

Article

Heterogeneity Study of the Visual Features Based on Geographically Weighted Principal Components Analysis Applied to an Urban Community

Yong Liu *, Shutong Yang and Shijun Wang

Shanghai Academy of Fine Arts, Shanghai University, Shanghai 200444, China; 13617801722@163.com (S.Y.); wangshijun975@163.com (S.W.)

* Correspondence: liuy533@shu.edu.cn; Tel.: +86-1861-6503-983

Abstract: Communities in urban space are the most basic living units. Community visual features directly reflect the local living quality and influence the perception of residents and visitors. The evaluation of the community visual features is of great significance to the space design under the guidance of urban landscape recognition and urban space perception. Based on the street view image data, this paper analyzes the composition of visual features in the community space scale by using the geographically weighted principal components analysis. GWPCA can not only reflect the global characteristics, but also analyze the local components, thus describing the visual features of the community in a comprehensive manner. The results show that: (1) community visual features have significant spatial heterogeneity at different statistical scales, and the spatial heterogeneity of community visual features can provide a basis for urban landscape planning and design; (2) the combination mode of dominant visual elements can reflect different community landscapes. The analysis of this paper further illustrates the effectiveness and application prospect of street view images in identifying the landscape composition mode of urban space from the medium-micro perspective. This conclusion is helpful for planners to learn the dominant visual features of the community through street view images, and, further, use the classification of elements of street view images to guide the planning and design of cityscape.

Keywords: street view image; community visual spatial features; geographical weighted principal components analysis; GWPCA



Citation: Liu, Y.; Yang, S.; Wang, S. Heterogeneity Study of the Visual Features Based on Geographically Weighted Principal Components Analysis Applied to an Urban Community. *Sustainability* **2021**, *13*, 13488. <https://doi.org/10.3390/su132313488>

Academic Editors: Tatsuhito Kono and Nao Sugiki

Received: 28 October 2021

Accepted: 26 November 2021

Published: 6 December 2021

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



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Community is one of the most basic space units in a city, and it is also the main place living and rest for urban residents in daily life. The streets and public spaces in communities are main spaces for the daily activities of urban residents. Communities have three basic attributes: built environment, functional attribute and visual appearance [1,2]. Built environment and functional attribute constitute the spatial layout and functional service of community. There are a series of researches on evaluating community function and built environment by using POI data, social media data, remote sensing data, LBS data and heat map data [3–7]. In contrast, there are relatively few studies on the visual appearance features of communities [3]. The visual appearance features of communities mainly refer to all kinds of things that can be seen in the local space of communities, including cars, bicycles and pedestrians which are dynamic as well as buildings, roads and plants which are static. These factors affect the life experience of residents and the impression of tourists on the community.

The innovation of sensor technology and digital technology promotes the emergency of new data collection and processing means, which provides high-resolution street real-life image data for the evaluation and research of urban spatial environment [8]. Street view images provide a new dimension for urban physical feature recognition on a large-scale

spatial scale, and they also provide a new opportunity for refined evaluation of urban physical features [9]. Compared with the traditional field survey data, street view images not only have a wide coverage and a large amount of data, but also provide API interface development. They can be downloaded free of charge from the interface provided by map operators. More importantly, the images contain abundant urban infrastructure information, which can provide artificial and natural landscapes at street level, and intuitively and accurately express the details of urban built environment. These advantages make street view image an important new data source in urban environmental assessment research [8,10,11].

The new data environment, the innovation of computer technology and geographic information technology provide new conditions for large-scale and meticulous examination of community visual features. Street view images record the built environment of cities in a relatively comprehensive manner. It will be a mainstream way of future urban research to use street view images to analyze and recognize urban space, which can respond to the existing theories and even explore new research fields.

The analysis of community visual features based on street view explores the visual feelings of street such as green viewing rate, sky opening degree and motorization degree [8,12,13]. There is a phenomenon of spatial differentiation in community visual features, and scholars believe that the visual appearance features of blocks are influenced by land use, urban planning, population mobility and aggregation [1,4], and there are significant differences in visual features in different areas of different cities [14,15].

Taking street view images in Huangpu district of Shanghai as an example, this paper, based on semantic segmentation of full convolution neural network and geographically weighted principal components analysis (GWPCA), analyzes the spatial distribution of dominant community visual features, judges the heterogeneity of community visual features in different spatial scales and tries to summarize the patterns of dominant visual features in local areas from the perspective of visual element composition.

2. Literature Review

The research on the community visual features is a kind of research topic on street space from the perspective of humanism [16], which pays attention to the visual perception elements and spatial differentiation of community, block and street. Traditional research methods of visual appearance features of communities, blocks and streets mainly focus on questionnaire survey, on-the-spot investigation and structural interview, and street visual features are evaluated and studied according to the subjective feelings and cognition of researchers and interviewees [17–20]. Chen et al. recorded the underlying interface variables and walking activities of commercial streets through on-the-spot observation and taking pictures, in order to explore the influence of street visual appearance on walking activities [17]; Xiong et al. took people as the core object of street evaluation, and made quantitative observation and evaluation based on people-oriented observation method on people's composition, distribution, behavior and color, so as to improve street quality [18]; Brown et al. learned the influence of the building environment on physical activities through telephone interviews, questionnaires and systematic observation [20]. These studies have certain limitations in objectivity of evaluation results, large-scale refined evaluation and humanistic perspective [9,21–23].

With the development of technology, to solve the limitations caused by field survey data, scholars began to use street view image data to analyze the visual features of communities. However, most existing studies use traditional regression analysis and principal components analysis methods to analyze community features. Salesses et al., based on Google Street View, used Moran's I statistics to measure the spatial features of urban environmental perception [24]; Li et al., based on Google Street View, used Poisson regression model with spatial filtering effect to study the relationship between violent crime and physical features of urban living environment [25]. Traditional principal components analysis can reduce the dimension of high-dimensional data, keep the most important

features, and remove redundant parts and relatively unimportant features in the data, so as to improve the data processing speed and the interpretability of the results. However, by analyzing and summarizing the community features, it is difficult to explain the high correlation among visual features by relying on the traditional regression analysis and principal components analysis [5]. For example, there may be high collinearity between the width of streets and the width of sidewalks, and between the height of buildings and the visibility of the sky. Traditional statistical analysis is difficult to eliminate this collinearity, and it is less convincing in explaining the composition of the community's visual features. Furthermore, in communities affected by socio-economic phenomena, the visual features are spatially heterogeneous and relevant, so it is necessary to analyze the spatial dimensions of visual features. However, the traditional principal components analysis is a statistical analysis method based on the overall situation [26], which ignores the spatial variation of the main load factor vector and variance contribution, and cannot realize the analysis of local spatial distribution features.

Based on the former researches, a small number of scholars have optimized their research methods, and replaced the traditional regression analysis and principal components analysis methods with geographically weighted principal components analysis to realize the analysis of local spatial distribution features. Chao et al., based on Baidu Street View, analyzed the community visual features at different scales by using geographically weighted principal components, and considered that under the statistical scale of 1 km, the visual perception evaluation based on Street View has a strong explanatory power, and has application scenarios in judging the spatial laws and designing the styles at the regional scale [26]. Geographically weighted principal components analysis (GWPCA) is a local statistical analysis method [5], which can help to understand the structural features of geographical changes in socio-economic phenomena, identify the maximum load factor variables of the principal components of spatial changes, visualize the local spatial distribution features of the variance contribution of principal components, and optimize the selection of geographically weighted regression variables. It can not only make up for the deficiency of principal components analysis in spatial distribution interpretation, but also enhance the analysis and interpretation of socio-economic geographical differentiation.

3. Methods

3.1. Study Area

The study area of this paper is the whole Huangpu District of Shanghai (Figure 1). The Huangpu District has both historical and cultural blocks and modern commercial and living blocks. It is adjacent to Huangpu River, with rich and diversified landscape features, which can meet the needs of analyzing the heterogeneity of visual features at the micro and macro scales.

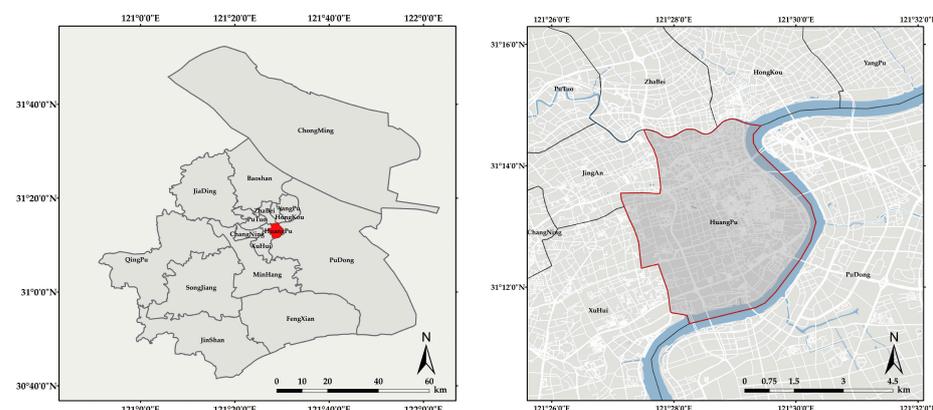


Figure 1. Study area in Shanghai.

3.2. Data Sources

The research object is the street view image within the research scope. The road network data used comes from OpenStreetmap, and the street view image data comes from the open platform of Baidu Map.

A sampling point is generated every 100 m on the road network (Figure 2). According to the geographical coordinates of the sampling point, the Street View Image Application Programming Interface (API) provided by Baidu Map can be called to return street view in images. At each sampling point, call parameters are set along four horizontal directions (0° , 90° , 180° and 270°), with elevation angle of 12° , image width of 512 and image height of 1024. A total of 11,560 street view images are collected. Except for the residential area without road network, the sampling points of street view images cover most areas of Huangpu District, which can represent the visual features of Huangpu District.



Figure 2. Study area and distribution of 2890 sampling points of street view images.

This study collected the above data in March, 2021. Baidu map provided the street view data in May, 2019, and OSM platform provided the road network data in 2021. Therefore, some sampling points may return null values, and these sampling points need to be deleted. In the end, 11,388 of the 11,560 street view images are valid.

3.3. Research Methods

In this paper, based on street view images, semantic segmentation is used to analyze street view images, and the proportion of various visual features at the corresponding spatial points of each street view image is identified. Based on the concept of 15-min life circle put forward by the Chinese government, Chao et al. counted street view images from three scales of 1 km, 500 m and 200 m. The re-search argued that the smaller the statistical unit is, the more space with similar visual features could be identified [26]. Based on this, this study further reduced the research scale, and used three scales of 100 m, 150 m and 200 m to count street view images to carry out analysis with a higher degree of refinement. Then, the GWmodel package based on R platform is used to realize geographically weighted principal components analysis at three scales, in order to identify whether the community visual features at different scales have spatial differentiation and aggregation, and to test the effectiveness of this method in identifying community visual features.

3.3.1. Semantic Segmentation Based on Full Convolution Neural Network (FCN)

Image semantic segmentation is the cornerstone technology of image understanding and one of the important tasks in the field of computer vision. It plays an important role in automatic driving system (specifically street scene recognition and understanding), unmanned aerial vehicle application (landing point judgment) and wearable device application. At present, there are many high-quality open-source frameworks that can be used. In the semantic segmentation of street view images, FCN and SegNet are commonly used.

The core idea of full convolution neural network is to learn the mapping from pixel to pixel. It is a framework for training end-to-end and pixel to pixel full convolution neural network, which is an extension of CNN model (Figure 3). The existing FCN model can identify up to 150 street view elements [27].

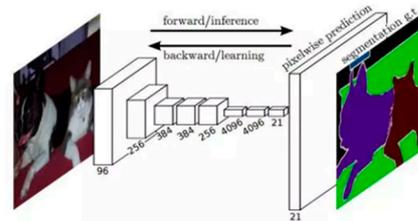


Figure 3. Model architecture of full convolution neural network (FCN).

3.3.2. Geographic Weighted Principal Components Analysis (GWPCA)

The statistical concept of geographical weighting was given by Fotheringham, Brunsdon, Charlton and Lloyd, etc. [28]. Geographically weighted principal components analysis introduces a series of formulas of geographical weighting, and brought the interaction of geographical positions among variables into the calculation, thus effectively solving the problem of spatial heterogeneity in multivariate data. Through the calculation of geographically weighted mean, geographically weighted variance and covariance of different locations, the calculation results of geographically weighted principal components analysis are obtained [29].

Before using geographically weighted principal components analysis, we need to carry out collinearity test and traditional principal components analysis on the data. KMO and Bartlett tests are commonly used. After identifying the collinearity and correlation features of the data, the general KMO value needs to reach more than 0.5. Then, it is necessary to carry out principal components analysis on the model to further identify the influence of variable features on the whole. According to the results of principal components analysis, whether to use geographically weighted principal components analysis can be determined.

Generally, for a series of analysis variables x_i , the coordinates at the spatial position i are (u, v) . GWPCA model performs coherent processing on variables x_i , and positions u and v correlations. Through calculation of the mean vector $\mu(u, v)$ and the local variance $\Sigma(u, v)$, the local covariance matrix is decomposed to obtain the local eigenvalues and local eigenvectors, and then obtain the local principal components.

The formula for calculating the local covariance matrix is:

$$\Sigma(u, v) = X^T W(u, v) X \quad (1)$$

wherein, X is the cortege matrix of original variables and sample units. For the i th line of the matrix, the local feature vector of region i is used to obtain the local principal component score of the i th line. $W(u, v)$ is the geographically weighted diagonal matrix, and the principal component of (u, v) is:

$$L(u_i, v_i) V(u_i, v_i) L(u_i, v_i)^T = \Sigma(u_i, v_i) \quad (2)$$

wherein, u_i, v_i is the local eigenvector matrix, $V(u_i, v_i)$ is the diagonal matrix of local eigenvalues and $\Sigma(u_i, v_i)$ is the local covariance matrix. According to the results of geographically weighted principal components analysis, a series of local components variances and loads can be obtained, which can be visualized in space to help to identify the spatial variation features of multivariate data structures. Therefore, geographically weighted principal components analysis can evaluate: (1) the spatial explanation of variance contribution of principal components; (2) local influence of load factor on principal components.

4. Results and Discussion

4.1. Results

4.1.1. Data Processing

The semantic segmentation model [27] developed by Yao et al. is used to semantically segment the collected street view images. The pixel contrast accuracy of this model is 81.44% in the training data set and 66.84% in the test data set. The result of semantic segmentation contains a total of 151 elements. According to the research needs and the main components of actual visual composition [26], a total of 13 main features are selected, including the sky, high-rise buildings, ordinary buildings, bicycles, railings, building floors, road floors, sidewalks, street lamps, shrubs, trees, walls and roadways. The results of semantic segmentation are shown in Figure 4. The selected visual elements generally account for more than 90% of the images, which can represent the dominant visual features. Each sampling point contains images in four directions, and 13 types of features are averaged as indicators of sampling points. Then, the sampling points included are counted by using grids at three scales (100 m, 150 m and 200 m). Each grid represents the average level of street view elements in the region.



Figure 4. Semantic segmentation schematic diagram.

4.1.2. Overall Features Based on Traditional Principal Components Analysis

Before principal components analysis and geographically weighted principal components analysis, it is necessary to test the correlation and collinearity of data. Here, KMO and Bartlett tests are used (Table 1). The test results show that although KMO index is low and the accuracy of the highest street scale data is only 0.59, all the results are significant at the level of 1%. KMO index represents the distribution of test data and the independence between variables. According to scholars' research and judgment, the proportion of visual features in street view images based on semantic segmentation has variable distribution interval of [0, 1], and it is easy to form strong collinearity among variables, so the judgment interval of KMO index can be appropriately relaxed [26]. This indicates that there is a linear correlation between the original variables, and the principal components analysis can be performed on original variables representing visual features. Meanwhile, with the increase of spatial scale, KMO index also increases, which indicates that the greater the spatial scale, the stronger the correlation between different visual features.

Table 1. KMO and Bartlett's test on data of different spatial scales.

		Grid of 100 m	Grid of 150 m	Grid of 200 m
Kaiser-Meyer-Olkin		0.44	0.47	0.49
Bartlett	χ^2	47,698.00	27,384.63	17,986.37
	df	12	12	12
	Sig.	0.000	0.000	0.000

For traditional principal components analysis. First, traditional principal components analysis requires to standardize the standard deviation of variables. Since the area ratio of each element is obtained after semantic segmentation of street view images, the calibration

of model variables is percentage, so standardization is unnecessary. In factor analysis, principal components are extracted in the form of variance–covariance matrix, and the cumulative variance and load factor coefficients can be obtained (Tables 2 and 3). It can be seen from the table that the cumulative contribution of the first three principal components gradually increases with the increase of spatial scale, and the cumulative contribution of the first three principal components exceeds 50% with the data accuracy of 200 m, which indicates that the first three principal components can explain some features of community visual features.

Table 2. Summary of global PCA at three spatial scales.

	Grid of 100 m			Grid of 150 m			Grid of 200 m		
	Comp.1	Comp.2	Comp.3	Comp.1	Comp.2	Comp.3	Comp.1	Comp.2	Comp.3
Standard deviation (STD)	1.684	1.552	1.353	1.605	1.389	1.261	1.575	1.417	1.290
Proportion of variance (PPTV)	0.185	0.157	0.119	0.223	0.167	0.138	0.221	0.179	0.149
Cumulative proportion (CPTV)	0.185	0.342	0.462	0.223	0.390	0.528	0.221	0.401	0.549

Table 3. The component matrix of each visual load factor under three spatial scales.

	Grid of 100 m			Grid of 150 m			Grid of 200 m		
	Comp.1	Comp.2	Comp.3	Comp.1	Comp.2	Comp.3	Comp.1	Comp.2	Comp.3
x1, Wall	0.006	0.107	0.183	0.123	0.173	0.322			0.187
x2, Building		−0.619	−0.348	0.593		−0.389	0.604	−0.119	−0.222
x3, Sky	−0.202		0.544		−0.595	0.264		0.652	
x4, Floor	0.450		0.130		0.141	0.233	0.103		0.267
x5, Tree	−0.296	0.563	−0.136	−0.649	0.262	−0.112	−0.630	−0.337	
x6, Road	−0.360		0.492	−0.255	−0.616		−0.317	0.491	−0.297
x7, Pave	−0.377		−0.239	−0.256		−0.643	−0.177	−0.284	−0.585
x8, Ground	0.116	0.276	−0.186	−0.120	0.136				0.125
x9, Plant		0.319	−0.373	−0.197	0.276	0.180	−0.232	−0.223	0.372
x10, Fence	−0.132	0.260		−0.106			−0.144	0.166	0.186
x11, Skyscraper									
x12, Street-light	−0.175				−0.208			0.172	−0.247
x13, Bike	−0.168	−0.162	−0.157			−0.377		−0.120	−0.399

Note: “x” stands for different main features.

In the table, at the first principal component level, the data of 100-m accuracy can only explain 18% of the data structure features. At the spatial scale of 200-m accuracy, 22% of the data structure features can be explained, and the model effect has been improved to some extent, indicating that the street view visual feature model may have a better explanatory effect with the expansion of the research scale. In Table 3, the load factor coefficients of the first three principal components under the three research scales are also retained. Among them, the correlation values of walls, buildings, trees, ground and plants are more comprehensive, which is consistent with traditional cognition, and the visual perception of residents mainly comes from factors such as fences, buildings and plants in the community.

Generally speaking, although the total proportion of the top three principal components is small, due to the particularity of street view data, the proportion of each element (in percentage) may change linearly with the spatial position, which requires further analysis from the perspective of spatial position. Based on PCA model, further geographically

weighted principal components analysis should be performed to identify the local spatial features and interpretation effectiveness of these characteristic elements.

4.1.3. Local Features Based on Geographically Weighted Principal Components Analysis

On R platform, the operation of geographically weighted principal components analysis at three spatial scales is realized, and the results of principal component parameters representing visual features are visualized on GIS platform. Figures 5–7 show the street visual feature variables corresponding to the cumulative contribution of variance and the maximum load factor of each spatial unit. The spatial bandwidths of 100 m, 150 m and 200 m are 323, 228 and 118, respectively. This shows that the smaller the spatial statistical scale, the greater the spatial heterogeneity and complexity of the visual feature variables, and the higher the possibility of changes in visual features caused by spatial and positional changes.

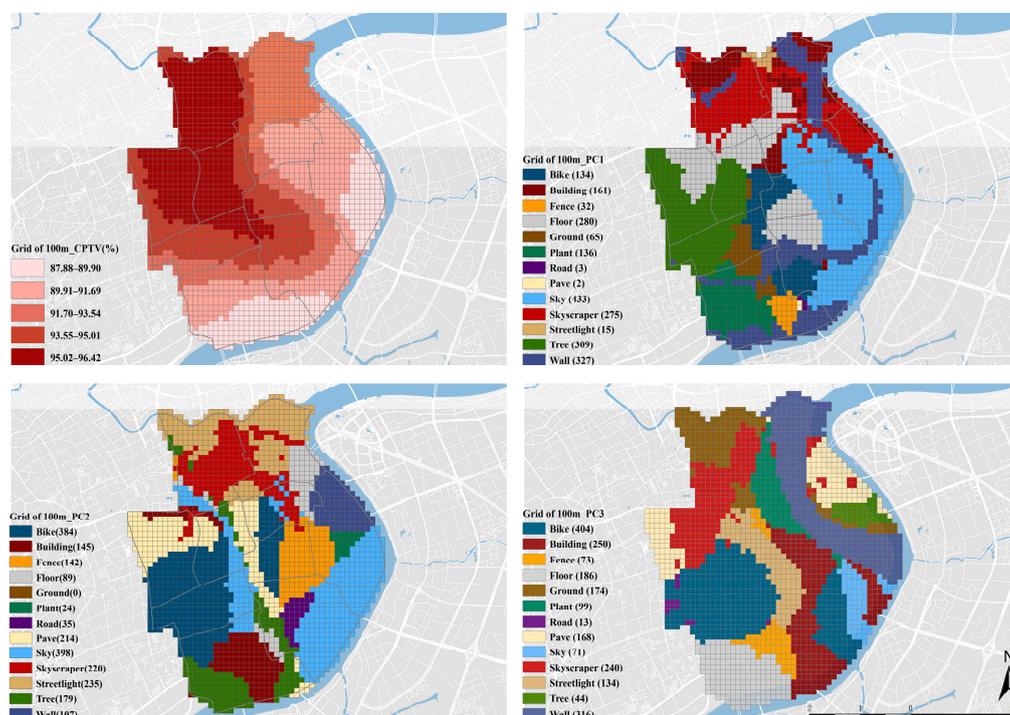


Figure 5. In GWPC1-PC3, the cumulative contribution of variance at grid of 100 m and the spatial distribution of load factors of visual features.

Based on the results of traditional principal components analysis, the first three principal components at three spatial scales are retained in the study (Table 4). From the change of first principal component in spatial scale, it can be seen that the value of first principal component also expands with the expansion of spatial scale (the increase of spatial statistical scale). This is mainly because of the increasing correlation of variables with the expansion of spatial scale. What is more, the variables can already represent most of the visual spatial features before dimension reduction.

According to the spatial variation of local cumulative contribution of variance, the cumulative contribution of variances of Nanjing Road Street, Huaihai Middle Road Street and Ruijin Second Road Street have certain stability and high values, which indicates that the community visual features of these areas are similar, while Wuliqiao Street, Bansongyuan Street and Xiaodongmen Street which are along Huangpu River have lower cumulative contribution of variances which change greatly at different spatial scales. This indicates that the spatial visual features change greatly, which is also related to the fact that these regions are along the river on one side and buildings and plots of different functional types are arranged on the other side.

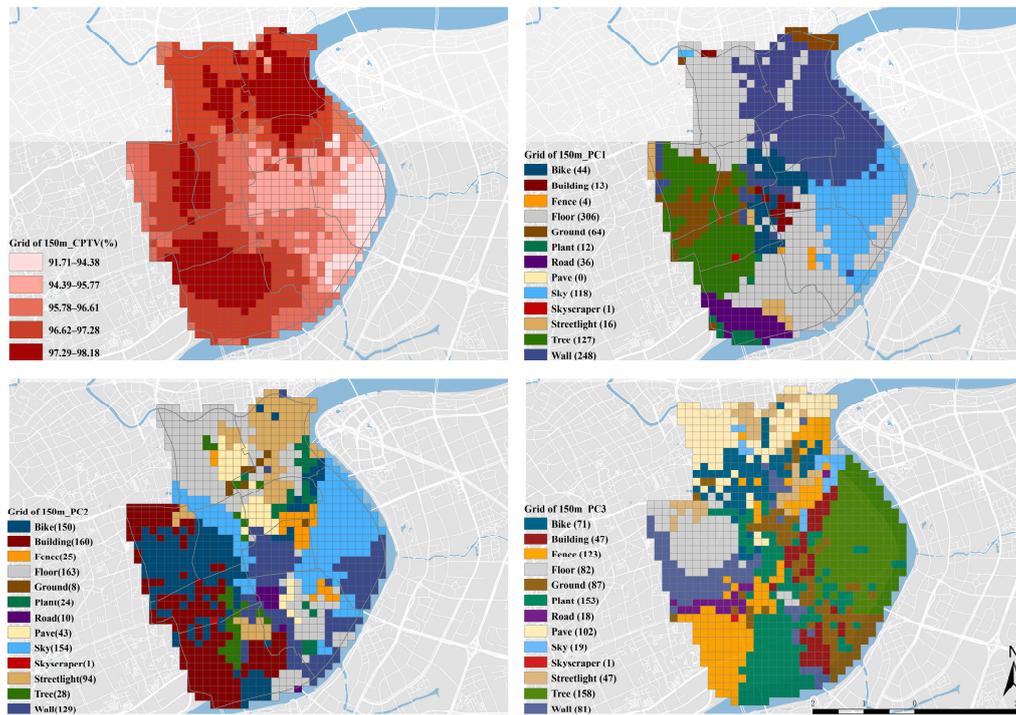


Figure 6. In GWPC1-PC3, the cumulative contribution of variance and the spatial distribution of load factors of visual features at the grid of 150 m.

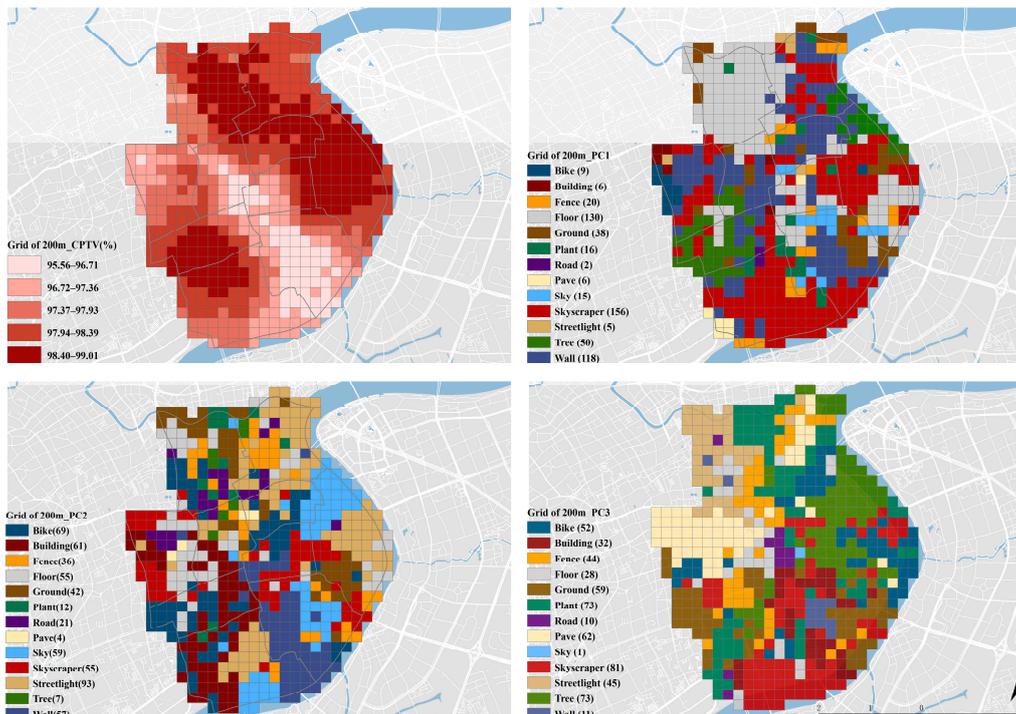


Figure 7. In GWPC1-PC3, the cumulative contribution of variance at grid of 200 m and the spatial distribution of load factors of visual features.

Table 4. Results of geographically weighted principal components analysis of community visual features at three spatial scales.

	Grid of 100 m			Grid of 150 m			Grid of 200 m		
	Comp.1	Comp.2	Comp.3	Comp.1	Comp.2	Comp.3	Comp.1	Comp.2	Comp.3
Standard deviation (STD)	1.480	1.344	1.265	1.510	1.357	1.296	1.615	1.347	1.317
Proportion of variance (PPTV)	0.169	0.139	0.123	0.175	0.142	0.129	0.201	0.140	0.133
Cumulative proportion (CPTV)	0.169	0.308	0.431	0.175	0.317	0.446	0.201	0.340	0.474
loadings									
x1, Wall	0.531	-	0.170	0.532	-	0.105	0.522	-	-
x2, Building	-	-0.593	-0.336	-	0.398	-0.555	-0.187	-0.583	0.169
x3, Sky	-0.249	-	0.577	-0.193	0.343	0.490	-0.175	0.298	0.490
x4, Floor	0.474	-	0.140	0.476	-	-	0.494	-	-
x5, Tree	-0.232	0.593	-0.190	-0.230	-0.614	-	-0.212	0.281	-0.570
x6, Road	-0.407	-	0.384	-0.390	0.182	0.351	-0.361	0.265	0.226
x7, Pave	-0.335	-	-0.336	-0.355	-0.254	-0.244	-0.349	-0.138	-0.292
x8, Ground	0.140	0.272	-	0.185	-0.284	0.171	0.183	0.213	-0.159
x9, Plant	-	0.344	-0.211	-	-0.326	-	-	0.231	-0.361
x10, Fence	-0.127	0.262	0.145	-0.112	-	0.258	-0.142	0.346	-
x11, Skyscraper	-	-	0.296	-	0.217	0.266	-	0.244	0.239
x12, Streetlight	-0.184	-	-	-0.192	-	-	-0.195	-	0.198
x13, Bike	-0.151	-0.124	-0.218	-0.159	-	-0.259	-0.160	-0.346	-

Note: "x" stands for different main features.

Under different spatial statistical calibers, the spatial distribution of cumulative contribution of variance of community visual feature variables is quite different. For example, at grid of 100 m, the cumulative contribution of variance decreases from northwest to southeast, and at grid of 150 m, it has a high value in the Bund Street and the Dapujiao street, and decreases to the periphery. The results at grid of 200 m are close to those at grid of 150 m.

Generally speaking, the results of geographically weighted principal components analysis show that there is obvious spatial heterogeneity in the community visual features in Huangpu District, and there is also spatial homogeneity in local visual features. The correlation and contribution among trees, buildings and sidewalks show sensitive differences with the change of spatial scale. For example, at grid of 100 m, high-rise buildings and plants have higher load factors, while at grid of 150 m, the higher load factor variables are bicycles and ground. There are similar changes at grid of 200 m.

In a word, the overall analysis of the results of geographically weighted principal components analysis shows that the meanings of the changes in community visual features that can be explained by geographically weighted principal components analysis are also different at different scales. Among it, the grid of 100 m can determine the activity intensity and vegetation coverage of people, and the grids of 150 m and 200 m can represent the functions of communities to a certain extent. Among it, although the first principal component value of the data at the grid of 200 m is relatively high in the model, there is a relatively serious loss of information according to the spatial distribution of the data.

4.2. Discussion

Further analysis of the spatial distribution of load factors shows that the dominant load factors are trees, high-rise buildings, sidewalks and bicycles at grid of 100 m, trees, walls and ground at grid of 150 m and trees, high-rise buildings and ground at grid of 200 m. Although the spatial distribution of the load factors is quite different, from the homogeneity and heterogeneity of the variables of the main visual features, the high-rise buildings and trees are the main community visual features of the Huangpu District, which is also in line with the reality.

According to the analysis of the spatial distribution of load factors at grid of 100 m, in terms of the first principal component, there are more high-rise buildings in Nanjing East Road (Figure 8a,b) and the Bund (Figure 8c,d), and there is wide ground in Huaihai Middle Road Street (Figure 8e,f). Besides, there is more complete sky around Laoximen and Xiaodongmen (Figure 8g,h), and more green space in Ruijin Second Road Street, Dapuqiao Street and Wuliqiao Street (Figure 9). The following figures list street view images of some points located in these areas with higher load factors in the street view data.

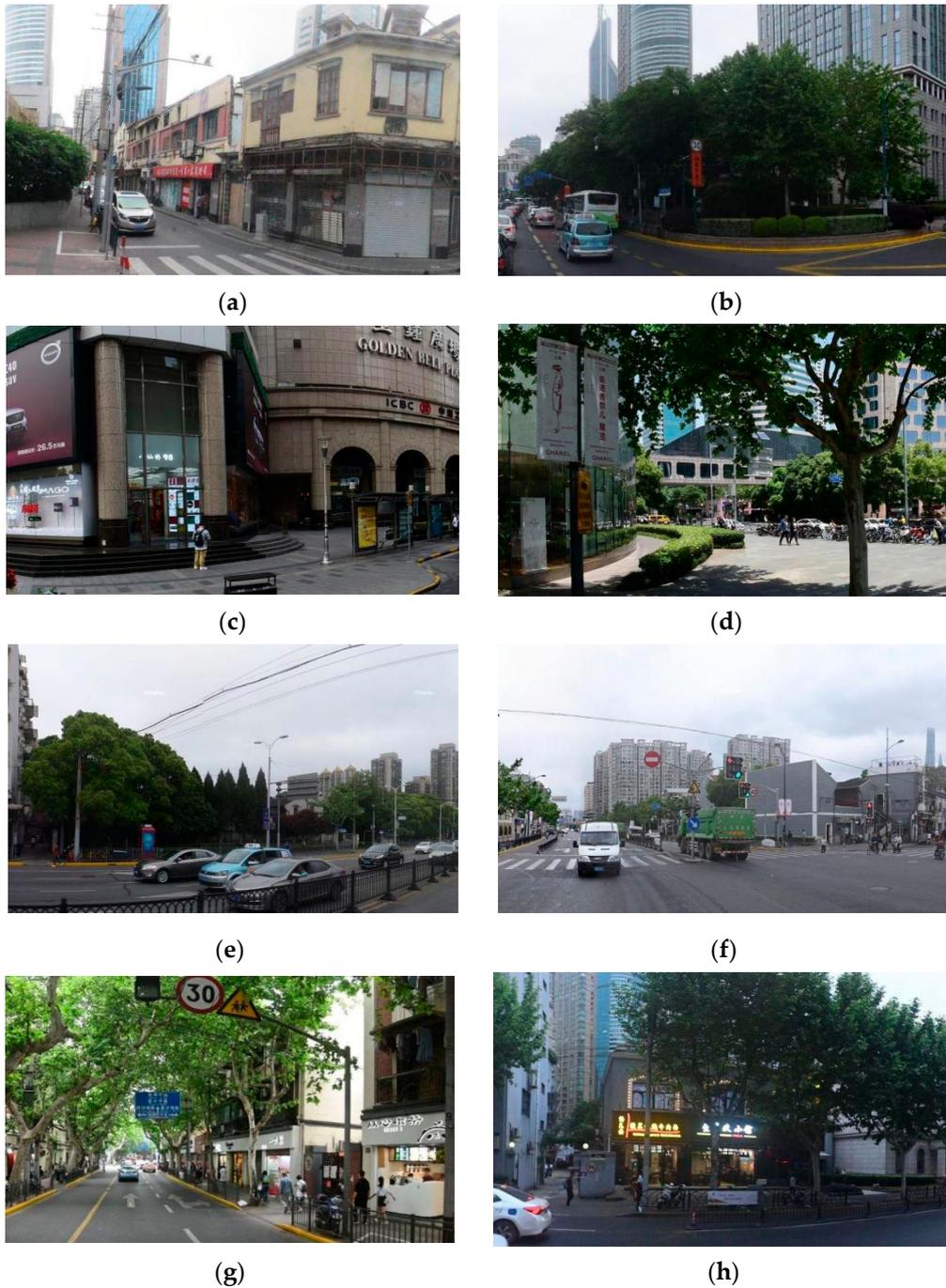


Figure 8. (a,b) Typical street view images of Nanjing East Road and the Bund Street; (c,d) typical street view images of Huaihai Middle Road Street; (e,f) typical street view images of Laoximen and Xiaodongmen streets; (g,h) typical street view images of Ruijin Second Road, Dapuqiao Street and Wuliqiao Street.

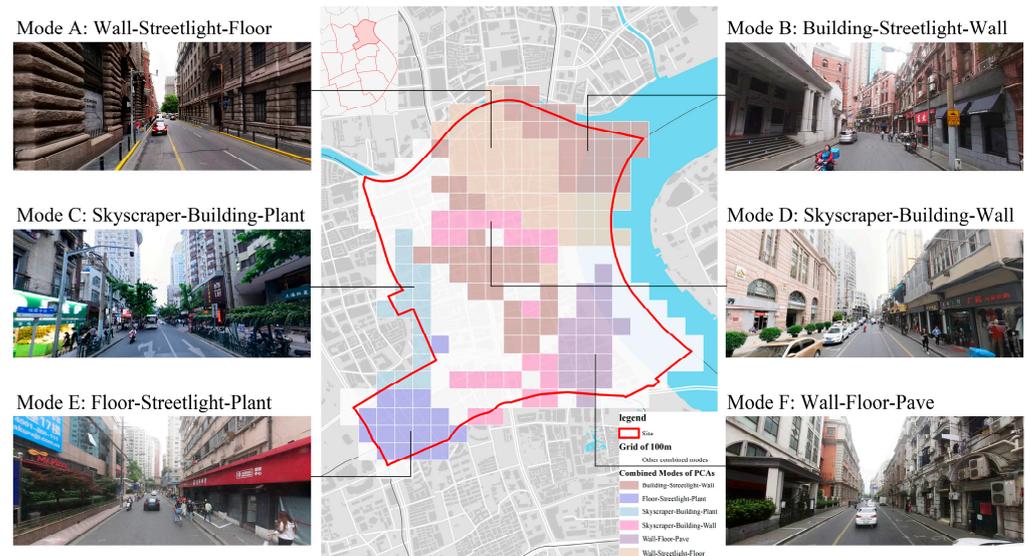


Figure 9. Six principal component combination modes of visual features of the Bund Street and reference street view images.

Although there is no strong correlation among variance contributions at different scales, they can be interpreted with certain credibility at their respective spatial statistical scales independently, and the spatial distribution of variance contribution further confirms the spatial differentiation state of community visual elements. The above geospatial statistical analysis is based on the street view image data of Huangpu District. At different spatial statistical scales, the application of community visual feature analysis has different scenarios. For example, in the model calculated by grid of 100 m, although the contribution of the first three principal components is not high, further analysis of the model results at grid of 100 m shows that the calculation results summarize the visual features well.

Taking the Bund Street as an example, the Bund area with traditional style, historical street and shops along the street as its main styles also has different visual features. Among the first three principal component composition modes in the 100 m grid model, except the combination mode with the proportion less than 5%, six principal component combination modes can be summarized from 267 statistical units (2172 in Huangpu District) (Table 5). Among the 267 units, 54 units are of “Building–Streetlight–Wall” combination mode. The combination mode with lower proportion is “Skyscraper–Building–Plant”, accounting for only 5.62%. The former shows the spatial features of small blocks, dense road network and low-rise buildings in the Bund area, while the latter shows the landscape visual features that there are more high-rise buildings around the Bund area and less proportion of green plants. Comparing with the actual street view images in the Bund area (Figure 9), it is observed that street view images with similar styles are also classified into different combination modes in the model because of their different dominant visual components, and the spatial heterogeneity of combination modes has shaped various visual features in the Bund area.

Table 5. Proportion of the composition modes of the first three principal components in the Bund Street at the grid of 100 m.

Combination Modes of PCAs				%
1	Building	Streetlight	Wall	20.22%
2	Wall	Streetlight	Floor	16.48%
3	Skyscraper	Building	Wall	8.61%
4	Floor	Streetlight	Plant	7.12%
5	Wall	Floor	Pave	7.12%
6	Skyscraper	Building	Plant	5.62%

Through the analysis of street view visual features, it is verified once again that it is feasible to identify the spatial heterogeneity of visual features based on street view, and the spatial distribution of visual features heterogeneity can reflect the differences of landscape styles in urban areas: (1) in the global analysis, to verify the universality of spatial heterogeneity of visual features, this paper uses street view images to further understand the relationship between visual features and space from a new dimension, and finds that the spatial heterogeneity of visual features is universal, and different spatial statistical scales will only affect the heterogeneity; (2) in the local analysis, to explore the main visual features of the local area, this paper selects the Bund area in Huangpu District for detailed analysis. The research shows that buildings, street walls and street lamps are the main single visual elements, and “Building-Streetlight-Wall” and “Wall-Streetlight-Floor” are the main visual combination features in areas with historical buildings and small-scale streets as the main styles. The research shows that buildings, street walls and street lamps are the main single visual elements, and “Building-Streetlight-Wall” and “Wall-Streetlight-Floor” are the main visual combination features in areas with historical buildings and small-scale streets as the main styles.

Similar to most researches on street view images, the proportion of trees, high-rise buildings, sky and street walls are the main influencing factors when identifying the elements that reflect the visual features of the community. At all these scales, the results show that trees, high-rise buildings and street walls have higher contribution in visual features. Although the sky proportion does not have high local variance contribution, the stability of changes of its load factor indicates that it may also be one of the main factors affecting community visual features.

5. Conclusions

In this paper, the method integrating machine learning, geographically weighted principal components analysis and spatial analysis is used to give full play to the advantages of different methods in landscape aesthetics research. Based on the research of Chao’s larger-scale statistical units [26], the smaller-scale statistical units of 100 m, 150 m and 200 m are demonstrated, and the combination mode of the first three principal components extracted after principal components analysis is tried to judge the dominant community visual features.

This study also has some limitations, which are reflected in the data sources and research methods: (1) Firstly, the limitation of street view picture data. Focusing on the spatial distribution of the dominant visual features of street view, we have not yet considered the variables other than street view images, such as sound, smell and temperature. Street view images also have their limitations, and scenes in residential areas, yards and parks have not been collected in the data. (2) Secondly, the limitation of principal components analysis. Principal components analysis mentioned in this paper is an objective quantitative method of built environment elements. Its results can affect pedestrians’ subjective perception experience, but cannot fully represent their subjective cognition. Dominant visual features mean that the proportion of principal components is high, but elements with high proportion of principal components do not mean that they also have rich practical significance. The model has not taken into account of the color and practical significance of elements. For example, although traffic lights have extremely small proportion of principal components, they are still the main visual centers of pedestrians and vehicles at intersections because of their guiding significance.

In the follow-up study, we will consider including images from social platforms, street view images taken by ourselves and city scene images obtained by other methods into the research to further expand the research samples. Seen from the research results of heterogeneity of visual features, street view images have good application scene in identifying the landscape composition pattern of the local space of the city. Future research should focus on the following two points: (1) At the macro level, trying to guide the cityscape zoning and planning in urban design based on the classification of elements in

street view images. (2) At the micro level, cooperating with the refined design and built environment element control in urban renewal practice.

Author Contributions: Conceptualization, Y.L.; Investigation, S.W.; Software, S.Y.; Supervision, Y.L.; Validation, Y.L. and S.W.; Visualization, S.Y.; Writing—original draft, S.Y.; Writing—review & editing, S.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Shanghai Planning Office of Philosophy and Social Science, grant number 2018BCK002.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: <https://api.map.baidu.com/panorama/v2?> (accessed on 30 March 2021).

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

References

1. Youngan, K.; John, R.H. Street Ego-hood: An Alternative Perspective of Measuring Neighborhood and Spatial Patterns of Crime. *J. Quant. Criminol.* **2020**, *36*, 29–66.
2. Friedmann, J. Place and Place-Making in Cities: A Global Perspective. *Plan. Theory Pract.* **2010**, *11*, 149–165. [[CrossRef](#)]
3. Liu, H.; Xu, Y.; Tang, J.; Deng, M.; Huang, J.; Yang, W.; Wu, F. Recognizing urban functional zones by a hierarchical fusion method considering landscape features and human activities. *Trans. GIS* **2020**, *24*, 1359–1381. [[CrossRef](#)]
4. Huang, Y.; Shao, S.; Lei, Y.; Tian, J. Identification of Urban Functional Zones Using Network Kernel Density Estimation and Kriging Interpolation. *J. Geomat.* **2019**, *44*, 14–18.
5. Zhang, X.; Du, S.; Wang, Q. Hierarchical Semantic Cognition for Urban Functional Zones with Vhr Satellite Images and Poi Data. *Isprs J. Photogramm. Remote Sens.* **2017**, *132*, 170–184. [[CrossRef](#)]
6. Yao, Y.; Li, X.; Liu, X.; Liu, P.; Liang, Z.; Zhang, J.; Mai, K. Sensing Spatial Distribution of Urban Land Use by Integrating Points-of-interest and Google Word2vec Model. *Int. J. Geogr. Inf. Sci.* **2016**, *31*, 825–848. [[CrossRef](#)]
7. Wei, L.; Luo, Y.; Wang, M.; Cai, Y.; Su, S.; Li, B.; Ji, H. Multiscale Identification of Urban Functional Polycentricity for Planning Implications: An Inte-grated Approach Using Geo-big Transport Data and Complex Network Modeling. *Habitat Int.* **2020**, *97*, 102134. [[CrossRef](#)]
8. Zhang, L.; Pei, T.; Chen, Y.; Song, C.; Liu, X. A Review of Urban Environmental Assessment Based on Street View Images. *J. Geo-Inf. Sci.* **2019**, *21*, 46–58.
9. Long, Y.; Tang, J. Large-scale Quantitative Measurement of the Quality of Urban Street Space: The Research Progress. *City Plan. Rev.* **2019**, *43*, 107–114.
10. Liu, L.; Wang, H. Application of Computer Vision in Urban Studies: Review and Prospect. *City Plan. Rev.* **2019**, *43*, 117–124.
11. Chester, H.; Lisa, A. Measuring Urban Streetscapes for Livability: A Review of Approaches. *Prof. Geogr.* **2016**, *68*, 149–158.
12. Zhang, L.; Ye, Y.; Zeng, W.; Chiaradia, A. Human-scale Quality on Streets: A Large-scale and Efficient Analytical Approach Based on Street View Images and New Urban Analytical Tools. *Urban Plan. Int.* **2019**, *34*, 18–27.
13. Dai, Z.; Hua, C. The Improvement of Street Space Quality Measurement Method Based on Streetscape. *Planners* **2019**, *35*, 57–63.
14. Naik, N.; Philipoom, J.; Raskar, R.; Hidalgo, C. Streetscore—Predicting the perceived safety of one million streetscapes. In Proceedings of the IEEE Conference on Computer Vision and pattern Recognition Workshops, Las Vegas, NV, USA, 27 June 2014; pp. 779–785.
15. Robinson, M.O. Measuring Urban Design: Metrics for Livable Places, By Reid Ewing and Otto Clemente. *J. Am. Plan. Assoc.* **2015**, *81*, 80–81. [[CrossRef](#)]
16. Long, Y.; Ye, Y. Human-scale Urban Form:measurements, performances, and Urban Planning & Design Interventions. *South Archit.* **2016**, *36*, 41–47.
17. Chen, Y.; Zhao, X. Research on Ground-floor Interfaces along Streets from the Perspective of Pedestrians: A Case Study of Huaihai Road in Shanghai. *City Plan. Rev.* **2014**, *38*, 24–31.
18. Xiong, W.; Yan, W.; Liu, X.; Ma, R. The Improvement of Spatial Quality on Beijing Historical Streets Based on People-oriented Observation. *Urban. Archit.* **2018**, *06*, 57–61.
19. Ewing, R.; Handy, S.L.; Brownson, R.C.; Clemente, O.; Winston, E. Identifying and Measuring Urban Design Qualities Related to Walkability. *J. Phys. Act. Health* **2006**, *3*, S223–S240. [[CrossRef](#)] [[PubMed](#)]
20. Brownson, R.C.; Hoehner, C.M.; Day, K.; Forsyth, A.; Sallis, J.F. Measuring the Built Environment for Physical Activity: State of the Science. *Am. J. Prev. Med.* **2009**, *36*, S99–S123. [[CrossRef](#)]
21. Michael, S. Learning to Make Liveable Cities. *J. Urban Des.* **2016**, *21*, 570–573.
22. Kang, Y.; Zhang, F.; Gao, S.; Lin, H.; Liu, Y. A Review of Urban Physical Environment Sensing Using Street View Imagery in Public Health Studies. *Annals Gis.* **2020**, *26*, 261–275. [[CrossRef](#)]

23. James, S.; Kevin, T.; Megan, F. Understanding Visual Engagement with Urban Street Edges Along Non-pedestrianised and Pedestrianised Streets Using Mobile Eye-tracking. *Sustainability* **2019**, *11*, 4251.
24. Philip, S.; Katja, S.; César, A.H. The Collaborative Image of the City: Mapping the Inequality of Urban Perception. *PLoS ONE* **2013**, *8*, e68400.
25. Li, H.; Antonio, P.; Desheng, L. Built Environment and Violent Crime: An Environmental Audit Approach Using Google Street View. *Comput. Environ. Urban Syst.* **2017**, *66*, 83–95.
26. Wu, C.; Peng, N.; Ma, X.; Li, S.; Rao, J. Assessing Multiscale Visual Appearance Characteristics of Neighbourhoods Using Geographically Weighted Principal Component Analysis in Shenzhen, China. *Comput. Environ. Urban Syst.* **2020**, *84*, 101547. [[CrossRef](#)]
27. Yao, Y.; Liang, Z.; Yuan, Z.; Liu, P.; Bie, Y.; Zhang, J.; Wang, R.; Wang, J.; Guan, Q. A human-machine adversarial scoring framework for urban perception assessment using street-view images. *Int. J. Geogr. Inf. Sci.* **2019**, *33*, 2363–2384. [[CrossRef](#)]
28. Fotheringham, A.S.; Brunson, C.; Charlton, M. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationship*; John Wiley & Sons: Hoboken, NJ, USA, 2002.
29. Lu, B.; Harris, P.; Charlton, M.; Brunson, C. The Gwmodel R Package: Further Topics for Exploring Spatial Heterogeneity Using Geo-graphically Weighted Models. *Geo-Spat. Inf. Sci.* **2014**, *17*, 85–101. [[CrossRef](#)]