Article

Urban Analytics and City Science

EPB: Urban Analytics and City Science
2022, Vol. 49(4) 1245–1265

DOI: 10.1177/23998083211059838
journals.sagepub.com/home/epb
SAGE

Interplay between auditory and visual environments in historic districts: A big data approach based on social media

Hui Xie and Yi He
Faculty of Architecture and Urban Planning, Chongqing University, Chongqing, China; Key Laboratory of New Technology for Construction of Cities in Mountain Area, Chongqing University, Chongqing, China. Email: xh@cqu.edu.cn

Xueying Wu
Department of Architecture and Civil Engineering, City University of Hong Kong, Hong Kong, 999077, China

Yi Lu
Department of Architecture and Civil Engineering, City University of Hong Kong, Hong Kong, 999077, China.

Abstract

Historic districts play a vital role in stimulating urban economic development, conserving regional culture, and enhancing public participation. Both auditory and visual environments, and the interplay between them, are critical to visitors' perception and evaluation of historic districts. However, most studies have explored either the auditory or visual environments separately. The handful of existing studies on audiovisual interaction were confined to laboratory environments, leading to limited external validity. Here, we performed a data-driven study of the features of auditory and visual environments and the interaction between them in 17 historic towns in China using posts containing soundscape-related keywords and streetscape photos from a popular Chinese social media platform. First, we found that the auditory environments in historic districts mainly consist of man-made sounds from folkloric activities, the sounds of street shop vendors, and natural sounds from running water and birds. Second, street greenery, spatial enclosure, and presence of pedestrian in visual environment are positively associated with emotional feedback of the soundscape. This study and others support the importance of studying the auditory and visual environments of historic districts in conjunction. The innovative methods used in this study can be used in further studies in the field.

Corresponding authors:
Department of Architecture and Civil Engineering, City University of Hong Kong, 83 Tat Chee Avenue, Kowloon Tong, Hong Kong, 999077, China.
Email: xueyingwu8-c@my.cityu.edu.hk

Department of Architecture and Civil Engineering, City University of Hong Kong, 83 Tat Chee Avenue, Kowloon Tong, Hong Kong, 999077, China.
Email: yilu24@cityu.edu.hk

Keywords
Audio–visual environments, historic districts, soundscape, auditory environments, visual environments

Introduction

Historic districts have attracted growing attention in both developed and developing countries in recent years (Fung et al., 2017). They are more than mere collections of distinctive historic buildings; they play an increasingly important role in economic, cultural, and societal respects (Xu et al., 2020). The adaptive reuse of historic buildings in these districts and the improvement of the relevant infrastructure can effectively boost tourism, drive economic growth and enhance real-estate development (Leichenko et al., 2001; Yung et al., 2014). Well-developed historic districts can also promote local cultural heritage and contribute to forging a cultural identity (Shipley and Snyder, 2013), provide numerous opportunities for public participation by local residents, and play an important role in enhancing the sense of belonging and social cohesion (Yung et al., 2014).

The rich and unique experience of the audio-visual environments in historic districts is vital to build a distinctive cultural atmosphere, attract visitors, and increase the districts' vitality (Gastal, 2020; Navarrete Escobedo, 2020). This aspect of the environment also has a significant effect on the protection and restoration of historic and cultural heritage (*N*. Li, 2010). Visually, the unique architectural styles of historic buildings and diverting streetscapes are important environmental elements that provide tourists with a rich experience (Y. Zhang, 2018). The auditory environment, including the sound of traditional folkloric and commercial activities, also affects the sensory experience of tourists. High-quality audio-visual environments offer tourists a fuller experience of local customs and improve their visitor satisfaction with historic districts, which can further increase the visitor flow and promote economic development (L. Lu et al., 2015; T. Zhou et al., 2017).

Visual and auditory environments are interlinked (Y. Liu et al., 2019). The perceived quality of visual environments is partially mediated by auditory environments, and vice versa. However, researchers have tended to discuss the two aspects separately. For example, studies that focus on visual environments have often investigated the changes in spatial patterns of historic districts or changes in streetscapes as a result of streets traffic expansion and building facade renovation (Oba and Iseki, 2020). Meanwhile, studies that focus on auditory environments have often explored the classification and distribution of sounds or visitor preferences regarding soundscapes in historic districts (Huang and Kang, 2015; Montazerolhodjah et al., 2019). Few studies have considered audio-visual environments simultaneously and explored their potential interaction.

Combined audio-visual studies of historic districts have been carried out, but they were mostly based on questionnaires or interviews, resulting in limited study areas (usually a few street blocks within one historic district) and sample sizes (usually no more than 500 participants) (Z. Zhou et al., 2014). Recently, however, the availability of social media data has provided us with new opportunities to cover more historic districts and survey more participants.

To address these research gaps, we explored the characteristics and relationships of the auditory and visual environments of 17 historic towns in southwest China. First, we collected the auditory and visual data from a large Chinese social media network, Sina Weibo. Second, the posts and street view image data were processed using data mining and machine learning techniques. Third, the characteristics of the audio-visual environments in those historic towns were then summarized, and the relationship between them was explored.

Literature review

Study of audio-visual environments in historic districts

"Historic district" generally refers to areas containing a group of buildings with traditional architectural features or local ethnic characteristics, and such districts have attracted much academic attention in many countries (Fung et al., 2017). Questionnaire surveys or structured interviews remain the main methods to collect subjective evaluations of audio-visual environments (Filippi, 2015; Jayantha and Yung, 2018; Zhai and Ng, 2013). Such data are often used to guide the design and planning of historic districts. For example, Han et al. (2011) explored city dwellers' preferences for audio-visual elements in different seasons and found that natural audio-visual elements were preferred in spring, summer, and autumn, whereas artificial audio-visual elements were preferred in winter. Elsewhere, Z. Zhou et al. (2014) used questionnaires to analyze the main factors affecting the audio-visual environments in historic districts, such as human activities, public sanitation, and landscaping and green areas.

Big data from social media platforms provide an alternative method to collect subjective evaluations of various aspects of the urban environment, although few such studies have focused on historic districts. Data from a range of social media platforms (e.g., Foursquare, Google Places, Twitter, and Instagram) can be used to assess different urban dynamics (Martí et al., 2019). For example, Donahue et al. (2018) found that the water features and green elements in urban parks had a positive relationship with the visit frequency and subjective evaluation of parks by analyzing text and photos posted by Flickr and Twitter users.

Impact of visual environments on human behavior and perception

Optimizing the visual environments in historic districts has long been the goal of designers. In recent years, quantitative research based on visual quality and the proportion of visual elements has been gradually introduced into this field (Ewing et al., 2006). It has been found that spatial enclosure and greenery have the greatest influence on people's behavior and perception (Badland et al., 2010; Haapakangas et al., 2020).

Spatial enclosure is mainly determined by building density, building height, and street width (Qaoud and Djamal, 2017). Yin and Wang (2016) calculated the proportion of buildings and sky in Google Street View images as a proxy for spatial enclosure and observed that spatial enclosure had a significant impact on walking behavior. In fact, spatial enclosure has a negative relationship with mental health. An excessive sense of spatial enclosure may lead to negative emotional conditions such as depression and anxiety (Wang et al., 2019).

Greenery, as an important visual element, plays a vital role in the quality of urban landscapes. Well-designed street greenery can improve the quality of life for urban residents (X. Li et al., 2015). Urban residents tend to walk and exercise more in places with more greenery, which indicates the important influence of this urban feature on their living habits and behavior patterns (Y. Lu and Ye, 2019). Furthermore, urban greenery has a significant impact on physical and mental health and can effectively reduce the risk of chronic diseases, accelerate the recovery of patients, and alleviate mental stress (Mitchell and Popham, 2008; Thompson et al., 2012).

Impact of soundscapes on human behavior and perception

The perceived quality of the auditory environment is not only affected by physical quantities such as sound pressure level and frequency but also related to factors such as scenery and setting and personal differences in social and cultural background. Therefore, the term *soundscape* has been advocated in recent studies (Yang and Kang, 2005; Yu and Kang, 2014). Soundscape refers to the auditory environment that is perceived, experienced, or understood by one person or a group of people in context (ISO, 2014). The soundscape plays several important roles in creating an atmosphere, shaping a sense of place, preserving historic and cultural value, and improving the overall quality of historic districts (Montazerolhodjah et al., 2019).

First, the soundscape of historic districts has a definite influence on individuals' psychological cognition and sense of belonging. A highly appealing, well-designed soundscape can create a unique atmosphere in a place, enabling psychological immersion in the environment (D. Zhang et al., 2013). It can also enhance visitors' psychological state, strengthening the connection between their own subjective consciousness and social activities (Huang and Kang, 2015).

Second, the soundscape of historic districts directly influences people's lifestyle and daily behavior. J. Liu et al. (2018) pointed out that the soundscape has a significant impact on pedestrians' walking routes and behavior patterns. A high-quality soundscape can effectively guide and organize pedestrians along favorable sightseeing routes and greatly improve their sensory experience. Studies in several historic districts of Xiguan and Guangzhou, China, found that typical soundscapes, such as the sounds of Cantonese opera, Cantonese folk songs, chess players in alleyways, and traditional workshops, reflected the daily activities of local residents and the cultural ecology of the city (Zhou, 2014).

Interplay of auditory and visual environments

Both auditory and visual environments, as well as the interplay between them, are critical to perception and evaluation of urban environments (Huang and Kang, 2015; Ren et al., 2018). Although visual elements are the dominant sensory input, they are related to, and interact with, the soundscape (Viollon et al., 2002).

Previous studies of audio-visual interaction mainly used questionnaires to explore the impact of various audio-visual environmental elements on the subjective satisfaction of visitors to a site. Forest parks and urban green spaces are common research settings. Some studies have found that natural landscapes combining natural sounds from birds, insects, and water can significantly enhance tourists' sense of immersion and satisfaction in forest parks (Y. Liu et al., 2019; Y. Liu et al., 2020). Visitors were more satisfied with the audio-visual environments in parks and in areas near water than with those in urban built environments (S. Zhang et al., 2019). The development of virtual reality provides a new technology to reproduce audio-visual environments in the laboratory. Using this method, it was found that both visual and auditory elements affected the participants' evaluation of urban environments, although not equally: the impact of the two types of elements on the overall satisfaction with urban environments was 76% and 24%, respectively (Jin Yong Jeon, 2020).

Research gaps

There are three major research gaps in the study of urban audio-visual environments. First, most studies analyzed the visual and auditory environments separately and ignored the interaction between them. This may have led to an oversimplified assessment of audio-visual environments. Second, existing studies have been confined to laboratory environments or have used subjective assessment methods such as traditional questionnaires or interviews, leading to a limited sample size. Today, big data from social media provide an opportunity to retrieve subjective evaluations of visual and auditory environments from large samples. Third, little is known about the major features of the visual and auditory environments in historic districts.

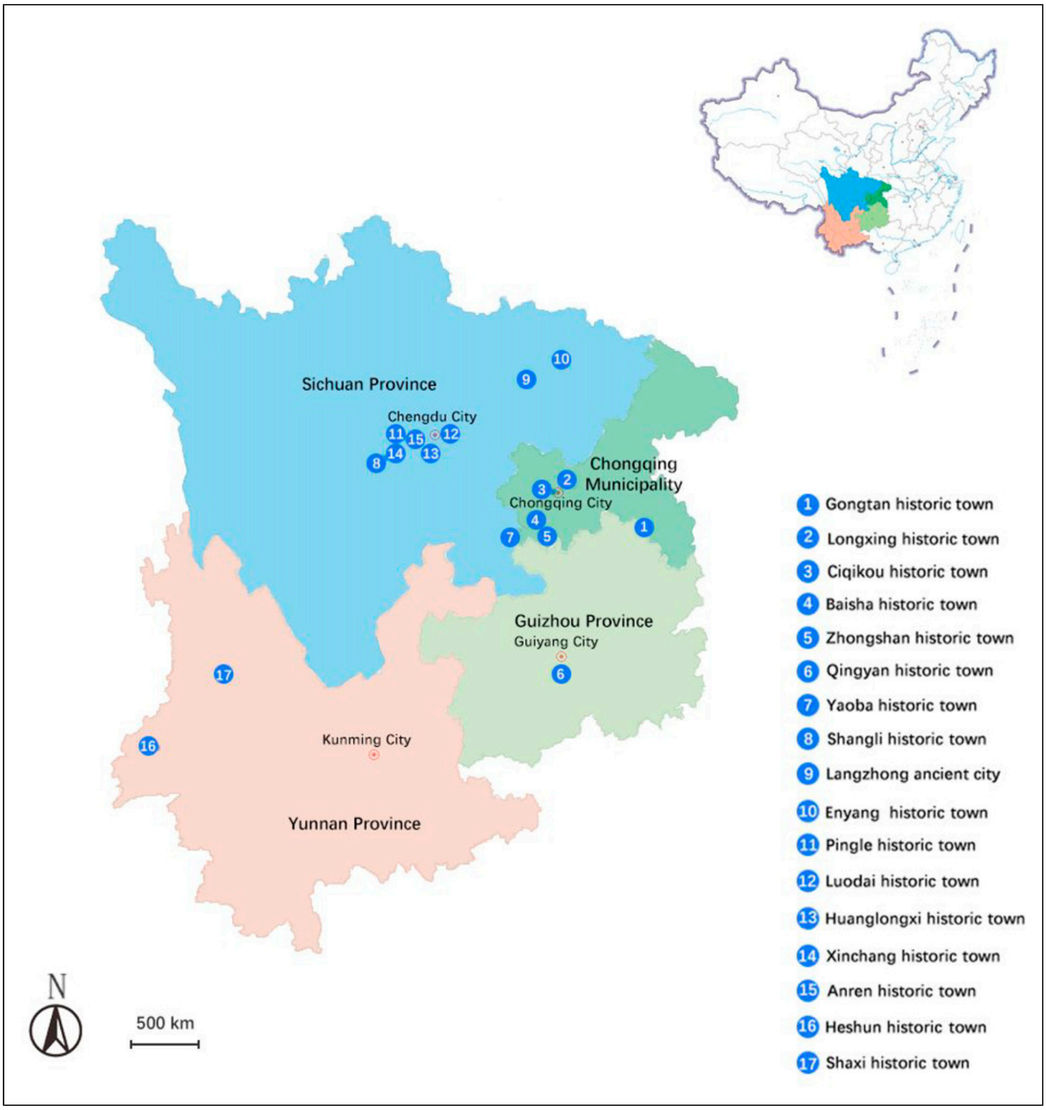

Figure 1. The locations of the 17 studied historic districts in southwest China.

Hence, in this study, we explored the features of the auditory and visual environments and the interaction between them in 17 historic towns in China, using posts that contained soundscape-related keywords and streetscape photos from Sina Weibo, a popular Chinese social media platform.

Method

Data

A total of 17 historic districts officially recognized by the Ministry of Housing and Urban-Rural Development and the State Administration of Cultural Heritage of China were selected as the study areas (Appendix I). All are in southwest China, specifically in Yunnan, Sichuan, Chongqing, and

Table 1. The number of posts and images in 17 studied historic districts.

	Soundscape text post			Street view image		
Localtion	Raw data	Retained data	Retaining ratio, %	Raw data	Retained data	Retaining ratio, %
Shaxi historic town	409	126	30.8	1469	287	19.5
Heshun historic town	504	85	16.9	923	299	32.4
Anren historic town	97	47	48.5	321	50	15.6
Xinchang historic town	60	45	75.0	115	61	53.0
Huanglongxi historic town	181	141	77.9	209	181	86.6
Luodai historic town	121	56	46.3	401	75	18.7
Pingle historic town	451	80	17.7	1021	72	7.1
Enyang historic town	96	29	30.2	332	112	33.7
Langzhong ancient city	237	117	49.4	394	168	42.6
Shangli historic town	318	37	11.6	713	60	8.4
Yaoba historic town	51	13	25.5	92	27	29.3
Qingyan historic town	515	106	20.6	2216	195	8.8
Zhongshan historic town	327	147	45.0	363	147	40.5
Baisha historic town	250	42	16.8	730	63	8.6
Ciqikou historic town	480	334	69.6	955	444	46.5
Longxing historic town	69	25	36.2	306	69	22.5
Gongtan historic town	179	47	26.3	413	85	20.6
Total	4345	1477	34.0	10,973	2395	21.8

Guizhou provinces (Figure 1). Buildings and streets are particularly well preserved in these historic districts. Folkloric activities remain in all of the studied historical towns, and attract large flow of visitors.

Sina Weibo has the most users of any social media platform in China. By March 2020, the number of active users had reached 550 million, with approximately 130 million posts and 120 million photos uploaded daily, providing a rich resource for big data research (Sina Weibo Data Center, 2018).

In this study, we extracted posts from 16 August 2009 to 28 February 2021 containing soundscape-related keywords together with location tags of the above-named historic districts using the web crawler software Locoyposter v7.3.1. The soundscape-related keywords included two types: 1) the noun words that directly describe specific sounds, such as the sound of running water, bird, and music and 2) the adjective words that describe how people feel about soundscape or sound environment, such as "quiet" and "noisy." Then, both the texts of the user-uploaded posts and corresponding streetscape photos were collected.

A total of 4345 posts and 10,973 images were retrieved in the raw dataset. This dataset was further filtered with two steps. First, all duplicated posts were excluded. Second, the posts that are not related to visitors' perceived soundscape, such as advertisements, were excluded based on expert judgment. Finally, 1477 posts containing soundscape-related keywords and 2395 associated images were selected (Table 1). The selected posts and images account for 34.0% of raw posts and 21.8% of raw images, respectively.

We also consider data privacy and ethic of using social media data. All retrieved data were posted by Sina Weibo users, and were publicly available online. The personal information of Sina Weibo

Table 2. Scores and corresponding mood levels.

Scores calculated by ROSTCM6	Level of mood
$x > 20$	Highly positive
$10 < x \leq 20$	Moderately positive
$0 < x \leq 10$	Mildly positive
$x = 0$	Neutral
$-10 \leq x < 0$	Mildly negative
$-20 \leq x < -10$	Moderately negative
$x < -20$	Highly negative

users, such as age and gender, was not collected to comply with privacy requirements. The collected data were further de-identified. Only the aggregate results were reported here in and hence cannot be traced back to individual Weibo user.

Text processing of soundscape-related posts

In this study, we mainly focused on analyzing the semantic adjectives from the collected posts. ROSTCM6 is a text analysis tool developed by Wuhan University (Zhan et al., 2018). We used the semantic network analysis and emotion analysis modules of this tool to evaluate the collected text.

The semantic network analysis module of ROSTCM6 identifies keywords or phrases that frequently co-occur at the semantic network level. The use frequencies of all possible keywords or phrase combinations are first counted. ROSTCM6 then visualizes keyword pairs of high frequency and generates a schematic diagram of the semantic network.

The emotion analysis module of ROSTCM6 works as follows. First, this function splits the text of each user's posts into equivalent semantic units in Chinese. Second, words or phrases that represent the user's emotion are extracted. Third, the corresponding scores of these words or phrases are retrieved from the built-in subjective emotion lexicon of ROSTCM6. In this way, each post obtains a single score, which represents the soundscape score, based on the average value of all emotional words in the post. Finally, the scores of the posts are automatically classified into seven mood categories—highly positive, moderately positive, mildly positive, neutral, mildly negative, moderately negative, and highly negative (Table 2).

Identification of street view image elements

Street view images uploaded by users can directly portray the visual environments that they experienced while visiting historic districts. We used pyramid scene parsing network (PSPNet) to identify different visual elements in the street view images uploaded by users (Zhao et al., 2017).

Nine visual elements, including streets, buildings, vegetation, sky, pedestrians, fence and traffic signs, natural terrain, cars, and bicycles, in the street view images were identified with the PSPNet extraction function as shown in Figure 2b. PSPNet can achieve a pixel-wise accuracy of 85% or better (Zhao et al., 2017). The proportion of each element in the images, calculated by using the number of identified pixels of each element divided by the total number of pixels of the whole image (a value ranges from 0 to 1), quantifies the visual environments that people see in the historic districts. In a pilot test with 50 randomly selected images from our dataset, nine visual elements were manually delineated by experts with Photoshop. The results of expert

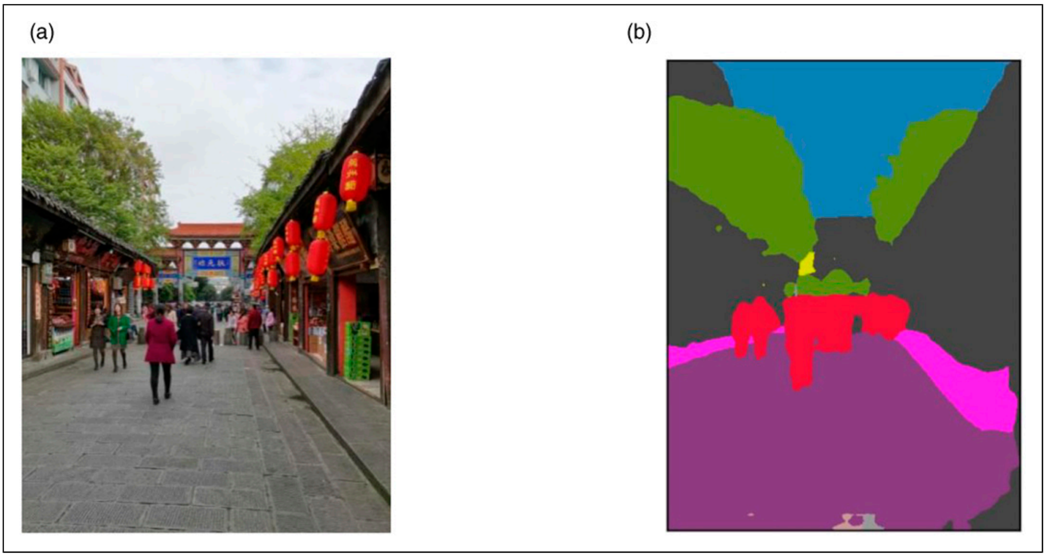

Figure 2. Visual element recognition in street view photos: (a) raw photo uploaded by Sina Weibo user and (b) the same photo segmented with PSPNet.

assessment were compared with the PSPNet results, and a pixel-wise accuracy of 87% was obtained.

Modeling the soundscape score with visual elements

To explore how visual and auditory environments interact with each other, we used nine visual elements extracted from the uploaded images to model the corresponding soundscape scores. In addition to the visual elements, the soundscape may be affected by the characteristics of different historic towns, such as the different location, town area, the level of economy, or the basic traffic condition. For example, the more population in a historic town, the more sounds it may produce and then may affect the soundscape. Similarly, the distance between a historic town and a tourist transportation station may also affect its soundscape. In a nutshell, both the visual environment observed by the individual and the characteristics of the historical towns may influence the soundscape score.

In order to control the impact of differences between the 17 historic towns, in this study, five town-level variables were added into the regression model, including town area, town population, the GDP of city where the historic town is located, distance to urban center, and distance to bus station. All independent and dependent variables are listed in Table 3.

Hierarchical linear modeling (HLM) was frequently conducted when the data are at varying hierarchical levels and different groups (Woltman et al., 2012). In this study, the individual posts and related photos (i.e., soundscape score and visual elements) are nested in different historical towns; hence, HLM was adopted to explore the relationship between soundscape and visual environment elements while considering the hierarchical structure of our dataset.

Two steps were included in HLM analysis. First, a regression model at individual level with random intercept was conducted to examine if the soundscape score was significantly associated with different visual elements, including streets, building, vegetation, sky, pedestrians, fence and traffic signs, natural terrain, cars, and bicycles. Second, the town-level variables, including town area, town population, GDP, distance to urban center, and distance to bus station, were added into regression model to explore the relations between soundscape score and them.

Table 3. Definition of the independent and dependent variables.

		Definition and calculation	Unit
Dependent variable			
Soundscape score		The scores calculated by ROSTCM6 using the uploaded text	Number
Independent variables			
Individual level	Streets Buildings Vegetation Sky Pedestrians Fence and traffic signs Natural terrain Cars Bicycles	The number of identified pixels of each element divided by the total pixels of the images	Ratio
Town level	Town area	The total area of each historic town	Km^2
	Town population	The total population in each historic town	10 thousand
	GDP	The GDP of city where the historic town is located	100 million
	Distance to urban center	The Euclidean distance from historic town to nearest urban center	Km
	Distance to bus station	The Euclidean distance from historic town to nearest bus station (tourist terminal or bus station)	Km

The model for individual level is as following

$$y_{ij} = \beta_{0j} + \sum_{q=1}^{9} \beta_{qj} X_{qij} + \varepsilon_{ij}$$

where

y_{ij} is the soundscape score;
β_{qj} (q=0, 1, …,9) are individual-level coefficients;
X_{qij} is the individual-level variable q for case i in town j; and
ε_{ij} is the individual-level random error.

The town level model is as following

$$\beta_{0j} = \gamma_{00} + \sum_{s=1}^{5} \gamma_{0s} W_{sj} + u_{0j}$$

where

γ_{0s} (s=0, 1, …, 5) are town-level coefficients;
W_{sj} is the town-level variable; and
u_{0j} is the town-level random error.

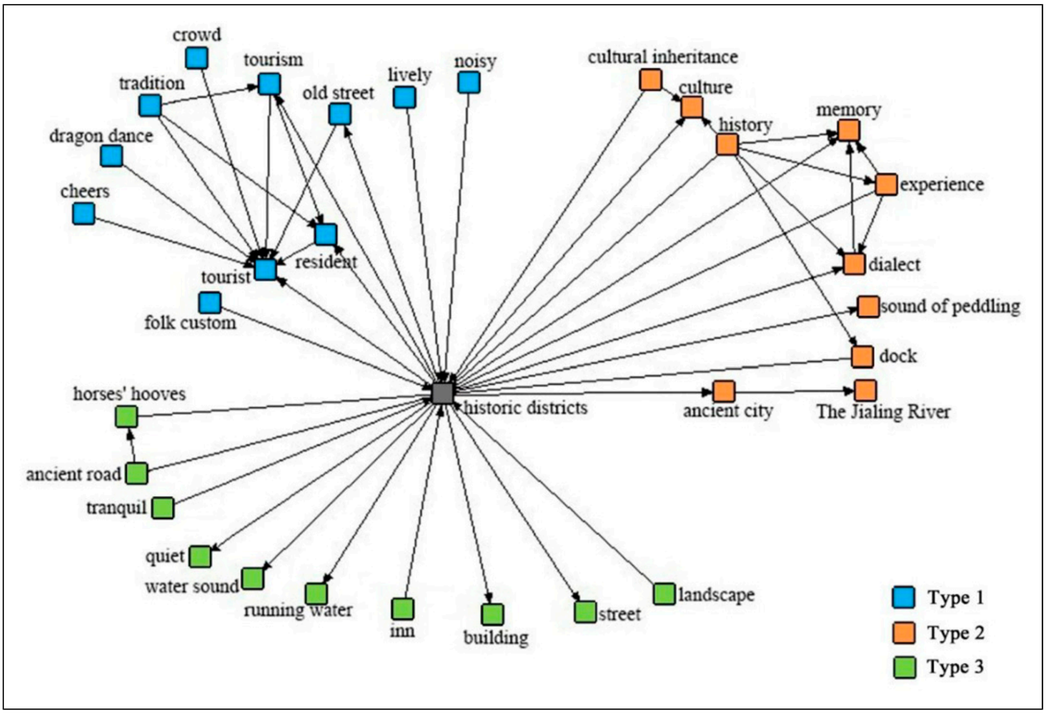

Figure 3. Semantic network analysis diagram of soundscape-related text elements.

Results

Types of soundscape

Figure 3 shows the semantic network of the sound keyword in the posts for our 17 historic districts, with each arrow linking two highly associated keywords. The semantic network analysis revealed three major clusters of soundscape keywords, that is, three types of soundscape. They are sounds related to traditional folkloric activities, street shop vendors, and natural sounds, which account for 33.4%, 34.5%, and 32.1% of all soundscape keywords, respectively. The classification is in line with an international standard for soundscape (ISO, 2016). Figure 4 presents the proportion of three soundscape types in 17 historic towns.

Type 1 mainly consists of man-made sounds produced by traditional folkloric activities, such as the dragon dance and the cheers of the audience. This type of sound is mainly found in historic towns in Sichuan province, and accounts for 45.6% of sound keywords in the towns in Sichuan. Type 2 mainly consists of man-made sounds produced by street shop vendors, such as shop assistants speaking in the local dialects. This type of sound is mainly found in historic towns in Chongqing and Guizhou province, and accounts for 75.2% of keywords. Type 3 mainly consists of natural sounds, such as running water or horses' hooves. This type of sound is mainly found in historic towns in Yunnan province, and accounts for 51.7% of keywords.

Perceived quality of soundscape

We analyzed the subjective evaluation of soundscapes using the emotion analysis module in ROSTCM6. Figure 5 shows the proportion of soundscape with different mood levels in each historic town.

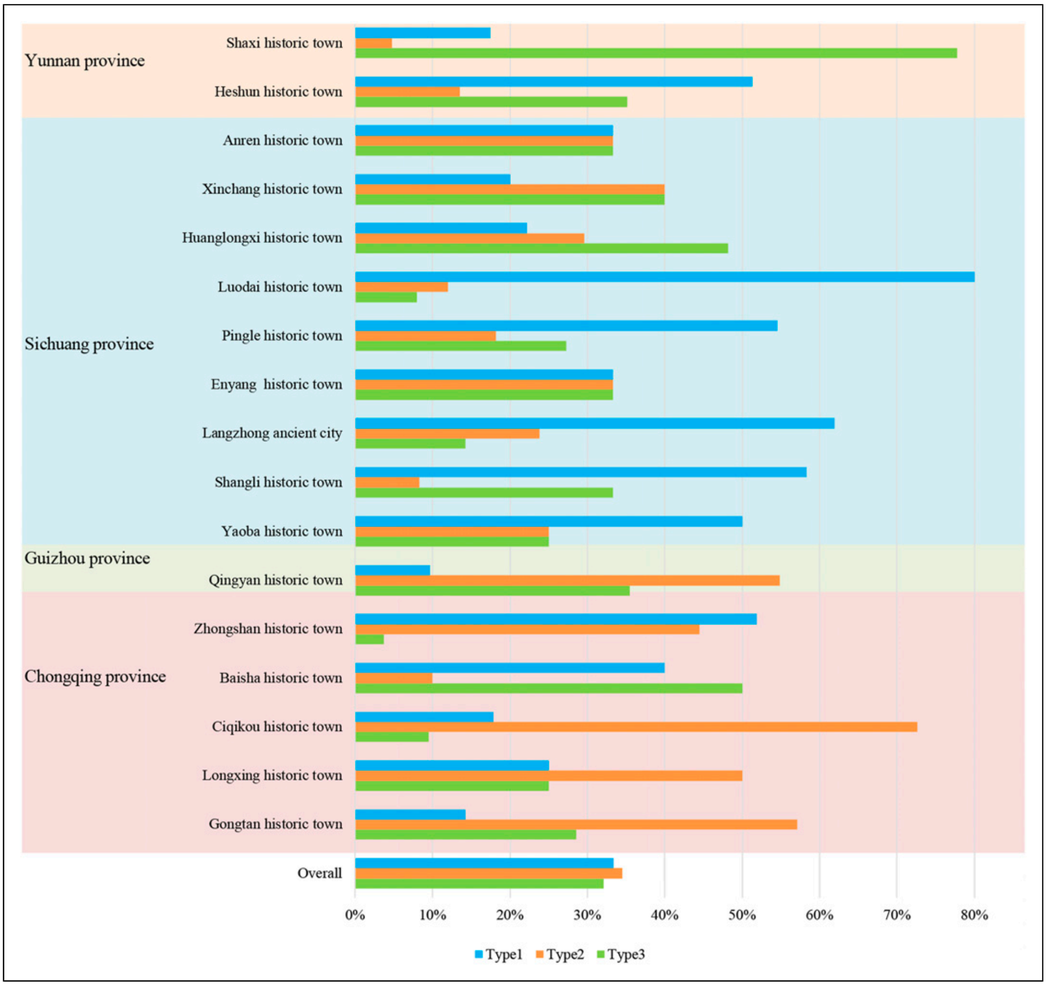

Figure 4. Proportion of soundscape types in each historic town.

Historic towns in Yunnan province received the highest proportion of positive emotional feedback. For example, the proportion of positive emotions in Heshun historic town was as high as 90.43%, among which the proportion of highly positive emotions was 42.55%. The historic towns in Sichuan were classified into a range of levels. Langzhong ancient city had the highest evaluation, with positive emotional feedback reported by 79.31% of users, whereas that of Enyang had a relatively poor evaluation, with positive emotional feedback from less than 60% of users. The historic towns in Chongqing and Guizhou were generally well appraised. Positive emotional feedback was given by more than 60% of users, except in Longxing historic town in Chongqing.

The average soundscape scores calculated from the users' posts in the same historic towns are shown in Figure 6. The soundscape scores in Yunnan were relatively high, ranging from a maximum of 23.44 to a minimum of 14.05. This can most likely be attributed to the natural sounds, such as running water, that are prevalent in Yunnan, which make people feel more comfortable and connected to the environment than man-made sounds. However, the soundscape scores in Sichuan showed a much wider range, with the highest score of 14.33 in Yaoba historic town and the lowest score of 7.5 in Xinchang and Shangli historic town. The soundscapes in this area are mainly characterized by

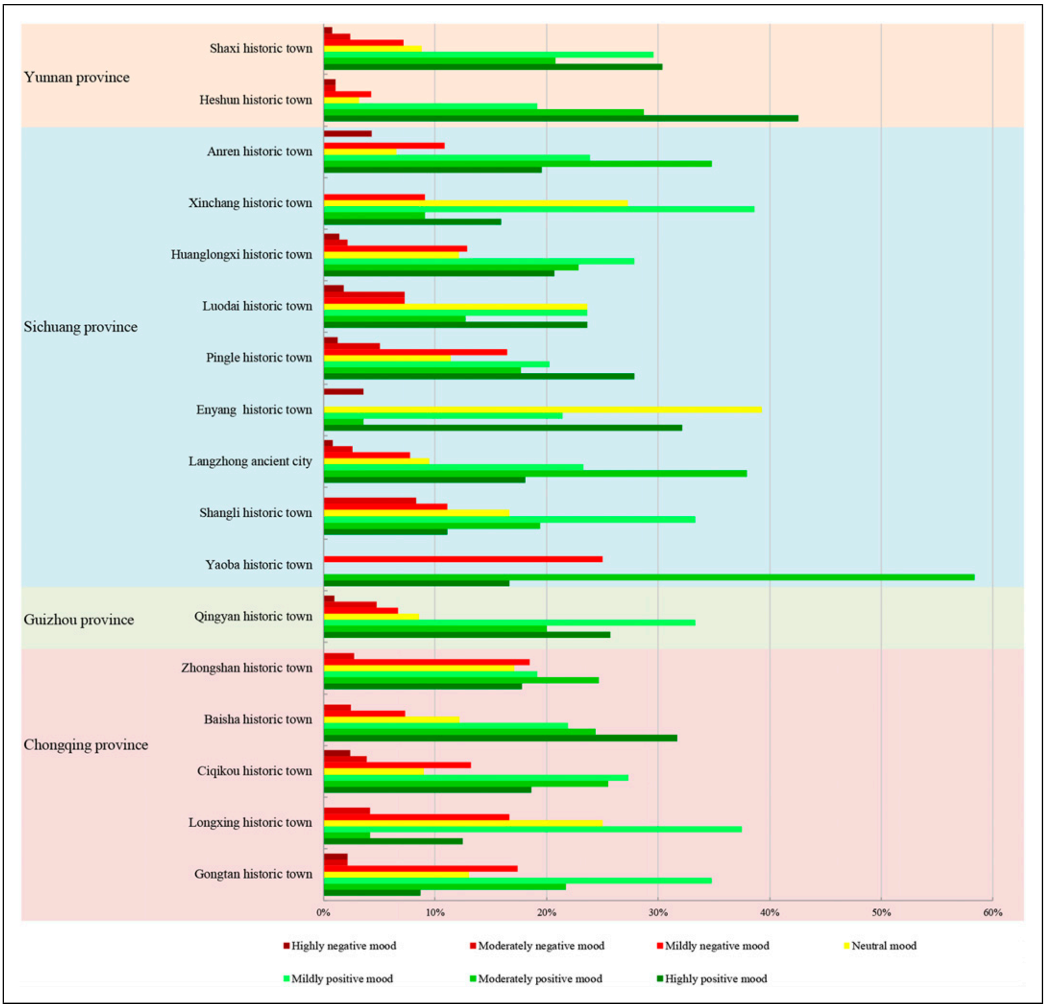

Figure 5. Results of soundscape subjective evaluation.

traditional folk activities, but because the area receives tourists from all over the country with a diverse spectrum of backgrounds, the local folk traditions may not fully resonate with all of them. Therefore, tourists have mixed opinions on this type of soundscape. The situation in Chongqing and Guizhou was similar to Sichuan, where the soundscape scores also varied considerably.

Streetscape images analysis

The calculated proportions of the nine elements (see *Method*) in the street view images of the 17 historic towns are provided in Table 3. All the proportions of visual elements in Table 4 are the average value in each historic town.

Little difference was found between the historic towns in terms of the proportion of streets, which fell within the range of 9%–16.4%. However, the proportion of buildings varied widely, implying significant differences in the sense of enclosure that would be felt by visitors. For example, the proportions of building in historic towns in Sichuan and Chongqing, such as those of Huanglongxi,

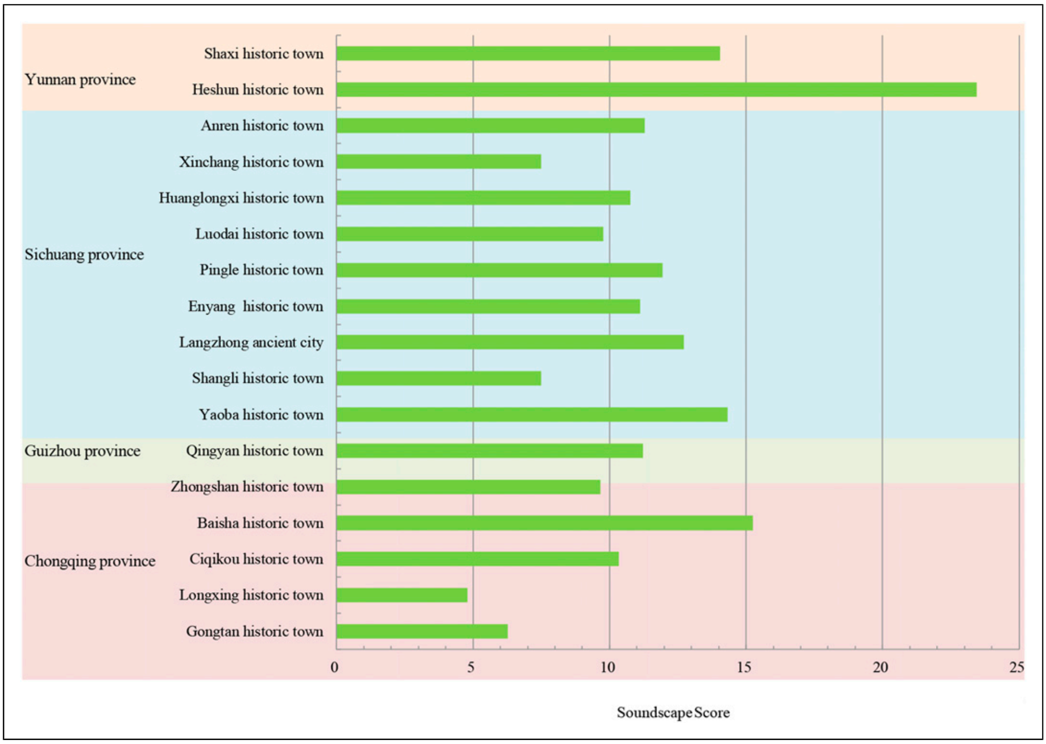

Figure 6. Average soundscape scores.

Pingle, and Longxing, were all above 60%, whereas the proportions of sky were all below 10%. This indicated that these three ancient towns had relatively high building density, narrow streets, and thus a strong sense of enclosure. The two historic towns in Yunnan showed the opposite characteristics, with buildings taking up less than 45% of the street view images and the sky taking up more than 10%. This implies that these towns would evoke much less sense of enclosure.

In terms of greenery, historic towns in Sichuan had the least, as represented by Huanglongxi and Pingle towns, whose greenery accounted for only 6.1% and 3.6%, respectively. Meanwhile, the proportions of pedestrians in the historic towns of Pingle, Langzhong, Zhongshan, and Ciqikou were all above 9%, which showed that all four attracted a large number of tourists and enjoyed relatively high vitality and popularity as areas for walking.

Fence and traffic signs, natural terrain, cars, and bicycles made up a relatively small proportion of the street images in all historic towns. The corresponding proportions were no more than 6.8%, 1.9%, 5.3%, and 0.5%, respectively, indicating their minor influence on the visual environment.

Regression analysis between soundscape score and visual elements

To explore how visual and auditory environments interact with each other, we conducted a HLM analysis between the soundscape scores of each user and the corresponding visual elements extracted from the uploaded images. The results of HLM analysis are shown in Table 5. The building and town population were excluded from the independent variables due to the collinearity problem. The remaining independent variables all passed the VIF test (VIF $<$ 4).

Table 4. Proportion of visual elements.

Name of historic town	Streets, %	Building, %	Vegetation, %	Sky, %	Pedestrian, %	Fence and traffic sign, %	Natural terrain, %	Car, %	Bicycle, %
Shaxi historic town	13.7	44.4	20.8	10.1	2.8	2.8	1.1	4.1	0.2
Heshun historic town	11.3	36.7	25.6	15.1	1.1	5.8	1.6	2.8	0.1
Anren historic town	12.6	40.6	25.1	5.8	5.6	4.2	0.7	5.3	0.0
Xinchang historic town	11.2	49.7	18.4	7.3	5.9	3.4	0.9	3.0	0.1
Huanglongxi historic town	12.7	69.6	6.1	6.4	1.5	2.2	0.9	0.3	0.2
Luodai historic town	10.8	48.3	13.7	11.3	8.5	3.6	0.3	3.4	0.1
Pingle historic town	10.1	63.2	3.6	6.4	11.9	3.1	0.1	1.5	0.2
Enyang historic town	15.2	34.5	27.9	7.0	7.2	5.5	1.4	1.2	0.1
Langzhong ancient city	12.3	41.0	16.2	10.1	9.2	6.5	0.3	4.0	0.2
Shangli historic town	11.6	41.2	23.0	8.4	5.6	6.8	0.8	2.5	0.1
Yaoba historic town	14.0	32.6	28.4	9.8	3.9	5.3	1.6	4.0	0.5
Qingyan historic town	10.3	39.9	25.9	15.6	3.1	2.6	0.8	1.6	0.1
Zhongshan historic town	13.6	40.6	24.3	5.3	10.5	3.1	1.2	1.3	0.1
Baisha historic town	16.4	44.2	16.6	15.4	3.4	2.2	1.2	0.5	0.1
Ciqikou historic town	9.0	52.7	15.8	7.4	9.3	2.7	0.3	2.6	0.1
Longxing historic town	12.3	63.8	10.2	8.0	2.4	2.2	0.4	0.8	0.1
Gongtan historic town	12.8	43.0	21.1	9.2	4.0	3.6	1.9	3.9	0.4

Table 5. Regression results of HLM.

Independent variables		Model 1 (standardized coefficients)		Model 2 (standardized coefficients)	
		B (95% CI)	*p*-Value	B (95% CI)	*p*-Value
Individual level	Streets	0.007 (−0.067, 0.082)	0.847	0.005 (−0.069, 0.079)	0.890
	Fence	0.073 (−0.004, 0.150)	0.064	0.064 (−0.013, 0.141)	0.105
	Vegetation	0.094 (0.017, 0.171)	0.016 *	0.092 (0.015, 0.169)	0.019 *
	Natural terrain	−0.007 (−0.081, 0.066)	0.843	−0.011 (−0.085, 0.063)	0.770
	Sky	−0.086 (−0.163, −0.008)	0.030 *	−0.085 (−0.162, −0.008)	0.031 *
	Pedestrian	0.147 (0.066, 0.228)	<0.001 ***	0.144 (0.064, 0.225)	<0.001 ***
	Cars	0.037 (−0.037, 0.110)	0.329	0.036 (−0.037, 0.110)	0.331
	Bicycle	−0.033 (−0.106, 0.040)	0.378	−0.031 (−0.104, 0.042)	0.405
Town level	Town area			−0.237 (−0.489, 0.015)	0.062
	GDP			−0.061 (−0.192, 0.070)	0.335
	Distance to urban center			−0.103 (−0.286, 0.079)	0.230
	Distance to bus station			0.272 (−0.013, 0.558)	0.059

Note: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

The proportions of vegetation and pedestrian were positively associated with the soundscape score in both Model 1 and Model 2. The proportion of sky was negatively significantly associated with soundscape score in both models. The rest of visual elements and town-level variables were not significant.

Discussion

Appealing audio–visual environments in historic districts can attract visitors and sustain the districts' vitality by engendering a distinctive cultural atmosphere. However, existing studies have tended to explore the auditory and visual environments independently, whereas most studies that have focused on audio–visual environments in historic districts were based on questionnaires or structured interviews with limited sample size. In this paper, we explored the interactions between auditory and visual environments in 17 historic districts in China using posts from Sina Weibo containing soundscape-related keywords and photos of the corresponding districts. Three distinct types of soundscape were found to prevail among these historic towns. The characteristic features and the relationships between auditory and visual environments were summarized. Specifically, four major findings emerged.

First, the three types of soundscape in the historic districts were characterized by sounds from folkloric activities, sounds from shop vendors, and natural sounds, respectively. Previous studies of soundscapes have paid most attention to the positive impact of natural sounds (Alvarsson et al., 2010). Running water, birdsong, and other natural sounds are frequently considered as a way to offset unwanted noise in soundscape design, whereas the potential role of the sounds of human activities has been neglected (Abdalrahman and Galbrun, 2020). However, we found that the unique soundscapes in our historic towns carried a distinct flavor relating to traditional human activities. The retention of man-made sounds (e.g., local dialects and folkloric activities) can make a distinct and vivid impression on tourists and enhance the appeal of historic districts. Preserving the authentic, unique auditory environments of historic districts should be taken into consideration during cultural renovation of these areas.

Second, street greenery in visual environment has a positive effect on perceived soundscape scores in historic districts. This is consistent with previous findings that street greenery may effectively improve the perceived quality of an environment (Haapakangas et al., 2020). Indeed, urban greenery has well-established physical and mental benefits, including reduced risk of heart disease and diabetes, lower rates of obesity, improved mental well-being, and decreased stress and depression (James et al., 2015; Triguero-Mas et al., 2015; Twohig-Bennett and Jones, 2018). Greenery can also effectively reduce air pollution and provide shade for pedestrians (Markevych et al., 2017). In addition, the presence of street greenery reduces the perceived noise (Haapakangas et al., 2020). All these benefits may encourage positive emotional evaluations of historic districts.

Third, contrary to our expectation, pedestrians in visual environment were positively associated with perceived soundscape scores. We originally thought that large crowds may lead to a noisy environment. This can be explained by Jane Jacobs' urban vitality theory, which emphasized the importance of pedestrian activities to maintain the urban vitality (Jacobs, 1963). Large number of pedestrians and related human sounds indicate a high-level vitality and popularity of historic towns (Hong and Jeon, 2015). Besides, as these photos were uploaded by tourists in historic towns, these photos are more likely to include selfies and other tourists taken in some favorable locations, which may also result in positive relation between pedestrians and soundscape scores.

Fourth, the level of spatial enclosure, which is in inverse relationship with proportion of sky, may have a positive effect on the soundscape. One previous study found that the space enclosure is positively associated with the willingness of walking due to better perceived pedestrian enjoinment (Yin and Wang, 2016). The most distinctive and representative feature in the historic towns is historical buildings with various styles. Hence, in historic districts, high space enclosure is usually associated with narrow streets and dense historical building, which leads to higher urban vitality and better perceived soundscapes.

This paper has both strengths and limitations. We collected crowdsourced data from a social media platform rather than using a traditional questionnaire, which granted access to a much larger number of samples with low cost and fast speed. Data mining technology can process text-based posts efficiently. Machine learning was used to automatically identify streetscape elements with satisfactory accuracy. These innovative methods can inspire further studies about the perceived quality of historical towns or other built environments.

However, this study also has limitations. First, the content of uploaded images and the emotion that the texts trying to convey may be biased in some cases. For instance, the text description may be a description of the overall visual and auditory environments, but the uploaded images may only serve as a selfie of the visitor or only involve a partial visual scene. Second, social media platforms are not popular among the elderly. Young participants may have been oversampled in this study. Hence, our results may not be generalizable to the whole population. Third, it is difficult to extract personal information (e.g., local residents vs. visitors and sociodemographic information) from Weibo data. Thus, we cannot identity whether different user groups may have the different emotional feedbacks for visiting a same historic town. Furthermore, the data we collected range over 10 years. There are possibly temporal changes in sound and visual environment over such time span. In addition, the data are subject to recall bias because the social media posts from tourists may have been made many days after their visits to these historic sites.

Conclusion

In this paper, we innovatively explored the interaction between auditory and visual environments in 17 historic towns in China using user-uploaded posts and associated photos from Sina Weibo. We found that the three main types of auditory environment were characterized by artificial sound from folk activities, street shop vendors, and natural sound. More importantly, the visual environments affected the perceived quality of the soundscapes. The street greenery, pedestrian, and sense of spatial enclosure had a positive effect on the perceived soundscape quality. These findings point out that it is critical to focus on the interplay between the auditory and visual environments of historic districts in future studies.

Acknowledgments

The work described in this paper was fully supported by grants from the Key Laboratory of New Technology for Construction of Cities in Mountain Area, Chongqing University (No. LNTCCMA-20200101).

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The work described in this paper was fully supported by grants from the Key Laboratory of New Technology for Construction of Cities in Mountain Area, Chongqing University (No. LNTCCMA-20200101).

ORCID iD

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

References

Abdalrahman Z and Galbrun L (2020) Audio-visual preferences, perception, and use of water features in open-plan offices. *The Journal of the Acoustical Society of America* 147(3): 1661–1672. DOI: 10.1121/10.0000892.

Alvarsson JJ, Wiens S and Nilsson ME (2010) Stress recovery during exposure to nature sound and environmental noise. *International Journal of Environmental Research and Public Health* 7(3): 1036–1046. DOI: 10.3390/ijerph7031036.

Badland HM, Opit S, Witten K, et al. (2010) Can virtual streetscape audits reliably replace physical streetscape audits? *Journal of Urban Health-Bulletin of the New York Academy of Medicine* 87(6): 1007–1016. DOI: 10.1007/s11524-010-9505-x.

Donahue ML, Keeler BL, Wood SA, et al. (2018) Using social media to understand drivers of urban park visitation in the Twin Cities, MN. *Landscape and Urban Planning* 175: 1–10. DOI: 10.1016/j.landurbplan.2018.02.006.

Ewing R, Handy S, Brownson RC, et al. (2006) Identifying and measuring urban design qualities related to walkability. *Journal of Physical Activity and Health* 3(s1): S223–S240. DOI: 10.1123/jpah.3.s1.s223.

Filippi M (2015) Remarks on the green retrofitting of historic buildings in Italy. *Energy and Buildings* 95: 15–22. DOI: 10.1016/j.enbuild.2014.11.001.

Fung IWH, Tsang YT, Tam VWY, et al. (2017) A review on historic building conservation: A comparison between Hong Kong and Macau systems. *Renewable and Sustainable Energy Reviews* 71: 927–942. DOI: 10.1016/j.rser.2016.12.121.

Gastal SdA (2020) *Attracting visitors to ancient neighbourhoods. Creation and management of the tourist historic city of Plymouth, UK*. Plymouth, UK: Periplo Sustentable, 38, pp. 69–73.

Haapakangas A, Hongisto V and Oliva D (2020) Audio-visual interaction in perception of industrial plants-Effects of sound level and the degree of visual masking by vegetation. *Applied Acoustics* 160: 107121. DOI: 10.1016/j.apacoust.2019.107121.

Han M-H, Joo M-K, Kim J-T, et al. (2011) Seasonal preferences and interactions of visual/auditory elements of streetscape in Namwon City, Korea. *Indoor and Built Environment* 20(1): 162–170. DOI: 10.1177/1420326x10394481.

Hong JY and Jeon JY (2015) Influence of urban contexts on soundscape perceptions: A structural equation modeling approach. *Landscape and Urban Planning* 141: 78–87. DOI: 10.1016/j.landurbplan.2015.05.004.

Huang L and Kang J (2015) The sound environment and soundscape preservation in historic city centres-the case study of Lhasa. *Environment and Planning B-Planning and Design* 42(4): 652–674. DOI: 10.1068/b130073p.

Jacobs J (1963) *The Death and Life of Great American Cities*. New York: Vintage.

James P, Banay RF, Hart JE, et al. (2015) A review of the health benefits of greenness. *Current Epidemiology Reports* 2(2): 131–142.

Jayantha WM and Yung EHK (2018) Effect of revitalisation of historic buildings on retail shop values in urban renewal: an empirical analysis. *Sustainability* 10(5): 1418. DOI: 10.3390/su10051418.

Jin Yong Jeon HIJ (2020) Effects of audio-visual interactions on soundscape and landscape perception and their influence on satisfaction with the urban environment. *BMC Ecology* 169: 106544. DOI: 10.1016/j.buildenv.2019.106544.

Leichenko RM, Coulson NE and Listokin D (2001) Historic preservation and residential property values: An analysis of Texas cities. *Urban Studies* 38(11): 1973–1987. DOI: 10.1080/00420980120080880.

Li N (2010) Preserving Urban landscapes as public history: the Chinese context. *Public Historian* 32(4): 51–61. DOI: 10.1525/tph.2010.32.4.51.

Li X, Zhang C, Li W, et al. (2015) Assessing street-level urban greenery using google street view and a modified green view index. *Urban Forestry and Urban Greening* 14(3): 675–685. DOI: 10.1016/j.ufug.2015.06.006.

Liu J, Xiong Y, Wang Y, et al. (2018) Soundscape effects on visiting experience in city park: A case study in Fuzhou, China. *Urban Forestry and Urban Greening* 31: 38–47. DOI: 10.1016/j.ufug.2018.01.022.

Liu Y, Hu M and Zhao B (2019) Audio-visual interactive evaluation of the forest landscape based on eye-tracking experiments. *Urban Forestry and Urban Greening* 46: 126476. DOI: 10.1016/j.ufug.2019.126476.

Liu Y, Hu M and Zhao B (2020) Interactions between forest landscape elements and eye movement behavior under audio-visual integrated conditions. *Journal of Forest Research* 25(1): 21–30. DOI: 10.1080/13416979.2019.1707341.

Lu L, Chi CG and Liu Y (2015) Authenticity, involvement, and image: Evaluating tourist experiences at historic districts. *Tourism Management* 50: 85–96. DOI: 10.1016/j.tourman.2015.01.026.

Lu Y and Ye Y (2019). Can people memorize multilevel building as volumetric map? A study of multilevel atrium building. *Environment and Planning B: Urban Analytics and City Science*, 46(2), 225–242 doi:doi: 10.1177/2399808317705659

Markevych I, Schoierer J, Hartig T, et al. (2017) Exploring pathways linking greenspace to health: Theoretical and methodological guidance. *Environmental Research* 158: 301–317. DOI: 10.1016/j.envres.2017.06.028.

Martí P, Serrano-Estrada L and Nolasco-Cirugeda A (2019) Social media data: challenges, opportunities and limitations in urban studies. *Computers, Environment and Urban Systems* 74: 161–174.

Mitchell R and Popham F (2008) Effect of exposure to natural environment on health inequalities: an observational population study. *Lancet* 372(9650): 1655–1660. DOI: 10.1016/s0140-6736(08)61689-x.

Montazerolhodjah M, Sharifnejad M and Montazerolhodjah M (2019) Soundscape preferences of tourists in historical urban open spaces. *International Journal of Tourism Cities* 5(3): 465–481. DOI: 10.1108/ijtc-08-2018-0065.

Navarrete Escobedo D (2020) Foreigners as gentrifiers and tourists in a Mexican historic district. *Urban Studies* 57: 3151–3168. DOI: 10.1177/0042098019896532.

Oba T and Iseki H (2020) Transportation impacts on cityscape preservation: spatial distribution and attributes of surface parking lots in the historic central districts. *Journal of Urban Planning and Development* 146(2): 1943–5444. DOI: 10.1061/(asce)up.1943-5444.0000559.

Qaoud R and Djamal A (2017) The impact of constructivism density of the urban tissue in improving the physical urban ambience of the free space-The street-for Saharian cities. The study of the case city of Biskra. In: Salame CT, Aillerie M and Papageorgas P (eds.), *International Conference on Technologies and Materials for Renewable Energy, Environment and Sustainability, Tmrees17*, 119, pp. 201–213.

Ren X, Kang J, Zhu P, et al. (2018) Effects of soundscape on rural landscape evaluations. *Environmental Impact Assessment Review* 70: 45–56. DOI: 10.1016/j.eiar.2018.03.003.

Shipley R and Snyder M (2013) The role of heritage conservation districts in achieving community economic development goals. *International Journal of Heritage Studies* 19(3): 304–321. DOI: 10.1080/13527258.2012.660886.

Sina Weibo Data Center (2018) *Weibo User Development Report*. Sina Technology (China) Co., Ltd. (Beijing: 2019). Available at: https://data.weibo.com/report/reportDetail?id=433

Thompson CW, Roe J, Aspinall P, et al. (2012) More green space is linked to less stress in deprived communities: Evidence from salivary cortisol patterns. *Landscape and Urban Planning* 105(3): 221–229. DOI: 10.1016/j.landurbplan.2011.12.015.

Triguero-Mas M, Dadvand P, Cirach M, et al. (2015) Natural outdoor environments and mental and physical health: Relationships and mechanisms. *Environment International* 77: 35–41. DOI: 10.1016/j.envint.2015.01.012.

Twohig-Bennett C and Jones A (2018) The health benefits of the great outdoors: A systematic review and meta-analysis of greenspace exposure and health outcomes. *Environmental Research* 166: 628–637. DOI: 10.1016/j.envres.2018.06.030.

Viollon S, Lavandier C and Drake C (2002) Influence of visual setting on sound ratings in an urban environment. *Applied Acoustics* 63(5): 493–511. DOI: 10.1016/s0003-682x(01)00053-6.
Wang R, Lu Y, Zhang J, et al. (2019) The relationship between visual enclosure for neighbourhood street walkability and elders' mental health in China: Using street view images. *Journal of Thrombosis and Haemostasis* 13: 90–102. DOI: 10.1016/j.jth.2019.02.009.
Woltman H, Feldstain A, MacKay JC, et al. (2012). An introduction to hierarchical linear modeling. *Tutorials in Quantitative Methods for Psychology* 8(1): 52–69.
Xu Y, Rollo J, Jones DS, et al. (2020) Towards sustainable heritage tourism: a space syntax-based analysis method to improve tourists' spatial cognition in chinese historic districts. *Buildings* 10(2): 29. DOI: 10.3390/buildings10020029.
Yang W and Kang J (2005) Acoustic comfort evaluation in urban open public spaces. *Applied Acoustics* 66(2): 211–229. DOI: 10.1016/j.apacoust.2004.07.011.
Yin L and Wang Z (2016) Measuring visual enclosure for street walkability: Using machine learning algorithms and google street view imagery. *Applied Geography* 76: 147–153. DOI: 10.1016/j.apgeog.2016.09.024.
Yu C-J and Kang J (2014) Soundscape in the sustainable living environment: A cross-cultural comparison between the UK and Taiwan. *Science of the Total Environment* 482: 501–509. DOI: 10.1016/j.scitotenv.2013.10.107.
Yung EHK, Chan EHW and Xu Y (2014) Community-initiated adaptive reuse of historic buildings and sustainable development in the inner city of Shanghai. *Journal of Urban Planning and Development* 140(3): 0000174. doi: 10.1061/(asce)up.1943-5444.0000174.
Zhai B and Ng MK (2013) Urban regeneration and social capital in China: A case study of the drum tower muslim district in Xi'an. *Cities* 35: 14–25. DOI: 10.1016/j.cities.2013.05.003.
Zhang D, Liu D, Cui Z, et al. (2013) Analysis of the present situation of research on soundscape and Han-Chinese buddhist temples. In: Hou H and Tian L (eds.), *Architecture, Building Materials and Engineering Management*, 388, pp. 357–360.
Zhang S, Zhao X, Zeng Z, et al. (2019) The influence of audio-visual interactions on psychological responses of young people in urban green areas: a case study in two parks in China. *International Journal of Environmental Research and Public Health* 16(10): 1845. DOI: 10.3390/ijerph16101845.
Zhan Y, Wu J and Du H (2018) Researching on Satisfaction of Tibetan Characteristic Agricultural Products Based on Consumer Comment from Jingdong Mall: Taking the Example of highland barley rice. *Proceedings of the International Symposium on Big Data and Artificial Intelligence* 29–32. doi: 10.1145/3305275.3305281, In press.
Zhang Y (2018) Negotiating authenticity in China's urban historic preservations-the case of the Kuan and Zhai Alleys in Chengdu. *Heritage and Society* 11(2): 79–104. DOI: 10.1080/2159032x.2019.1587968.
Zhao H, Shi J, Qi X, et al. (2017) Pyramid scene parsing network. In: 30th Ieee Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21-26 July 2017, pp. 6230–6239.
Zhou T, Zhou Y and Liu G (2017) Comparison of critical success paths for historic district renovation and redevelopment projects in China. *Habitat International* 67: 54–68. DOI: 10.1016/j.habitatint.2017.06.008.
Zhou M (2014) Research on Soundmarks of Soundscape at Historical District. *Applied Mechanics and Materials (Vols*. Trans Tech Publications, Ltd., 584–586, 521–524. Available at: https://doi.org/10.4028/www.scientific.net/amm.584-586.521.
Zhou Z, Kang J and Jin H (2014) Factors that influence soundscapes in historical areas. *Noise Control Engineering Journal* 62(2): 60–68. DOI: 10.3397/1/376206.

ISO (International Organization for Standardization). (2014) *Acoustics — Soundscape — Part 1: Definition and Conceptual Framework*. ISO/FDIS 12913-1:2014. Geneva: International Organization for Standardization.

ISO (International Organization for Standardization). (2016) *Acoustics – Soundscape – Part 2: Data Collection and Reporting Requirements*. ISO/CD 12913-2:2016(E). Geneva: International Organization for Standardization.

Author Biographies

Dr. Hui XIE obtained his first degree and MEng from Chongqing University, and his PhD from the School of Architecture, University of Sheffield. Before coming to Chongqing under the 100 Talents Plan in 2012, he worked as a Research Associate at Sheffield and Cambridge. His main research interests are the health effects of built environment, architectural acoustics, urban sound environment and soundscape. Till now, he has published more than 80 academic papers, along with 30+ international conference papers. He has been the principle investigators for 30+ research projects funded by China and UK's national research councils, including 3 NSFC projects, and has been a consultant for 30+ major acoustics and noise control projects in the UK, Middle East and China. He has been the key contributor to several national and industry standards.

Yi HE obtained his Master Degree from the Faculty of Architecture and Urban Planning Chongqing University in June 2021. His research interests focus on the audio-visual environment of historical district, especially the acoustic environment and the soundscape. His current research is to explore how the soundscape affects the subjective evaluation of tourists in the historical district, and how to improve the overall spatial quality based on the soundscape.

Xueying WU is s a PhD candidate in the Department of Architecture and Civil Engineering at the City University of Hong Kong. She has received the master degree in urban and rural planning in Harbin Institute of Technology in 2019. She has also received the "Outstanding Students" in HIT. Her current research is concerned with the relationship between different types of human behavior (including bicycle-transfer cycling behavior, protesting behavior, travel behavior under the covid-19 situation) and built environment in urban space.

Dr. Yi LU is Associate Professor in Department of Architecture and Civil Engineering at the City University of Hong Kong. He received his bachelor degree of architecture from Southeast University in Nanjing China, his master degree from National University of Singapore, and his PhD from Georgia Institute of Technology, USA. He practiced in architectural firms in Singapore and China for several years. Dr. LU is interested in the behavioral impact of built environment; such as how site design of an urban smart development project influence walking, how nursing unit design influence staff behavior. He has published in leading journals such as Environment and Behavior, Environment and Planning B. He has also developed rigorous spatial description of physical environment and related assessment tools.