predicting perception scores from the image, we took a rule-based approach using only high-level features (i.e., streetscapes) to align with urban design measures (Ewing & Handy, 2009) to ensure interpretability. Nevertheless, we notice that the mainstream in image-based perception studies is to use both low- and high-level features (Ito & Biljecki, 2021; Rossetti, 2019; Verma et al., 2020). The generic features such as HSL histogram, Blob detection, and edge detection can complement high-level features and improve prediction accuracy. In addition, thanks to our reviewers' comment, a future study should consider an in-person field audit as a cross validation procedure to evaluate the reliability of scores predicted from ML models based on streetscape features extracted.

Last, both the subjective scores and objective view indexes were difficult to interpret, limiting our viability in providing more actionable policy recommendations in urban design and city planning. For example, what does it mean to the Floor Area Ratio (FAR) in the masterplans and zoning codes if the sky view decrease by 10% or the enclosure score increase by 10%? How will the street façade change if the complexity score increases by a 10%? Future studies can correlate conventional urban form and density metrics such as POI density, floor area ratio, street width and block size (Cervero & Kockelman, 1997; Qiu, 2021) with subjective scores and objective view indexes to provide more informative and actionable references to inform urban design guidelines that ultimately facilitate better streets environment.

6. Conclusion

While an increasing number of housing prices studies adopted CV frameworks and SVI data, they have been limited to objective measures only accounting for single or multiple visual elements' influences. We comprehensively compared the strengths of associations between subjective and objective measures of streetscape qualities on housing prices for Shanghai. Our framework integrated classical urban design measures (Ewing & Handy, 2009) with new urban analytical tools with AI, crowdsourcing, and opensource SVI to effectively inform how urban scenes affect housing prices. With the proposed framework, we collected and evaluated five subjective perceptions based on extracting visual elements presented in SVIs. We then discussed the complementary effects of using both subjective perceptions and objective view indexes to inform housing price studies. Our study addressed the necessity and effectiveness of incorporating subjectively measured street perceptions for housing prices literature, which is less discussed. Future studies on housing price should include important subjective perceptions, while carefully dealing with the strong multicollinearity within subjective measures. Our study provides a useful reference to policymakers, economists, and planners for choosing street environment indicators.

6.1. Significance of Subjective and Objective Measures

First, both objective view indexes and subjective measures explain more variance than conventional structural attributes, while objective measures collectively exhibit stronger strength over the subjective counterpart. Our findings suggest that developers obtain a price premium from surrounding street environments. A street environment tax can be levied to compensate public budgets used on street environments maintenance. Specifically, subjective measures of human scale, imageability and complexity, and objective indicators like road and tree view showed stronger explanatory power over conventional variables which have been indicated important (e.g., number of bathrooms, interior decoration quality, south-facing, and floor height). Meanwhile, most street quality variables were positively related to housing prices, except for "enclosure", "human scale", and "sidewalk". It could be resulted from the fact that these indicators were negatively related to "openness" - a quality preferred by home buyers in China (Chen, 2020; Fu, 2019; Ma, 2021).

6.2. Variable Selection and Correlations

Second, our findings provided important references in constructing subjective perceptions as well as objective view indexes selection. For example, less discussed perceptions like "human scale", "imageability" and "complexity" exhibited stronger strength of association with housing prices than the more widely-studied "safety" and "enclosure", indicating that more subtle perceptions could be investigated in future studies. However, less ambiguous perceptions (i.e., "safety" and "enclosure") also resulted in wider score variance, higher average scores, and more accurate predictions using ML.

For objective view indexes selection, less ubiquitous visual elements like person and fence exhibited stronger associations with human perceptions and housing prices than the most-studied features like sky, tree, and building. Furthermore, we found "sky" view exhibits serious multicollinearity with "building" and "tree", with a VIF value larger than 10. Therefore, prior studies might result in biased estimations where sky, tree and building views stand together to explain housing prices (Chen, 2020; Fu, 2019). While more comprehensive perceptions should be incorporated, and less ubiquitous and subtle visual elements should be tested, future studies must carefully deal with potential multicollinearity issues. In general, correlations between subjective perceptions were positive and strong, while that of view indexes were mostly negligible.

6.3. Complementary Effects between Two Measures

On the one hand, subjective perceptions showed larger standardized coefficients. They individually might represent more underlining mechanisms related to street users, capturing more unexplained built environment factors. However, strong correlations exist within subjective perceptions due to the ambiguity and overlap with their definitions, limiting the further incorporation of more perceptions. On the other hand, objective view indexes collectively explained more price variance, while having negligible or low correlations. The types of visual elements presented and the intensity of each element's presence across geospatial regions are relatively stable and consistent. Therefore, view indexes can complement subjective human perceptions while reducing multicollinearity with subjective measures.

When selecting objective indicators to complement or replace subjective measures, decisions should be made based on the definition of perceptual qualities, for example, how comprehensive a definition is, and how familiar is the definition to daily street users. More comprehensive perceptions exhibit stronger strength in explaining human behavior and housing prices, however, with stronger multicollinearity. Meanwhile, view indexes like "sky" and "tree" are found to effectively represent more self-evident perceptions such as "openness" and "greenness" (Ma, 2021; Yin & Wang, 2016). Therefore, less comprehensive perceptions could be substituted by selected view indexes to supplement missing information while reducing multicollinearity.

Institutional Review Board Statement:

Ethical review and approval were waived for this study due to the analyzed datasets are properly anonymized, no participant can be identified.

Informed Consent Statement:

Written informed consent was waived due to the analyzed dataset are properly anonymized, no participant can be identified.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to thank the participants who helped to rank SVI perception qualities in the Digital Future Workshop (2020) as well as the sponsor and organizer of the workshop- Prof. Philip YUAN and Tongji University.

Funding

This research received Kermit C. & Janice I. Parsons Scholarship (2019) from City and Regional Planning Department, Cornell University.

Appendix A

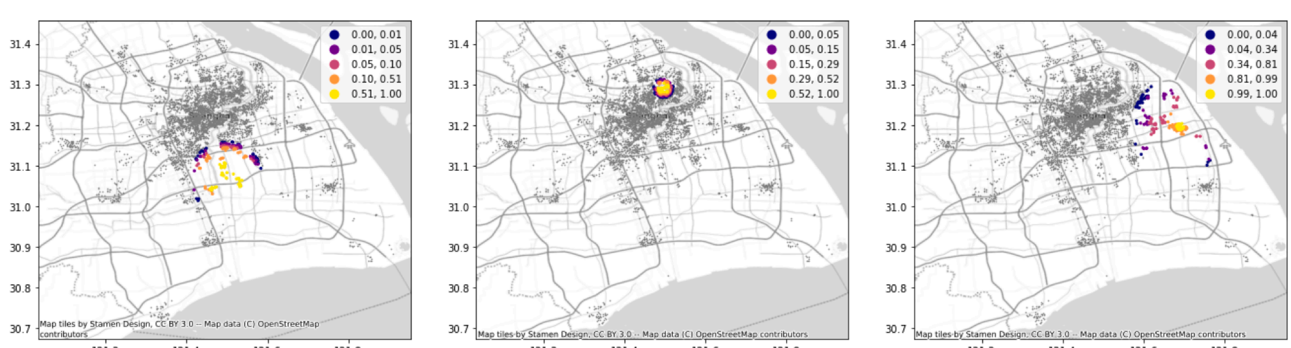


Fig. A1. The variations in the spatial scope for the best estimated bandwidth (k = 185).

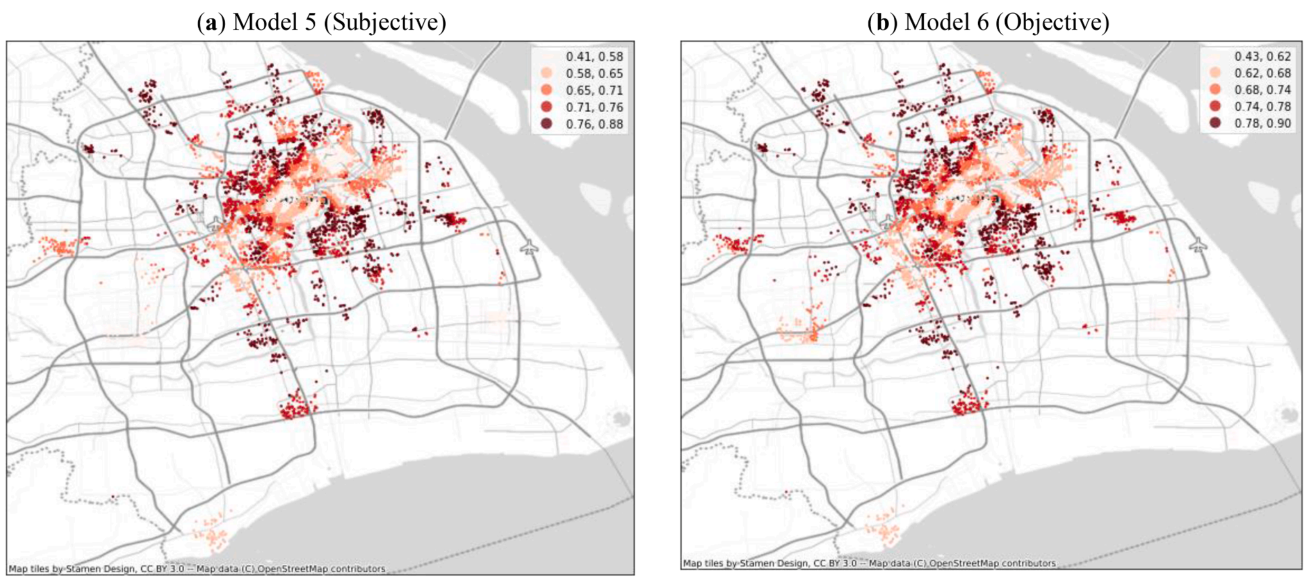


Fig. A2. Local R^2 of (**a**) Model 5 (Subjective Streetscape), (**b**) Model 6 (Objective Streetscape).

Table A1
Regression results and diagnosis.

	Model 0		Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	Baseline		Subjective		Objective		Subjective		Objective		Subjective		Objective	
Method	OLS		OLS		OLS		SAC		SAC		GWR		GWR	
Variables	Coef.	P > t	Coef.	P > t	Coef.	P > t	Coef.	P > t	Coef.	P > t	Coef.	P > t	Coef.	P > t
CONSTANT	0.4041		0.4298		0.6830	***	0.2555	***	1.0317	***	0.461	***	0.465	***
Structural Attributes														
FLAREA	−0.0001	***	−0.0001	**	−0.0001	**	0.0002	***	0.0004	***	0.0002	***	0.0002	***
BEDRM	−0.0040	*	−0.0037	**	−0.0047	**	−0.0053	***	−0.0018	***	−0.006	***	−0.006	***
BATH	0.0244	***	0.0238	***	0.0243	***	0.0252	***	−0.0214	***	0.027	***	0.028	***
CSTRYR	0.0022	***	0.0022	***	0.0021	***	−0.0002	**	−0.0010	***	0.002	***	0.002	***
ELEVTR	0.0350	***	0.0348	***	0.0338	***	−0.0310	***	0.0196	***	−0.112	***	0.039	***
HGHT	−0.0166	***	−0.0172	***	−0.0177	***	0.0003		0.0035		0.04	***	−0.016	***
TWR_SLB	−0.0670	***	−0.0650	***	−0.0645	***	−0.0163	***	−0.0755	***	−0.016	***	−0.072	***
STH_NTH	0.0084	***	0.0085	***	0.0085	***	−0.0124	***	0.0530	***	−0.072	***	0.002	*
DECOR	0.0295	***	0.0293	***	0.0289	***	0.0161	***	0.0579	***	0.002	*	0.028	***
Location Attributes														
LN(D2CBD)	−0.1143	***	−0.1102	***	−0.1136	***	−0.0008	**	0.0175	***	0.028	***	−0.108	***
LN(DENWRK)	0.0008	**	0.0006	**	0.0005	*	−0.0012	***	0.0116	***	0.0002		0.0002	
LN(DENSRV)	0.0047	***	0.0048	***	0.0043	***	−0.0027	***	−0.0129	***	0.004	***	0.002	**
LN(A2MTR)	0.0216	***	0.0234	***	0.0222	***	0.0031	***	0.0065	***	0.023	***	0.022	***
LN(A2SCH)	0.0646	***	0.0657	***	0.0670	***	0.0074	***	−0.0030	***	0.068	***	0.065	***
Subjective Streetscape Attributes														
LN(S1_ENCLS)	/		−0.0896	***	/		−0.0619	***	/		−0.206	***	/	/
LN(S2_HMSCL)	/		−0.1438	***	/		−0.1293	**	/		−0.249	***	/	/
LN(S3_CMPLX)	/		0.2147	***	/		0.1415	***	/		0.38	***	/	/
LN(S4_IMGBL)	/		0.1160	***	/		0.0286		/		0.24	***	/	/
LN(S5_SAFTY)	/		0.1339	***	/		0.1393	***	/		0.239	***	/	/
Objective Streetscape Attributes														
LN(O1_SKY)	/		/		/		/		/		/		/	
LN(O2_TREE)	/		/		0.0225	***	/		0.0455	***	/		0.0115	***
LN(O3_BLDG)	/		/		−0.0118	***	/		−0.0341	***	/		−0.007	*
LN(O4_PRSN)	/		/		0.0012	***	/		0.0035	***	/		0.001	***
LN(O5_CAR)	/		/		0.0068	***	/		0.0108	***	/		0.007	***
LN(O6_ROAD)	/		/		0.0360	***	/		0.0218	***	/		0.041	***
LN(O7_SDWK)	/		/		−0.0083	***	/		−0.0211	***	/		−0.006	***
LN(O8_FENC)	/		/		0.0028	***	/		0.0050	***	/		0.002	***
LN(O9_WALL)	/		/		−0.0002	***	/		−0.0012	***	/		0.001	***
LN(O10_SIGN)	/		/		0.0009	***	/		0.0105	***	/		0.009	***
Spatial Lag/Error														
LAG(LN(PRICE))			/				0.0201	***	0.0066	***				
LAMBDA			/				−0.2361	***	0.6432	***				
Regression Diagnosis														

(continued on next page)

Table A1 (*continued*)

	Model 0		Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	Baseline		Subjective		Objective		Subjective		Objective		Subjective		Objective	
Method	OLS		OLS		OLS		SAC		SAC		GWR		GWR	
Variables	Coef.	P > t	Coef.	P > t	Coef.	P > t	Coef.	P > t	Coef.	P > t	Coef.	P > t	Coef.	P > t
Adjusted R2 (Pseudo R2)	0.7393		0.7444		0.7502		(0.7907)		(0.7393)		0.897		0.905	
Moran's I on Residual (P value)	0.2034	***	0.1857	***	0.1678	***	−0.0021 (0.1276)		0.0007 (0.588)					
Robust LM (lag)	283.859	***	428.719	***	4038.061	***								
Robust LM (error)	16589.358	***	13433.823	***	14780.848	***								

Note: p-value in parentheses; ***, **, and * indicate significance level of 1%, 5% and 10%, respectively.

References

Anderson, S. T., & West, S. E. (2006). Open space, residential property values, and spatial context. *Regional Science and Urban Economics, 36*(6), 773–789. https://doi.org/10.1016/j.regsciurbeco.2006.03.007

Anselin, L. (2003). Spatial Econometrics. In *A Companion to Theoretical Econometrics* (pp. 310–330). John Wiley & Sons Ltd. https://doi.org/10.1002/9780470996249.ch15.

Anselin, L. (1988a) 'Model Selection in Spatial Econometric Models', in Anselin, L. (ed.) *Spatial Econometrics: Methods and Models*. Dordrecht: Springer Netherlands (Studies in Operational Regional Science), pp. 243–252. 10.1007/978-94-015-7799-1_14.

Anselin, L. (1988b) 'Model Validation and Specification Tests in Spatial Econometric Models', in Anselin, L. (ed.) *Spatial Econometrics: Methods and Models*. Dordrecht: Springer Netherlands (Studies in Operational Regional Science), pp. 223–242. 10.1007/978-94-015-7799-1_13.

Anselin, L. (1988c) 'Spatial Heterogeneity', in Anselin, L. (ed.) *Spatial Econometrics: Methods and Models*. Dordrecht: Springer Netherlands (Studies in Operational Regional Science), pp. 119–136. 10.1007/978-94-015-7799-1_9.

Beleites, C., et al. (2013). Sample size planning for classification models. *Analytica Chimica Acta, 760*, 25–33. https://doi.org/10.1016/j.aca.2012.11.007

Breiman, L. (2001). Random Forests. *Machine Learning, 45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Brownson, R. C., et al. (2009). Measuring the built environment for physical activity: State of the science. *American Journal of Preventive Medicine, 36*(4 Suppl), S99–123. e12. https://doi.org/10.1016/j.amepre.2009.01.005

Buonanno, P., Montolio, D. and Raya-vílchez, J.M. (2013) 'Housing prices and crime perception', *Empirical Economics*, 45(1), pp. 305–321. http://dx.doi.org.proxy.library.cornell.edu/10.1007/s00181-012-0624-y.

Cervero, R., & Kockelman, K. (1997). Travel demand and the 3Ds: Density, diversity, and design. *Transportation Research Part D: Transport and Environment, 2*(3), 199–219. https://doi.org/10.1016/S1361-9209(97)00009-6

Chen, L., et al. (2020). Measuring Impacts of Urban Environmental Elements on Housing Prices Based on Multisource Data—A Case Study of Shanghai, China. *ISPRS International Journal of Geo-Information, 9*(2), 106. https://doi.org/10.3390/ijgi9020106

Cleveland, W. S., & Devlin, S. J. (1988). Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting. *Journal of the American Statistical Association, 83*(403), 596–610. https://doi.org/10.1080/01621459.1988.10478639

Couclelis, H. (2009). Rethinking Time Geography in the Information Age. *Environment and Planning A: Economy and Space, 41*(7), 1556–1575. https://doi.org/10.1068/a4151

DigitalFUTURES (2020) *DigitalFUTURES WORLD: Manifesto*, *DigitalFUTURES.world*. Available at: https://www.digitalfutures.world/manifesto (Accessed: 16 February 2021).

Donovan, G. H., & Butry, D. T. (2010). Trees in the city: Valuing street trees in Portland, Oregon. *Landscape and Urban Planning, 94*(2), 77–83. https://doi.org/10.1016/j.landurbplan.2009.07.019

Dormann, F., C., et al. (2007). Methods to account for spatial autocorrelation in the analysis of species distributional data: A review. *Ecography, 30*(5), 609–628. https://doi.org/10.1111/j.2007.0906-7590.05171.x

Dubey, A. *et al.* (2016a) 'Deep Learning the City: Quantifying Urban Perception at a Global Scale', in Leibe, B. et al. (eds) *Computer Vision – ECCV 2016*. Cham: Springer International Publishing (Lecture Notes in Computer Science), pp. 196–212. 10.1007/978-3-319-46448-0_12.

Elhorst, J. P. (2010). Applied Spatial Econometrics: Raising the Bar. *Spatial Economic Analysis, 5*(1), 9–28. https://doi.org/10.1080/17421770903541772

Ewing, R., et al. (2006). Identifying and Measuring Urban Design Qualities Related to Walkability. *Journal of Physical Activity & Health, 3*(s1), S223–S240. https://doi.org/10.1123/jpah.3.s1.s223

Ewing, R., & Clemente, O. (2013). *Measuring Urban Design: Metrics for Livable Places*. Island Press.

Ewing, R., & Handy, S. (2009). Measuring the Unmeasurable: Urban Design Qualities Related to Walkability. *Journal of Urban Design, 14*(1), 65–84. https://doi.org/10.1080/13574800802451155

Fotheringham, A. S., Yang, W., & Kang, W. (2017). Multiscale Geographically Weighted Regression (MGWR). *Annals of the American Association of Geographers, 107*(6), 1247–1265. https://doi.org/10.1080/24694452.2017.1352480

Fu, X., et al. (2019). Do street-level scene perceptions affect housing prices in Chinese megacities? An analysis using open access datasets and deep learning. *PLOS ONE, 14*(5), Article e0217505. https://doi.org/10.1371/journal.pone.0217505

Geniaux, G., & Martinetti, D. (2018). A new method for dealing simultaneously with spatial autocorrelation and spatial heterogeneity in regression models. *Regional Science and Urban Economics, 72*, 74–85. https://doi.org/10.1016/j.regsciurbeco.2017.04.001

GeoDa (2020) *Spatial Regression with GeoDa*, *GeoDa Center*. Available at: https://geodacenter.github.io/documentation.html (Accessed: 26 May 2021).

Glaeser, E. L., et al. (2018). Big Data and Big Cities: The Promises and Limitations of Improved Measures of Urban Life. *Economic Inquiry, 56*(1), 114–137. https://doi.org/10.1111/ecin.12364

Gong, F.-Y., et al. (2018). Mapping sky, tree, and building view factors of street canyons in a high-density urban environment. *Building and Environment, 134*, 155–167. https://doi.org/10.1016/j.buildenv.2018.02.042

Griew, P., et al. (2013). Developing and testing a street audit tool using Google Street View to measure environmental supportiveness for physical activity. *International Journal of Behavioral Nutrition and Physical Activity, 10*(1), 103. https://doi.org/10.1186/1479-5868-10-103

Harvey, C. W. (2014). *Measuring Streetscape Design for Livability Using Spatial Data and Methods*. University of Vermont. Available at: https://scholarworks.uvm.edu/graddis/268.

He, L., Páez, A., & Liu, D. (2017). Built environment and violent crime: An environmental audit approach using Google Street View. *Computers, Environment and Urban Systems, 66*, 83–95. https://doi.org/10.1016/j.compenvurbsys.2017.08.001

Herbrich, R., Minka, T. and Graepel, T. (2006) 'TrueSkillTM: A Bayesian Skill Rating System', in *NIPS*.

Huang, Z., et al. (2017). Spatial and hedonic analysis of housing prices in Shanghai. *Habitat International, 67*, 69–78. https://doi.org/10.1016/j.habitatint.2017.07.002

Ito, K., & Biljecki, F. (2021). Assessing bikeability with street view imagery and computer vision. *Transportation Research Part C: Emerging Technologies, 132*, Article 103371. https://doi.org/10.1016/j.trc.2021.103371

Jackson, L. E. (2003). The relationship of urban design to human health and condition. *Landscape and Urban Planning, 64*(4), 191–200. https://doi.org/10.1016/S0169-2046(02)00230-X

Jacobs, J. (1992). *The Death and Life of Great American Cities*. Vintage Books.

Jain, A.K. and Chandrasekaran, B. (1982) '39 Dimensionality and sample size considerations in pattern recognition practice', in *Handbook of Statistics*. Elsevier (Classification Pattern Recognition and Reduction of Dimensionality), pp. 835–855. 10.1016/S0169-7161(82)02042-2.

Jansson, C. (2019) *Factors important to street users' perceived safety on a main street.* Available at: http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-260049 (Accessed: 16 February 2021).

Kang, Y., et al. (2020). Understanding house price appreciation using multi-source big geo-data and machine learning. *Land Use Policy*. , Article 104919. https://doi.org/10.1016/j.landusepol.2020.104919

Kang, Y., et al. (2021). Human settlement value assessment from a place perspective: Considering human dynamics and perceptions in house price modeling. *Cities, 118*, Article 103333. https://doi.org/10.1016/j.cities.2021.103333

Kelly, C.M. *et al.* (2013) 'Using Google Street View to audit the built environment: inter-rater reliability results', *Annals of Behavioral Medicine: A Publication of the Society of Behavioral Medicine*, 45 Suppl 1, pp. S108-112. 10.1007/s12160-012-9419-9.

Kim, H. and Carruthers, J.I. (2015) *Environmental Benefits of Green Space: Focusing on the Seoul Metropolitan Area*. SSRN Scholarly Paper ID 3511324. Rochester, NY: Social Science Research Network. Available at: https://papers.ssrn.com/abstract=3511324 (Accessed: 9 February 2021).

Kim, E. J., & Kim, H. (2020). Neighborhood Walkability and Housing Prices: A Correlation Study. *Sustainability, 12*(2), 593. https://doi.org/10.3390/su12020593

Kutner, M.H., Nachtsheim, C. and Neter, J. (2004) *Applied Linear Regression Models*. 4th edn. New York: McGraw-Hill/Irwin. Available at: https://books.google.com/books/about/Applied_Linear_Regression_Models.html?id=3_AlAQAAIAAJ (Accessed: 8 March 2021).

Larsen, J. E., & Blair, J. P. (2014). Price effects of surface street traffic on residential property. *International Journal of Housing Markets and Analysis [Preprint]*.. https://doi.org/10.1108/IJHMA-12-2012-0062

Lee, C., & Moudon, A. V. (2006). Correlates of Walking for Transportation or Recreation Purposes. *Journal of Physical Activity & Health, 3*(s1), S77–S98. https://doi.org/10.1123/jpah.3.s1.s77

Li, X., et al. (2015). Assessing street-level urban greenery using Google Street View and a modified green view index. *Urban Forestry & Urban Greening, 14*(3), 675–685. https://doi.org/10.1016/j.ufug.2015.06.006

Li, X., et al. (2019). A novel method for predicting and mapping the occurrence of sun glare using Google Street View. *Transportation Research Part C: Emerging Technologies, 106*, 132–144. https://doi.org/10.1016/j.trc.2019.07.013

Lin, L., & Moudon, A. V. (2010). Objective versus subjective measures of the built environment, which are most effective in capturing associations with walking? *Health & Place, 16*(2), 339–348. https://doi.org/10.1016/j.healthplace.2009.11.002

Lynch, K. (1960) *The Image of the City*. The MIT Press. Available at: https://mitpress.mit.edu/books/image-city (Accessed: 10 February 2021).

Ma, X., et al. (2021). Measuring human perceptions of streetscapes to better inform urban renewal: A perspective of scene semantic parsing. *Cities, 110*, Article 103086. https://doi.org/10.1016/j.cities.2020.103086

Manski, C. F. (1993). Identification of Endogenous Social Effects: The Reflection Problem. *The Review of Economic Studies, 60*(3), 531–542. https://doi.org/10.2307/2298123

McCluskey, J. J., & Rausser, G. C. (2001). Estimation of Perceived Risk and Its Effect on Property Values. *Land Economics, 77*(1), 42–55. https://doi.org/10.2307/3146979

Mehta, V. (2014). Evaluating Public Space. *Journal of Urban Design, 19*(1), 53–88. https://doi.org/10.1080/13574809.2013.854698

Microsoft (2005) 'TrueSkill^{TM} Ranking System', *Microsoft Research*. Available at: https://www.microsoft.com/en-us/research/project/trueskill-ranking-system/(Accessed: 13 May 2019).

Naik, N. et al. (2015) Do People Shape Cities, or Do Cities Shape People? The Co-Evolution of Physical, Social, and Economic Change in Five Major U.S. Cities. SSRN Scholarly Paper ID 2698292. Rochester, NY: Social Science Research Network. 10.2139/ssrn.2698292.

Naik, N., Philipoom, J., Raskar, R., & Hidalgo, C. (2014). Streetscore – Predicting the Perceived Safety of One Million Streetscapes. In *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops*. Washington, DC, USA: IEEE Computer Society. https://doi.org/10.1109/CVPRW.2014.121.

Nembrini, S., König, I. R., & Wright, M. N. (2018). The revival of the Gini importance? *Bioinformatics, 34*(21), 3711–3718. https://doi.org/10.1093/bioinformatics/bty373

Niu, T., Chen, Y., & Yuan, Y. (2020). Measuring urban poverty using multi-source data and a random forest algorithm: A case study in Guangzhou. *Sustainable Cities and Society, 54*, Article 102014. https://doi.org/10.1016/j.scs.2020.102014

NYU, UN Habitat and Lincoln Institute (2016) *Atlas of Urban Expansion – Shanghai, Shanghai*, *Atlas of urban expansion*. Available at: http://www.atlasofurbanexpansion.org/cities/view/Shanghai_Shanghai (Accessed: 17 February 2021).

Nyunt, M. S. Z., et al. (2015). Objective and subjective measures of neighborhood environment (NE): Relationships with transportation physical activity among older persons. *International Journal of Behavioral Nutrition and Physical Activity, 12*(1), 108. https://doi.org/10.1186/s12966-015-0276-3

Pandit, R., Polyakov, M., & Sadler, R. (2014). Valuing public and private urban tree canopy cover. *Australian Journal of Agricultural and Resource Economics, 58*(3), 453–470. https://doi.org/10.1111/1467-8489.12037

Park, K., et al. (2019). Street life and the built environment in an auto-oriented US region. *Cities, 88*, 243–251. https://doi.org/10.1016/j.cities.2018.11.005

Pedregosa, F., et al. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research, 12*(85), 2825–2830.

Pivo, G., & Fisher, J. D. (2011). The Walkability Premium in Commercial Real Estate Investments. *Real Estate Economics, 39*(2), 185–219. https://doi.org/10.1111/j.1540-6229.2010.00296.x

Qiu, W., et al. (2021). Subjectively Measured Streetscape Perceptions to Inform Urban Design Strategies for Shanghai. *ISPRS International Journal of Geo-Information, 10*(8), 493. https://doi.org/10.3390/ijgi10080493

Qiu, W., et al. (2022). Subjectively Measured Streetscape Qualities for Shanghai with Large-Scale Application of Computer Vision and Machine Learning. In P. F. Yuan (Ed.), *Proceedings of the 2021 DigitalFUTURES* (pp. 242–251). Singapore: Springer. https://doi.org/10.1007/978-981-16-5983-6_23.

Queralt, A., et al. (2021). Reliability of streetscape audits comparing on-street and online observations: MAPS-Global in 5 countries. *International Journal of Health Geographics, 20*(1), 6. https://doi.org/10.1186/s12942-021-00261-5

Rey, S. J., & Anselin, L. (2010). PySAL: A Python Library of Spatial Analytical Methods. In M. M. Fischer, & A. Getis (Eds.), *Handbook of Applied Spatial Analysis: Software Tools, Methods and Applications* (pp. 175–193). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-03647-7_11.

Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy, 82*(1), 34–55.

Rosen, S. (1979) 'Wage-based indexes of urban, quality of life', *Current issues i urban economics* [Preprint].

Rossetti, T., et al. (2019). Explaining subjective perceptions of public spaces as a function of the built environment: A massive data approach. *Landscape and Urban Planning, 181*, 169–178. https://doi.org/10.1016/j.landurbplan.2018.09.020

Roy, K. *et al.* (2016) 'Be aware of error measures. Further studies on validation of predictive QSAR models', *Chemometrics and Intelligent Laboratory Systems*, 152, pp. 18–33. 10.1016/j.chemolab.2016.01.008.

Rundle, A. G., et al. (2011). Using Google Street View to Audit Neighborhood Environments. *American journal of preventive medicine, 40*(1), 94–100. https://doi.org/10.1016/j.amepre.2010.09.034

Salazar Miranda, A., et al. (2021). Desirable streets: Using deviations in pedestrian trajectories to measure the value of the built environment. *Computers, Environment and Urban Systems.* , Article 101563. https://doi.org/10.1016/j.compenvurbsys.2020.101563

Salesses, P., Schechtner, K., & Hidalgo, C. A. (2013). The Collaborative Image of The City: Mapping the Inequality of Urban Perception. *PLOS ONE, 8*(7), Article e68400. https://doi.org/10.1371/journal.pone.0068400

Sander, H., Polasky, S., & Haight, R. G. (2010). The value of urban tree cover: A hedonic property price model in Ramsey and Dakota Counties, Minnesota, USA. *Ecological Economics, 69*(8), 1646–1656. https://doi.org/10.1016/j.ecolecon.2010.03.011

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

Shaw, S.-L., & Sui, D. (2020). Understanding the New Human Dynamics in Smart Spaces and Places: Toward a Splatial Framework. *Annals of the American Association of Geographers, 110*(2), 339–348. https://doi.org/10.1080/24694452.2019.1631145

Tian, H. *et al.* (2021) 'Evolution of Historical Urban Landscape with Computer Vision and Machine Learning: A Case Study of Berlin', *Journal of Digital Landscape Architecture*, p. 16. 10.14627/537705039.

Verma, D., Jana, A., & Ramamritham, K. (2020). Predicting human perception of the urban environment in a spatiotemporal urban setting using locally acquired street view images and audio clips. *Building and Environment, 186*, Article 107340. https://doi.org/10.1016/j.buildenv.2020.107340

Wang, R., et al. (2019). Perceptions of built environment and health outcomes for older Chinese in Beijing: A big data approach with street view images and deep learning technique. *Computers, Environment and Urban Systems, 78*, Article 101386. https://doi.org/10.1016/j.compenvurbsys.2019.101386

Wen, H., Xiao, Y., & Zhang, L. (2017). 'School district, education quality, and housing price: Evidence from a natural experiment in Hangzhou. *China', Cities, 66*, 72–80. https://doi.org/10.1016/j.cities.2017.03.008

Willis, K. G., Powe, N. A., & Garrod, G. D. (2005). Estimating the Value of Improved Street Lighting: A Factor Analytical Discrete Choice Approach. *Urban Studies, 42*(12), 2289–2303. https://doi.org/10.1080/00420980500332106

Wilson, J.Q. and Kelling, G.L. (2011) 'The police and neighborhood safety Broken Windows', in *Social, Ecological and Environmental Theories of Crime*. Routledge.

Wolch, J. R., Byrne, J., & Newell, J. P. (2014). Urban green space, public health, and environmental justice: The challenge of making cities "just green enough". *Landscape and Urban Planning, 125*, 234–244. https://doi.org/10.1016/j.landurbplan.2014.01.017

Yang, J., et al. (2021). The financial impact of street-level greenery on New York commercial buildings. *Landscape and Urban Planning, 214*, Article 104162. https://doi.org/10.1016/j.landurbplan.2021.104162

Yao, Y., et al. (2019). A human-machine adversarial scoring framework for urban perception assessment using street-view images. *International Journal of Geographical Information Science, 33*(12), 2363–2384. https://doi.org/10.1080/13658816.2019.1643024

Yao, Y., et al. (2021). Discovering the homogeneous geographic domain of human perceptions from street view images. *Landscape and Urban Planning, 212*, Article 104125. https://doi.org/10.1016/j.landurbplan.2021.104125

Ye, Y., et al. (2019). Daily Accessed Street Greenery and Housing Price: Measuring Economic Performance of Human-Scale Streetscapes via New Urban Data. *Sustainability, 11*(6), 1741. https://doi.org/10.3390/su11061741

Yin, L., & Wang, Z. (2016). Measuring visual enclosure for street walkability: Using machine learning algorithms and Google Street View imagery. *Applied Geography, 76*, 147–153. https://doi.org/10.1016/j.apgeog.2016.09.024

Yin, L., Cheng, Q., Wang, Z., & Shao, Z. (2015). 'Big data' for pedestrian volume: Exploring the use of Google Street View images for pedestrian counts. *Applied Geography, 63*, 337–345. https://doi.org/10.1016/j.apgeog.2015.07.010

Zhang, F., et al. (2018). Measuring human perceptions of a large-scale urban region using machine learning. *Landscape and Urban Planning, 180*, 148–160. https://doi.org/10.1016/j.landurbplan.2018.08.020

Zhang, Y., & Dong, R. (2018). Impacts of Street-Visible Greenery on Housing Prices: Evidence from a Hedonic Price Model and a Massive Street View Image Dataset in Beijing. *ISPRS International Journal of Geo-Information, 7*(3), 104. https://doi.org/10.3390/ijgi7030104

Zhao, H. *et al.* (2016) 'Pyramid Scene Parsing Network'. Available at: https://arxiv.org/abs/1612.01105v2 (Accessed: 10 April 2020).

Zhou, B., et al. (2018). Semantic understanding of scenes through the ADE20K Dataset. *Int J Comput Vis., 127*, 302–321.

Zhou, H., et al. (2019). Social inequalities in neighborhood visual walkability: Using street view imagery and deep learning technologies to facilitate healthy city planning. *Sustainable Cities and Society, 50*, Article 101605. https://doi.org/10.1016/j.scs.2019.101605