https://doi.org/10.15302/J-LAF-1-020064

RECEIVED DATE | 2021-06-09

Designing Attention —Research on Landscape Experience Through Eye Tracking in Nanjing Road Pedestrian Mall (Street) in Shanghai

CHEN Yiyan[1], CHEN Zheng[1,2,*], DU Ming[3]

1 College of Architecture and Urban Planning, Tongji University, Shanghai 200092, China

2 Key Laboratory of Ecology and Energy-saving Study of Dense Habitat (Ministry of Education), Tongji University, Shanghai 200092, China

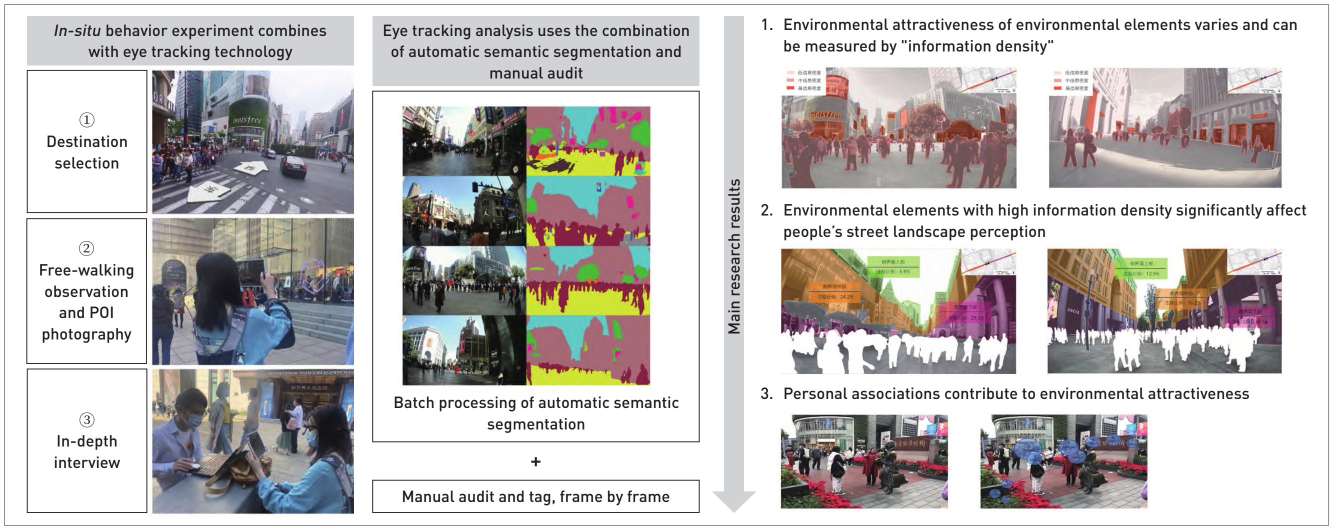
3 Tongji Architectural Design (Group) Co., Ltd., Shanghai 200092, China

*CORRESPONDING AUTHOR

Address: 1239 Siping Road, Yangpu District, Shanghai, China

Email: zhengchen@tongji.edu.cn

GRAPHICAL ABSTRACT



HIGHLIGHTS

- Eye movements help interpret cognition process that leads to experience and behavior
- Fixation content analysis combines automatic semantic segmentation with manual audit
- Information density measures the efficiency of environmental elements on attracting attentions
- Elements with high information density (e.g. business signs) play a larger role in experience

ABSTRACT

To understand the attention distribution and visual cognition in streets, this study conducted an experiment on Nanjing Road Pedestrian Mall (Street) using Head-mounted eye-trackers. Participants' attention distribution were analyzed via Area Of Interest (AOI) during real-life walking scenarios, combining with several experiment tasks (i.e., destination selection, point-of-interest photography, and in-depth interviews), to capture participants' naturalistic decision-making. The study combined automatic semantic segmentation with manual audit to code participants' attention fixation duration and proportion by AOIs. A new indicator "information density," which is the ratio of the attention percentage to the exposure percentage of a given environmental element, was introduced to describe the efficiency of environmental elements on attracting attentions. Findings revealed information density varies across environmental elements: higher information density was found in sign, building entrance, brand name, and poster; the lower was found in building, sky, and ground; while tree and person fell in between. Findings hence suggest environmental elements of higher information density (such as business signs) should be systematically designed to enhance desired experiences. Findings also indicated that personalized experiences are more likely to induce positive associations about environment which eventually lead to place attachment.

KEYWORDS

Street Landscape;
Nanjing Road Pedestrian Mall (Street);
Landscape Experience;
Head-mounted Eye-tracker;
In-situ Environmental Experience;
Spatial Perception;
Environmental Attractiveness

EDITED BY
Tina TIAN, WANG Moying
TRANSLATED BY
Tina TIAN, WANG Moying, GU Xiang

1 Introduction

Streets are not only an important part of urban traffic system, but also a major sort of place for social activities. With the rapid urban development, pedestrian malls/streets have gradually gone from simply providing commercial services toward magnifying specific landscape qualities and accommodating multiple activities, enriching citizens' social life[1].

Researchers have carried out cognitive experience analysis based on observational research on people's behavior and landscape experiences on streets, developing various research tools and scales to measure users' activities and behavioral psychology through behavioral observation, questionnaire, in-depth interview, etc. Research tools include Kevin Lynch's City Image and Its Five Elements and Cognitive Mapping[2][3], and methods include Jan Gehl's Public Space & Public Life Survey based on site surveys and behavioral observations[4][5] and Charles Egerton Osgood's semantic differential based on psycho-experiments[6]. With the development in Physiology and Psychology, innovative approaches such as emotion and experience capturing based on physiological data[7] and self-reported assessment with a focus on ecological validity[8]~[10] have been proposed. However, due to the immaterial nature of environmental experiences and feelings, although these studies measure users' experience in and emotional memory about a given space, they provide little help to support design practice from the perspective of spatial elements.

Eye tracking technology can accurately record the authentic movements of human eyes in millisecond responses, and reflect human rapid perception of the environment[11][12]. It helps researchers intuitively understand the visual cognition process from spatial element capture to behavioral decision-making, and has therefore been increasingly used in studies on evaluations of human factors[13], environmental experience[14], etc. However, in many of these studies related to landscape perception and street experience, participants were only asked to freely observe around without any specific activity scenarios[15]~[19], or orthogonal experiments involving multiple environmental factors (e.g., building, tree, sky, and ground) were set for specific purposes, in order to examine how environmental factors affect people's attention[20]~[22]. As people's perception of and preference for a same scene or image is not absolute[23], different spatial setting and research purposes would lead to varied cognition results[24]~[26], which means that without clear task guidance such experiments might fail to achieve the desired experimental outcomes. For example, in an European study[27] that used eye-tracking to investigate how people perceive and observe landscapes, participants were only asked to look at landscape photographs aimlessly, and the research results were nothing but that participants observed panoramic and detailed photographs differently from other types of photographs, which contributed little to the research on environmental elements or landscape experience. Therefore, this study designs an eye-tracking experiment that records people's authentic experience and feelings and behaviors as they wander through the street, attempting to better understand the cognition processes.

In summary, in order to map the distribution of people's attention on the pedestrian street and to learn the visual cognition process from environmental element capture to behavioral decision-making in a more intuitive way, this study uses eye-tracking technology to conduct an exploratory experiment to analyze participants' eye-movement attention. This paper attempts to answer questions, including 1) how attention varies between environmental elements; 2) which types of environmental elements and spaces can importantly affect participants' landscape experience; and 3) how environmental elements specifically guide participants' behavioral decision-making .

2 Research Methods

2.1 Experimental Design

In this study, an eye-movement field experiment was conducted using Ergoneers Dikablis head-mounted eye-trackers with a high sampling frequency (60 Hz)[28] and behavioral tasks were set up to elicit participants' authentic decisions as they wonder the street. Tasks include free-walking observations that use prompts to set specific scenarios and given behavioral tasks, both stimulating participants' daily behaviors on the street environment. In the former task, participants were asked to wear the eye-trackers and freely observe on the street to record their visual cognition during the wandering; the latter included destination selection, point-of-interest (POI) photography, and in-depth interviews, so as to understand their visual cognition process of important events related to spatial experiences.

As this study focuses on understanding pedestrians' real cognition and behaviors in a natural state, a small sample size was adopted for deep systematic exploration[29][30]. This type of in-depth analysis prioritizes the value of research findings for practice and therefore allows an appropriate reduction of sample size[31]. Referring to other exploratory studies, the experiment adopted a sample size of 6 ~ 8 participants and collected data under principles of theoretical sampling①[32] and theoretical saturation②[33].

2.2 Site Selection

This research selected the Nanjing Road Pedestrian Mall (the street hereafter) in Shanghai as the experiment site. Known as "the greatest street of China"[34], the street is a vibrant public space and a tourism destination in Shanghai full of dramatic informative and visual stimuli, and considered suitable for various behavioral experiments. The street was renovated at the beginning of the 21st century, bringing about a sharp visits growth[35] and huge economic and social benefits; in early September 2020, its east extension section was completed, stretching from South Xizang Road to the Bund, with a total length of 1,599 m[34][36].

Based on existing cognition studies on the street[37]~[41] and considering the contrasting landscape appearance difference between the middle section (in-between Hubei Road and Henan Middle Road, remaining the original style of the street) and the newly built east section (in-between Henan Middle Road and Zhongshan East No. 1 Road), this study selected parts of both sections—225 meters for the former and 175 meters for the latter (Fig. 1).

① Theoretical sampling can be understood as the process of data collection under grounded theory whereby the analyst jointly collects, codes, and analyses data and decides what data to collect next and where to find them in order to develop a theory as it emerges (Source: Ref. [32]).

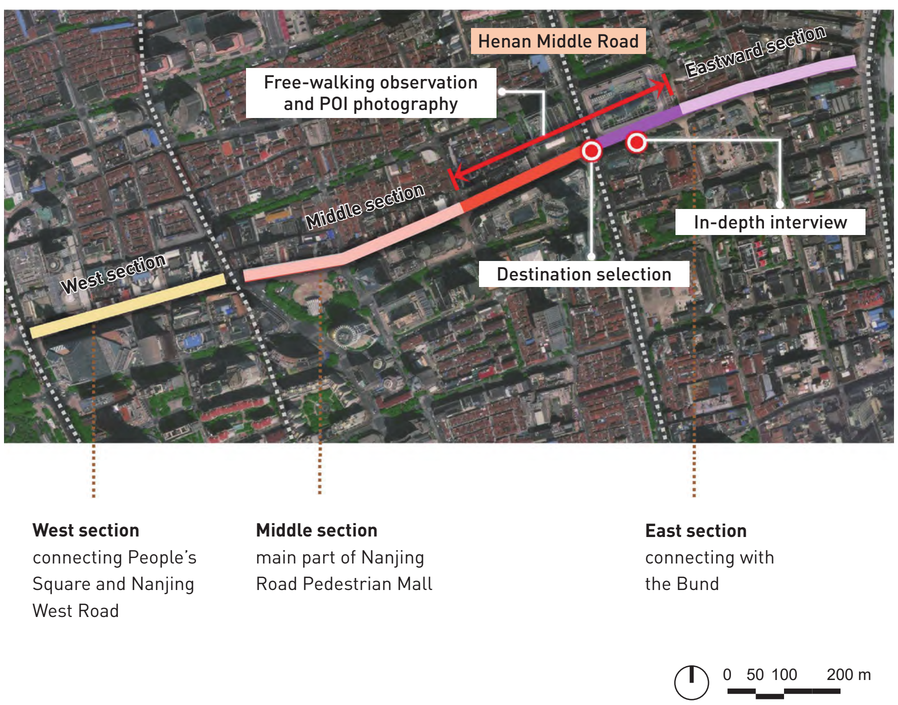
② Theoretical saturation means that in the iterative process of theoretical sampling, the sampling can be stopped when no additional data are being found whereby the analyst can develop properties of the category (Source: Ref. [33]).

2.3 Experiment Procedure

The experiment flow and task spots are shown in Figures 1 and 2. Considering that participants' free-walking observation and POI photography often occurred at the same time, participants were asked to perform both tasks simultaneously in the experiment.

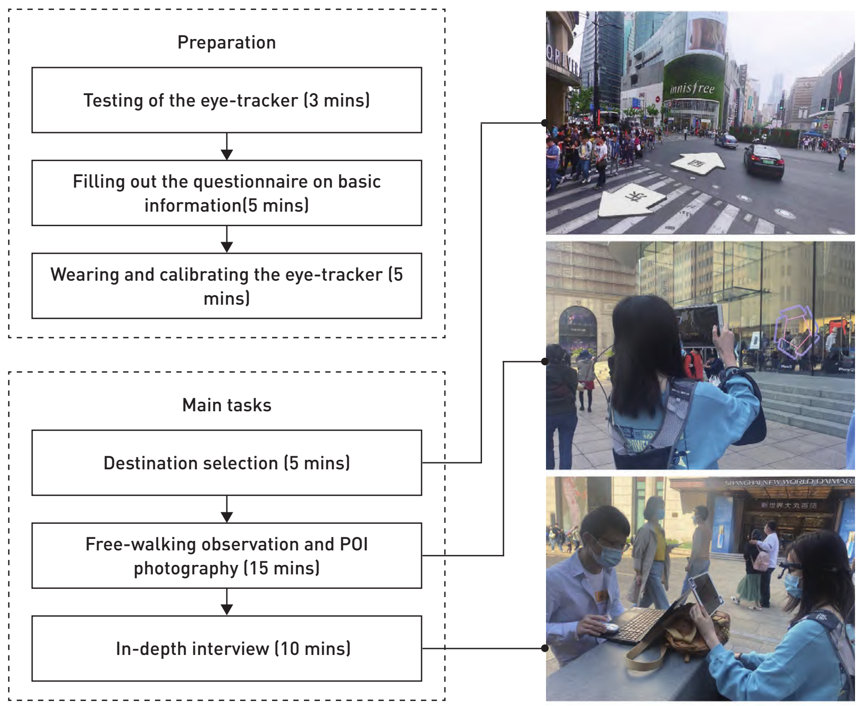
After filling in questionnaires about basic information and calibrating eye-trackers, participants were first asked to perform the destination selection task at the intersection of Henan Middle Road: after a short observation of both sections of the street, they were asked to choose the direction they wanted to wander first, and to perform free-walking observation and POI photography. During the free-walking observation, participants were guided by prompts about walking on the street as they casually do—such as "imagine yourself now walking around the street with your friends during leisure time" and "wander around and stop at any time if you would like to observe or approach POI"—and took pictures whenever they want. During this process, the eye-trackers recorded participants' attention distribution of the street environment, and a

1. Map of selected task spots of the experiment



1 © Chen Yiyan, Chen Zheng, Du Ming

2. Experiment procedure



2 © Chen Yiyan, Chen Zheng, Du Ming

data logging device was installed to record their performance. After walking around both sections, participants would complete the in-depth interviews at the designated location. The questions of in-depth interview were about the results of the destination selection and the reasons, the overall impression of the two sections during the free-walking observation, and the reasons for taking each photo. They were also asked to select three favorite photos from the ones they had taken, and explain the reasons. Researchers ensured that participants answered all the questions completely and clearly.

The experiment recruited on-site visitors as participants to ensure the representativeness and authenticity of the samples. According to existing studies, the majority of visitors to the street are young people (aged 18 ~ 40)[42], accounting for more than 50% of the total visitors, and the proportion keeps growing in recent years[43]; meanwhile, students and company employees are the most occupation types of the visitors as surveyed[42]. Thus, the recruited participants were required between 18 and 40 years old and being students or company employees by occupation. Physiologically, their naked or corrected defects of vision should be at least 1.0, and they should have no eye diseases (e.g., color blindness, color weakness) or hearing disabilities. A total of 7 participants, 3 females and 4 males/5 students and 2 company employees, recently living in Shanghai for at least 6 months, were recruited. The experiment was conducted on October 31 and November 1, 2020 with similar weather conditions during the experiment hours, and each participant completed his/her experiment in 40 minutes, accompanied by a researcher throughout the process.

2.4 Methodologies of Data Analysis

The data collected in this experiment were categorized into pure eye-movement data and behavior-related eye-movement data. Data that were less related to eye movements but also reflected participants' feelings and experience were categorized as pure behavioral data. Comprehensive analyses were conducted for the three categories of data.

1) Pure eye-movement data analysis: in order to visualize the distribution of attention, the research used the area of interest (AOI) analysis method, which is commonly used in eye-movement studies, i.e., by defining a specific extent or element of interest as an AOI, and calculating the accumulated fixation duration and the attention percentage of the given AOI[44].

2) In addition to AOI analysis, scan-path analysis is also used to process behavior-related eye-movement data, which investigates the cognition process associated with behavioral decision-making by analyzing the sequence of attention on environmental elements.

And 3) pure behavioral data, including photographs taken and in-depth interviews, were coded and categorized to systematically explore participants' cognition and perception.

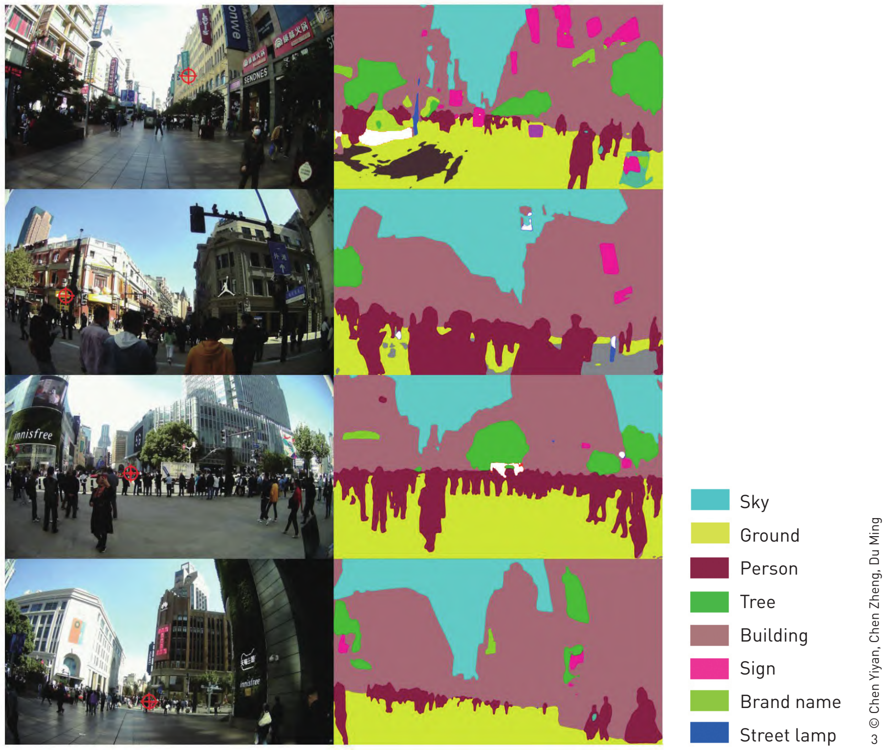
To ensure the data validity, the eye-movement data 5 seconds after the audio instruction in the destination selection task, the eye-movement data 3 seconds before taking photos in the POI photography task, and the eye-movement data throughout the free-walking observation task were extracted.

Eye-movement AOI analysis mainly consists of two steps: automatic semantic segmentation combined with manual audit, and the calculation of the information density of environmental elements.

Due to the huge amount of collected eye-movement data, data analysis process would be extremely time-consuming—for instance, usually an eight-hour manual audit is required for processing a piece of 10-minute data since the researcher needs to audit the data back-and-forth and frame by frame to audit different objects. Therefore, ADE20k dataset trained by Zhou Bolei et al.[45] was used for image semantic segmentation, which is the largest open-source dataset worldwide for semantic segmentation and scene parsing developed by the MIT Computer Vision research group. Its software combining computer semantic segmentation with automatic eye-movement AOI analysis, codevelopped by the Computer Vision Group at the School of Software Engineering, Tongji University, enables automatic batch recognition of attention to each environmental element, greatly improving the efficiency of data analysis (Fig. 3). However, due to the accuracy limitation of automatic semantic segmentation, some environmental elements and areas closely relevant to street experience (e.g., signs, building entrances, posters) could not be accurately recognized. Such AOIs and the corresponding fixation duration need to be manually audited. The total statistics were visualized overlapped on street scene images recorded by the front cameras of eye-trackers[46]. Thus, this study used a combination of automatic semantic segmentation and manual audit in the analyses of the attention distribution of each environmental element.

Additionally, the fixation duration of a given environmental element is largely affected by the visual space the element occupies[47]. To further analyze the attention attracting extent of each element, this study introduces the indicator "information density"③, the ratio of its attention percentage to its pixel (exposure)

③ Existing research has studied similar indicators. For example, in Ref. [44], the researchers described the indicator as "dwell time divided by the % surface areas of AOIs," but did not name it. The indicator is named "information density" in this study.



3. Semantic segmentation of environmental elements and automatic batch processing of AOIs

percentage, to describe the efficiency of information input per unit area. The formula is as follows:

$$I_i = \left(\frac{D_i}{\sum_{j=1}^{n} D_j} \right) / \left(\frac{A_i}{\sum_{j=1}^{n} A_j} \right),$$

where I_i is the information density of the ith element; D_i is the fixation duration of the ith element (ms); $\sum_{j=1}^{n} D_j$ is the sum of the fixation duration of all the n elements in the frame (ms); A_i is the pixel area of the ith element in an image (px); $\sum_{j=1}^{n} A_j$ is the sum of the pixel area of all the n elements in the image (px).

Whereas, both attention percentage and pixel percentage range from 0 to 100%. When the information density is 0 ~ 1, it means that the attention percentage is lower than the exposure percentage; vice versa when the information density is greater than 1. The studies on individual environmental elements can be expanded onto other AOIs that are more important to spatial design (e.g., building ground floor, street furniture), where information density can broadly describe people's intuitive information preferences, and the recognition and systematic intervention of preferences of environmental elements or AOIs can help the spatial design of streets to better direct users' attention, so as to stimulate specific feelings or behaviors.

3 Research Results

3.1 Information Density Differences Among Environmental Elements

Research results reveal the significant differences of information density among varied environmental elements in the destination selection task. In this study, 9 categories of environmental elements were identified according to ADE20k dataset, namely sky, ground/sidewalk/floor ("ground" hereafter), building, person, tree, building entrance, poster, brand name, and sign. The study classified the environmental elements by the mean value of information density of each element into low (0, 1], medium (1, 2], and high (2, ∞) levels (Table 1; Figs. 4, 5): the elements of sign, building entrance, brand name, and poster have the highest information density, followed by tree and person, and the elements of building, sky, and ground have the lowest information density.

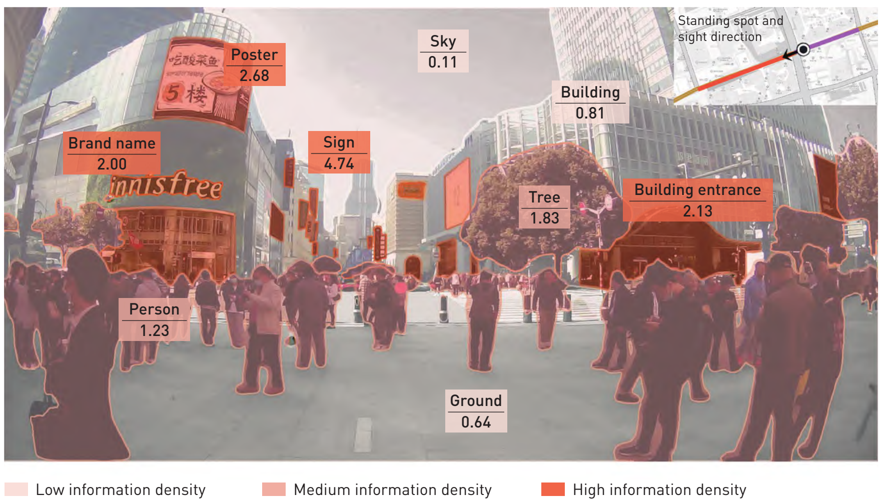
Among the high information density elements, sign had the highest total information density (4.89) while occupying a very small pixel area (0.97%), which was only found in the middle section of the street. The elements of poster and brand name also received considerable attention (2.31 and 2.16 for information density) with very low exposure percentages (2.38% and 1.26%). Environmental elements such as building entrance that are of high transparency and openness also received a lot of attention (20.98%) with a relatively lower exposure percentage (8.56%).

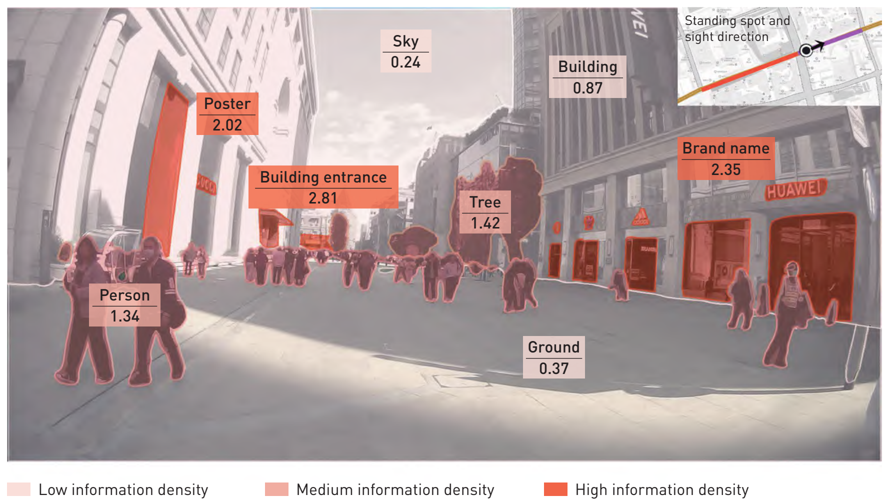
Among the low information density elements, building (after removing elements such as building entrances and signs) had the highest overall attention proportion (25.78%), but its total exposure proportion was also high (30.73%), resulting in a relatively low total information density; meanwhile, the elements of sky and ground had the lowest total information densities (0.16 and 0.49), especially sky, which received little attention while occupying large pixel areas.

These research results suggest that elements with the largest exposure proportion do not always receive the most attention, while some others can attract a lot of attention despite occupying a very small pixel area. It is reasonable to believe that when people wander to be informed on the street, they tend to pay attention to high information density elements such as building entrances and signs that reflect the information and status of spatial activities; tree and person to some extent can also provide spatial information as they also reflect the comfort level and the maintenance situation

Table 1: Information density of each environmental element

Environmental elements		Middle section			East section			Total		
		Attention percentage	Exposure percentage	Information density	Attention percentage	Exposure percentage	Information density	Attention percentage	Exposure percentage	Information density
Low information density (0, 1]	Sky	1.57%	13.90%	0.11	1.86%	7.82%	0.24	1.71%	10.86%	0.16
	Ground	13.04%	20.27%	0.64	9.65%	26.35%	0.37	11.41%	23.31%	0.49
	Building	21.50%	26.60%	0.81	30.41%	34.85%	0.87	25.78%	30.73%	0.84
Medium information density (1, 2]	Person	15.34%	12.52%	1.23	13.66%	10.17%	1.34	14.50%	11.35%	1.28
	Tree	9.70%	5.31%	1.83	3.30%	2.33%	1.42	6.50%	3.82%	1.70
High information density (2, ∞)	Brand name	2.86%	1.43%	2.00	2.56%	1.09%	2.35	2.72%	1.26%	2.16
	Poster	5.70%	2.13%	2.68	5.30%	2.63%	2.02	5.50%	2.38%	2.31
	Building entrance	19.28%	9.04%	2.13	22.68%	8.08%	2.81	20.98%	8.56%	2.45
	Sign	9.20%	1.94%	4.74	/	/	/	4.74%	0.97%	4.89





4. Information density of each environmental element in the middle section
5. Information density of each environmental element in the east section

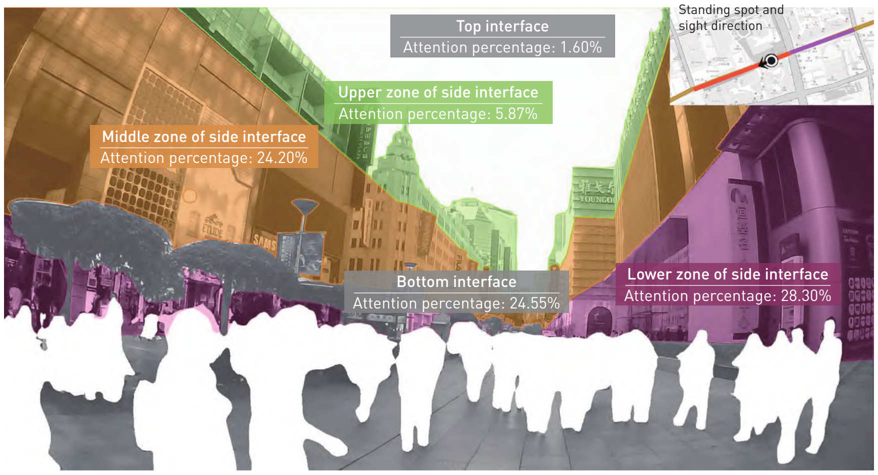
of the site; sky and ground that have a relatively high exposure proportion but less information are easily overlooked.

3.2 Stimulating Landscape Perception of Streets Through Systematic Design of Environmental Elements of High Information Densities

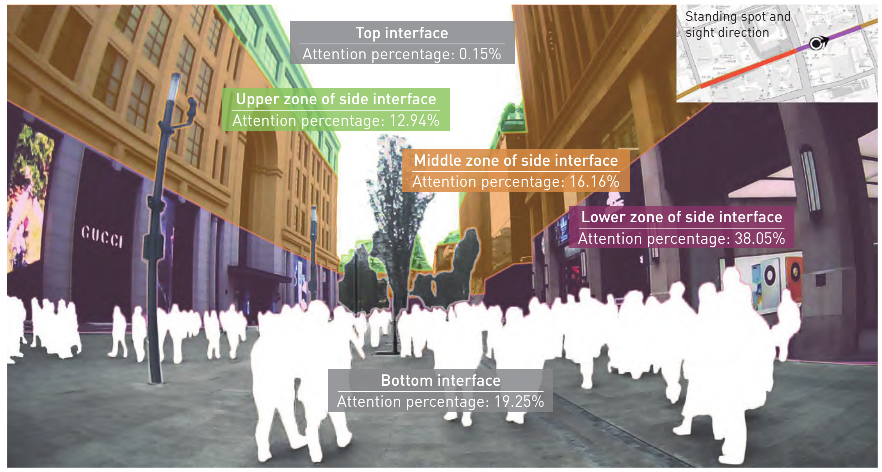
During the free-walking observation task, the researchers found that environmental elements of high information densities such as signs and posters were mostly found on the middle zone of side interfaces on the street. To probe into how does such a setting affect participants' attention and perception, this research spatially divided the street into top, bottom, and side interfaces and further divided the side interfaces into upper (top floor of podium and higher areas), middle, and lower zones (ground floors), to analyze

Table 2: Attention percentage at varied interfaces of the street during the free-walking observation task

	Interfaces				
	Side interface			Top interface	Bottom interface
	Lower zone	Middle zone	Upper zone		
Middle section	28.30%	24.20%	5.87%	1.60%	24.55%
East section	38.05%	16.16%	12.94%	0.15%	19.25%



6 © Chen Yiyan, Chen Zheng, Du Ming



7 © Chen Yiyan, Chen Zheng, Du Ming

6. Attention distribution at the upper, middle, and lower zones of the side interface in the middle section
7. Attention distribution at the upper, middle, and lower zones of the side interface in the east section

the attention distribution accordingly.

Overall, the side interface of the street received significantly more attention than the top and bottom interfaces; participants' attention was concentrated on the lower zone in the east section, and on the lower and middle zones in the middle section (Table 2; Figs. 6, 7).

Specifically, in the middle section, the lower zone of the side interface received the most attention (28.30%), the middle zone also received a similar proportion (24.20%). This does not prove the results of other existing research suggesting that people's attention on the street often concentrates on the ground floor of buildings[24][48] (i.e., the lower zone of the side interface). Participants described the middle section of the street as "vibrant and busy," "visually interesting and rich," and "very distinctive," for reasons such as "unique business signs," "large LED advertising posters," and "old-fashioned architectural decorations," all of which were the environmental elements appeared in the middle zone of the side interface. In addition, researchers also noticed that participants' body movements such as looking up, turning their heads, and turning around, in addition to their eye movements, exemplified that they were attracted by the middle and upper zones of buildings. Usually a pedestrian's horizontal viewshed is about 90° on each side, while the vertical viewshed is very limited—the one's visual axis would even shift downwards by 10° during a careful walk[4]—this explains why one's sight often concentrates on the lower zone of the side interface when walking. However, when the elements of high information density are arranged at a certain height, people would quickly adjust their sight towards where they can get more information. The middle section of the street gathers various kinds of distinctive and old-fashioned signs and posters in the middle zone of the side interface, including diverse architectural decorations such as customized gables and corners, creating a spatial style different from ordinary street spaces and an AOI of rich visual components that simulate people's curiosity. All these can lead to a series of observation behaviors and create a "unique," "rich," and "distinctive" landscape experience.

3.3 Personal Emotional Connections Would Contribute to Environmental Attractiveness

This research explored the visual cognition processes associated with environmental attractiveness by categorizing the photos taken by participants, combined with the analysis of eye-movement attention scan-paths before the photo-taking and the in-depth interviews. Based on the text encoding analysis on the photography reasons, the photos can be classified into core impression (42 pieces), unique experience (32 pieces), and personal association (8 pieces) according to their environmental attractiveness (Table 3, Fig. 8).

1) Photos of core impression reflect unique scenes or nodes of the street, while participants described them in the in-depth interviews as "representing the characteristics of the street" and "not seen elsewhere." These photos recorded, for example, the area with "various" and "colorful" business signs (photo No. 7) and the "crowded" and "massive" visitors (photo No. 13).

Table 3: Participants' self-reported reasons for taking photographs of personal association

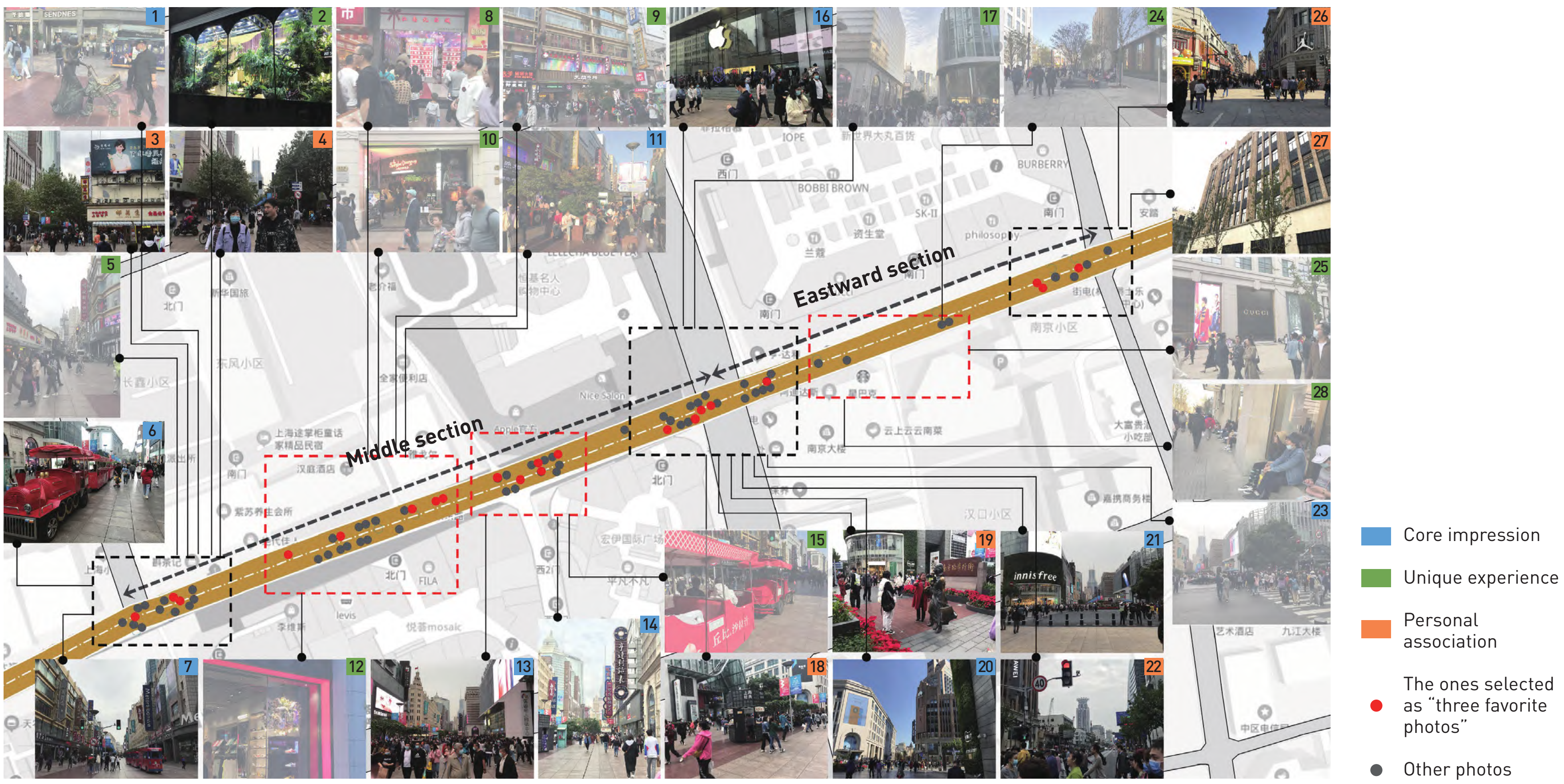
Photo number	Content	Participants' self-reported reasons for taking photographs	Favorite photo or not
3		"Because I think this Old Fengxiang billboard is quite retro... like the kind I had seen when I visited Shanghai as a child, at that time there were such billboards shown in TV commercials, very old Shanghai style."	Yes
4		"At the end of the street you can see the World Trade Center, the building is of an unusual shape—it has two "horns"—I can see this building all the way along the Nanjing East Road, as what I remember."	No
18		"When I was walking slowly waiting for the traffic lights, I passed by this kiosk and took a closer look. It seems that it is a shop of Shanghai facial creams, and my friends and I had bought some products somewhere else. It is interesting!"	Yes
19		"When I first came to Shanghai, I took a photo here with my friends on my first visit to the street. We really had fun till late night, and there were few people on the street. Such a memory! I miss that."	Yes
Not numbered because of the same content with photo No. 19		"When I was an undergraduate, my mother came to visit me, and I took her some photos right here on the street."	No

Continued

Table 3: Participants' self-reported reasons for taking photographs of personal association

Photo number	Content	Participants' self-reported reasons for taking photographs	Favorite photo or not
22		"I was trying to get a shot of the Bund Center in the distance... My favorite thing about touring here when I was a kid was this building, especially the beautiful lightings at night. The shape of the building is very extraordinary, I guess."	Yes
26		"The Peace Hotel is ahead! It is so close to the Bund. Last time my friend and I sat at the Bund for a long time. I always think the illuminated green roof of this building looks appealing at night."	Yes
27		"This building looks quite like the one in a TV series I watched a while ago. It is the main character's office, and almost the same one, particularly the building textures! I wonder if the TV series was took here!"	No

8. The taken spots of the photos of core impression, unique experience, and personal association, and the selected "three favorite photos" (those photos representing the same content are not repeatedly displayed in this figure).



8 © Chen Yiyan, Chen Zheng, Du Ming

2) Photos of unique experience record what participants perceived as interesting and different from everyday life, such as unique window decorations (photo No. 2) and a sales clerk inviting customers in a panda costume (photo No. 8).

And 3) photos of personal association reflect the scenes or nodes evoking participants' memories, described such as "having been there before" and "being associated with familiar films and television works," for example, the World Trade Center at the end of the street (photo No. 4) and the Bund Center (photo No. 22).

Although the percentage of personal association photographs was small (only 10% of the total number of photographs), they were selected as the "three favorite photos" by the highest percentage (63%), much higher than the 33% for core impression and 6% for unique experience. This suggests that when a scene/node evokes a participant's personal memory, it tends to build emotional connections and thus become a "favorite" part of the one's overall landscape experience.

The reasons for the selection of personal association photos as the favorite ones can be further understood combining with eye-tracking attention scan-path analysis. Photo No. 19 (Fig. 9-1) was taken near the iconic marble sign of the street, at the entrance of the middle section. The eye-tracking attention scan-path in the first 3 seconds (Fig. 9-2) shows that the participant firstly noticed the person taking photograph and the person being photographed in front of the sign, then observed the sign to recognize the photography act, and then observed the persons and the sign again—in the process the participant might recall the memories of taking similar photographs; latter the one observed the flower bed to consider the possibility of passing, and finally decided to take a photograph *in-situ*.

The results of the experiment show that when environmental experiences evoke personal association and motivate sentiment (i.e., emotional resonance), they tend to lead to the visitor's positive cognition outcomes and the enhancement of environmental attractiveness.

4 Conclusions

This study uses eye-tracking technology to intuitively understand the cognition process from environmental element capture to behavioral decision-making, providing new insights for the research on spatial experience of streets. At the same time, the combination of eye-tracking data analysis with in-depth interviews establishes a direct link between non-material cognitive factors (e.g., personal experiences, emotions, and memories) and material physical spaces with qualitative methods, providing an effective means for similar studies. In addition, the study set tasks of street experiences in real-life scenes to explore the relations between different environmental elements in cognition process for specific behavioral purposes, offering a deeper understanding of the overall cognitive process and a new research path. This study includes the following main findings.

1) The experiment site, Nanjing Road Pedestrian Mall (Street), has rich environmental elements that vary in information density: the highest for the elements of sign, building entrance, poster, and brand name, followed by tree and person. High information density elements, especially signs, have very low exposure percentages, but they significantly affect people's overall environmental perception of the street, and should be prioritized in the spatial design of streets.


© Chen Yiyan, Chen Zheng, Du Ming
9-1

© Chen Yiyan, Chen Zheng, Du Ming
9-2

9. Fig. 9-1 is taken by the participant; Fig. 9-2 is the views overlaid with fixation and scan-path maps (fixation duration showed in circles and saccade path in lines) to the original photo.

2) Systematic design of environmental elements of high information density would facilitate creating unique landscape experiences. On the street, the systematic design of high information density elements such as signs has greatly enhanced participants' attention to the middle zone of buildings, forming an extraordinary landscape style from other streets, largely improving the effects of design interventions. In recent years, Huangpu District Government of Shanghai has carried out a series of renovation and renewal projects on the street, during which such environmental elements in the middle zones of buildings were mostly preserved, protecting and sustaining the landscape identity and characteristics of the street. In future design of urban spaces, more attention should also be paid to identifying the environmental elements of high information density and their influence on visitors' spatial perception and behaviors, so as to stimulate and encourage certain feelings and activities, and to improve the overall environmental quality of designed spaces.

And 3) notably, among the spaces that attract visitors' attention, places that evoke emotions are more likely to be memorable and lead to positive cognition. In this study, participants' walking experience on the street was continuously excited by various environmental cues that associated with both sensory and emotional memories, reinforcing the emotional connection between the individual and the place, thus contributing to the environmental attractiveness of the street. In planning and design practices, recognizable environmental cues such as uniform, recurring icons or signs and street furniture can be used to improve such emotional connections.

RESEARCH FUND

- Precise Repairing of Urban Street Landscape via *In-situ* Environmental Experience Computing, the National Nature Science Foundation of China (No. 51878461)

REFERENCES

[1] Wei, M., & Zhang, Y. (2000). Creating a "People-Oriented" urban commercial pedestrian system. *Urban Planning Forum*, (1), 20–22.
[2] Lynch, K. (1962). *The Image of the City*. MIT Press.
[3] Appleyard, D., Lynch, K., & Myer, J. R. (1964). *The View From the Road*. MIT Press.
[4] Gehl, J. (2003). *Life Between Buildings*. Van Nosrand Reinhold.
[5] Zhao, C., Yang, B., & Liu, D. (2012). PSPL survey: The evaluation method for quality of public space and public life—The study on Jan Gehl's theory and method for public space design (Part 3). *Chinese Landscape Architecture*, 28(9), 34–38.
[6] Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The Measurement of Meaning*. University of Illinois Press.
[7] Venables, P. H. (1991). Autonomic activity. *Annals of the New York Academy of Sciences*, 620(1), 191-207. doi:10.1111/j.1749-6632.1991.tb51584.x
[8] Csikszentmihalyi, M., Larson, R., & Prescott, S. (1977). The ecology of adolescent activity and experience. *Journal of Youth and Adolescence*, 6(3), 281-294. doi:10.1007/BF02138940
[9] Haywood, K. M. (1990). Visitor-employed photography: An urban visit assessment. *Journal of Travel Research*, 29(1), 25-29. doi:10.1177/004728759002900106
[10] Strauss, A. L. (1987). *Qualitative Analysis for Social Scientists*. Cambridge University Press.
[11] Yarbus, A. L. (2013). *Eye Movements and Vision*. Springer.
[12] Hamamé, C. M., Vidal, J. R., Perrone-Bertolotti, M., Ossandón, T., Jerbi, K., Kahane, P., Bertrand, O., & Lachaux, J. (2014). Functional selectivity in the human occipitotemporal cortex during natural vision: Evidence from combined intracranial EEG and eye-tracking. *Neuroimage*, (95), 276-286. doi:10.1016/j.neuroimage.2014.03.025
[13] Shoval, N., Schvimer, Y., & Tamir, M. (2018). Real-time measurement of tourists' objective and subjective emotions in time and space. *Journal of Travel Research*, 57(1), 3-16.
[14] Scott, N., Zhang, R., Le, D., & Moyle, B. (2019). A review of eye-tracking research in tourism. *Current Issues in Tourism*, 22(10), 1244-1261. doi:10.1080/13683500.2017.1367367
[15] Potocka, I. (2013). The lakescape in the eyes of a tourist. *Quaestiones Geographicae*, 32(3), 85–97. doi:10.2478/quageo-2013-0018
[16] Nomoto, K., Shimosaka, T., & Sato, R. (2018). Comparison of gaze behavior among Japanese residents, Japanese visitors, and non-Japanese visitors while walking a street. *2018 Joint 10th International Conference on Soft Computing and Intelligent Systems (SCIS) and 19th International Symposium on Advanced Intelligent Systems (ISIS)* (pp. 693-697). IEEE. doi:10.1109/SCIS-ISIS.2018.00117
[17] Kiefer, P., Giannopoulos, I., Kremer, D., Schlieder, C., & Raubal, M. (2014). Starting to get bored: An outdoor eye tracking study of tourists exploring a city panorama. *Proceedings of the Symposium on Eye Tracking Research and Applications* (pp. 315-318). Association for Computing Machinery. doi:10.1145/2578153.2578216
[18] Berto, R., Massaccesi, S., & Pasini, M. (2008). Do eye movements measured

across high and low fascination photographs differ? Addressing Kaplan's fascination hypothesis. *Journal of Environmental Psychology*, 28(2), 185-191. doi:10.1016/j.jenvp.2007.11.004
[19] Franěk, M., Šefara, D., Petružálek, J., Cabal, J., & Myška, K. (2018). Differences in eye movements while viewing images with various levels of restorativeness. *Journal of Environmental Psychology*, (57), 10-16. doi:10.1016/j.jenvp.2018.05.001
[20] Noland, R. B., Weiner, M. D., Gao, D., Cook, M. P., & Nelessen, A. (2017). Eye-tracking technology, visual preference surveys, and urban design: Preliminary evidence of an effective methodology. *Journal of Urbanism: International Research on Placemaking and Urban Sustainability*, 10(1), 98-110. doi:10.1080/17549175.2016.1187197
[21] Nordh, H., Hagerhall, C. M., & Holmqvist, K. (2013). Tracking restorative components: Patterns in eye movements as a consequence of a restorative rating task. *Landscape Research*, 38(1), 101-116. doi:10.1080/01426397.2012.691468
[22] Valtchanov, D., & Ellard, C. G. (2015). Cognitive and affective responses to natural scenes: Effects of low level visual properties on preference, cognitive load and eye-movements. *Journal of Environmental Psychology*, (43), 184-195. doi:10.1016/j.jenvp.2015.07.001
[23] Yarbus, A. L. (1967). Eye Movements During Perception of Complex Objects. *Eye Movements and Vision* (pp. 171-211). Springer.
[24] Simpson, J., Freeth, M., Simpson, K. J., & Thwaites, K. (2019). Visual engagement with urban street edges: Insights using mobile eye-tracking. *Journal of Urbanism: International Research on Placemaking and Urban Sustainability*, 12(3), 259-278. doi:10.1080/17549175.2018.1552884
[25] Simpson, J. (2018). Street DNA: The who, where and what of visual engagement with the urban street. *Journal of Landscape Architecture*, 13(1), 50-57. doi:10.1080/18626033.2018.1476032
[26] Sun, C., & Yang, Y. (2019). A study on visual saliency of way-finding landmarks based on eye-tracking experiments as exemplified in Harbin Kaide Shopping Center. *Architectural Journal*, (2), 18-23.
[27] Dupont, L., Antrop, M., & Van Eetvelde, V. (2014). Eye-tracking analysis in landscape perception research: Influence of photograph properties and landscape characteristics. *Landscape Research*, 39(4), 417-432. doi:10.1080/01426397.2013.773966
[28] Bojko, A. (2013). *Eye Tracking the User Experience: A Practical Guide to Research*. Rosenfeld Media.
[29] Creswell, J. W. (1998). *Qualitative Inquiry and Research Design: Choosing Among Five Approaches*. Sage publications.
[30] Creswell J. W. (2003). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (2nd ed.). Sage.
[31] Kaplan, R. (1996). The Small Experiment: Achieving More With Less. In J. L. Nasar & B. B. Brown (Eds.), *Public and Private Places* (pp. 170-174). Environmental Design Research Association.
[32] Glaser, B. G., Strauss, A. L., & Strutzel, E. (1968). The discovery of grounded theory; strategies for qualitative research. *Nursing Research*, 17(4), 377-380.
[33] Rubin, H. J., & Rubin, I. S. (2011). *Qualitative Interviewing: The Art of Hearing Data*. Sage.
[34] Zheng, S., Wang, W., & Chen, Y. (2001). Pedestrian streets radiate urbanism. *Architectural Journal*, (6), 35–39. doi:10.3969/j.issn.0529-1399.2001.06.010
[35] Zheng, S., Qi, H., & Wang, W. (2000). Function upgrading and extension of urban space—Study on the transformation background of Nanjing East Road Pedestrian Street. *Urban Planning Forum*, 125(1), 13-19.
[36] Wang, Y., Jiang, Y., & Villadsen, K. S. (2015). Strategies for Street improvement in world-class cities: A case study of Huangpu District in Shanghai. *Urban Transport of China*, 13(1), 35–39. doi:10.13813/j.cn11-5141/u.2015.0106
[37] Chang, Q. (2005). *Begin of a Metropolis: A Study of the Bund Section of Nanjing Road in Shanghai*. Tongji University Press.
[38] He, S., & Wang, R. (1998). Create a pro-human urban public space-on the construction of a commercial pedestrian street on Nanjing East Road, Shanghai. *Architectural Journal*, (3), 20-25, 66-67.
[39] Wang, W. (2000). Creating a vibrant central business zone—Analysis on the environmental design of Nanjing East Road Pedestrian Street. *Planners*, 16(6), 28–31.
[40] Zhou, S., Yan, W., Liu, Y., & Meng, H. (2001). The design of Nanjing Road Pedestrian Mall in Shanghai. *New Architecture*, (3), 1–5.
[41] Zha, J., Wang, S., & Tang, W. (2020). Renovation and new appearance through history: Practice and reflection on public space renewal of the historical and cultural areas in shanghai and taking the east extension section of Nanjing Road pedestrian street as an example. *Architectural Practice*, (10), 72–81.
[42] Cai, J., Wang, D., & Zhu, W. (2011). Study on changes of consumer behavior in East Nanjing Road—A comparison of the year 2001 and 2007. *Human Geography*, 26(6), 89–97. doi:10.13959/j.issn.1003-2398.2011.06.014
[43] Tang, Y. (2021, May 28). *The investment platform of Nanjing Road pedestrian street launched; The post-80s and post-90s generations become its major guest groups*. Shanghai Observer.
[44] Duchowski, A. T. (2003). Eye Tracking Techniques. *Eye Tracking Methodology: Theory and Practice* (pp. 55-65). Springer. doi:10.1007/978-1-4471-3750-4_5
[45] Zhou, B., Zhao, H., Puig, X., Fidler, S., Barriuso, A., & Torralba, A. (2017). Scene Parsing Through ADE20k Dataset. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 5122-5130). IEEE. doi:10.1109/CVPR.2017.544.
[46] Lusk, A. (2002). *Greenways' places of the heart: Aesthetic guidelines for bicycle paths* (Doctoral dissertation). University of Michigan.
[47] Amati, M., Parmehr, E. G., McCarthy, C., & Sita, J. (2018). How eye-catching are natural features when walking through a park? Eye-tracking responses to videos of walks. *Urban Forestry & Urban Greening*, (31), 67-78. doi:10.1016/j.ufug.2017.12.013
[48] Zhang, L., Zhang, R., Yin, B., Wang, P., & Yan, S. (2021). Study on the style perception of the commercial historical street based on eye-tracking—A case of Arcade Street in Shapowei, Xiamen. *Urbanism and Architecture*, 18(16), 111–118, 148. doi:10.19892/j.cnki.csjz.2021.16.25
[49] Kim, J., & Fesenmaier, D. R. (2015). Measuring emotions in real time: Implications for tourism experience design. *Journal of Travel Research*, 54(4), 419-429.
[50] Duan, J., & Chen, W. (2012). Ambulatory-assessment based sampling method: Experience sampling method. *Advances in Psychological Science*, 20(7), 1110-1120. doi:10.3724/SP.J.1042.2012.01110

https://doi.org/10.15302/J-LAF-1-020064　中图分类号 | TU985　文献标识码 | A　收稿时间 | 2021-06-09

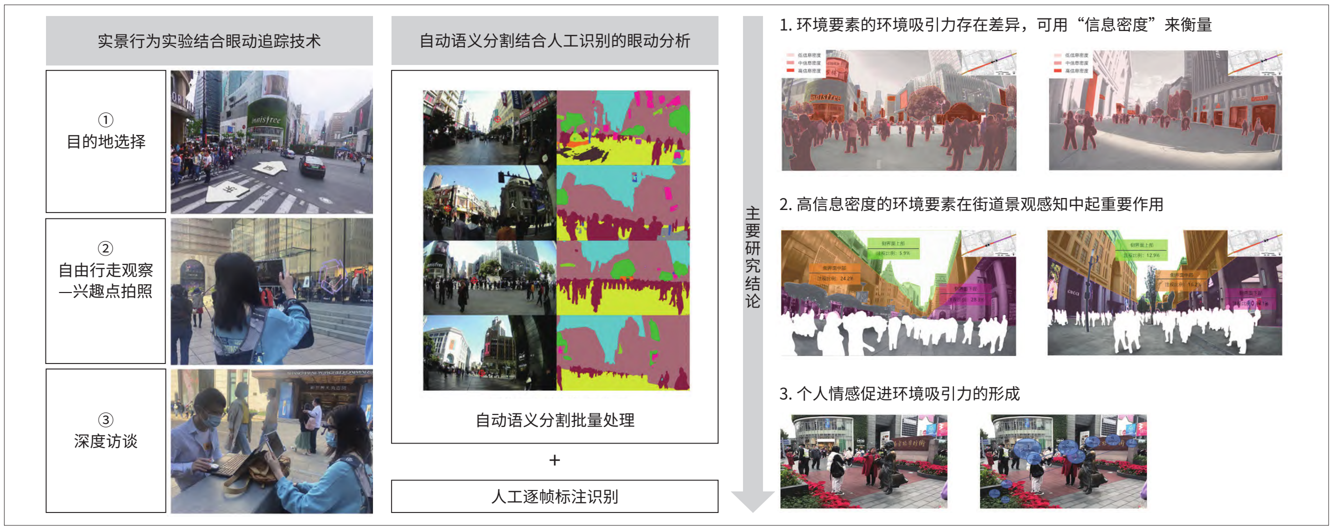
注意力的设计
——眼动追踪技术辅助下的上海市南京路步行街景观体验研究

陈奕言[1]，陈筝[1,2,*]，杜明[3]

1 同济大学建筑与城市规划学院，上海 200092
2 同济大学高密度人居环境生态与节能教育部重点实验室，上海 200092
3 同济大学建筑设计研究院（集团）有限公司，上海 200092

***通讯作者**
地址： 上海市杨浦区四平路1239号
邮编： 200092
邮箱： zhengchen@tongji.edu.cn

图文摘要



文章亮点

- 实景行为实验结合眼动追踪技术可更好理解环境感知和行为决策
- 眼动分析可采用自动语义分割和人工识别相结合的方式
- 环境要素的环境吸引力存在差异，可用“信息密度”来衡量
- 外挂店招等高信息密度的环境要素在街道景观感知中起重要作用

摘要

为了解人们在商业步行街上的注意力分布情况，理解街道体验相关的视觉认知过程，本研究选择上海市南京路步行街作为实验场地，采用头戴式眼动仪探究参与者自然状态下行走时的注意力关注点，同时结合目的地选择、兴趣点拍照、深度访谈多种实验任务，以激发参与者真实的行为决策。研究采用自动语义分割与人工识别结合的方法，分析并统计各环境要素的眼动兴趣区注视情况，并引入“信息密度”这一眼动新指标（即主观注视比例与客观画面面积占比的比值），以反映环境要素或区域对注意力的吸引程度。研究结果显示，环境要素的信息密度具有差异：外挂店招、建筑入口空间、品牌标识和电子屏幕的信息密度最高，其次是绿化和人群，建筑立面、天空和地面最低。通过对以外挂店招为代表的高信息密度环境要素进行系统性设计，有助于塑造特色鲜明的街道景观体验。此外，研究结果显示，当环境体验与个人经历相关并被赋予情感价值时，容易导致正向的认知判断结果，促进环境吸引力的形成。

关键词

街道景观风貌；
南京路步行街；
景观体验；
头戴式眼动仪；
实景实验；
空间认知；
环境吸引力

编辑 田乐，汪默英
翻译 田乐，汪默英，顾芗

1 引言

街道不仅是城市交通的重要载体，也是城市社会活动的主要场所。其中，随着城市快速发展，商业步行街逐渐从单纯提供商业服务向展现特定景观风貌、承载多元活动发展，成为城市社会生活的重要场所[1]。

国内外研究者针对街道中的行为活动和景观体验，开展了基于观察调研的认知学体验分析，开发了各种研究工具和量表，通过行为观察、问卷调研、深度访谈等方法采集数据，来衡量使用者活动和行为心理。研究工具包括广为人知的凯文·林奇提出的城市意象五要素（City Image and Its Five Elements）及认知地图（Cognitive Mapping）[2][3]，研究方法包括扬·盖尔提出的立足场地调查和行为观察的“公共空间—公共生活”调研法（Public Space & Public Life Survey）[4][5]，以及查尔斯·埃哲顿·奥斯顾德提出的立足心理实验的SD语义解析法（semantic differential）[6]。而随着生理学与心理学研究的推进，基于生理反馈的情绪体验捕捉[7]和注重生态效度的主观体验评价[8]~[10]等创新性方法也相继提出。然而，由于环境体验和感受本身的非物质特征，这些研究虽然一定程度上衡量了使用者在空间中的体验感受与情感记忆，但却很难有效支持具体空间要素层面的设计应用。

眼动追踪技术可以记录人眼球的客观运动，提供毫秒级的追踪响应，能够精准捕捉人对环境的快速认知加工[11][12]，可帮助研究者直观理解从空间要素到行为决策的视觉认知过程，因此近年来已在人因评价[13]、环境体验[14]等领域得到了越来越多的应用。其中，不少与景观感受和街道体验相关的研究不设计具体的行为活动情境，以进行自由观察为主[15]~[19]，或围绕具体目的设计建筑立面、绿化、天空、地面等多种环境因素的正交实验，考察这些环境因素对注视的影响[20]~[22]。但是，由于人们对于同一场景或画面的认知及偏好不是绝对的[23]，在空间的探索中，不同的情境和目的会产生不同的关注[24]~[26]，若没有清晰的任务引导，往往无法达到期望的实验效果。如欧洲一项研究[27]在利用眼动仪探究人们是如何感知和观察风景时，仅让参与者无目的地观察风景照片，最终只得出全景照片和细节照片的观察方式不同于其他类型照片的结论，没有揭示任何与环境要素或风景体验相关的有效结论。因此，本研究在眼动实验中设置行为任务还原人们在街道中游逛时的真实状态和行为活动，以准确理解认知过程。

综上，为了解人们在步行街上的注意力分布情况，更加直观地理解从环境要素到行为决策的视觉认知过程，本研究选用眼动追踪技术进行探索性的眼动实景实验，并在眼动实验中设置行为任务，在还原人们在街道中游逛时的真实状态和行为活动的同时激发真实的行为决策，以捕捉分析更加贴近实际情况、真实可靠的眼动情况。具体研究问题包括：1）人们对不同环境要素的注视情况有何差异；2）哪些环境要素和空间对景观体验影响较大；3）环境要素具体如何引导行为决策。

2 研究方法

2.1 实验总体设计

本研究采用眼动现场实验，使用60Hz高采样频率[28]的Ergoneers Dikablis头戴式眼动仪，并设置行为任务诱发行人在街道中的真实决策。任务包括用提示语限定具体生活场景的自由行走观察及特定行为任务两类，均真实还原了行人在街道环境中的常见行为：前者使参与者佩戴眼动仪在实景环境中自由观察，还原其日常游逛时的视觉认知状态；后者通过特定行为任务（包括目的地选择、兴趣点拍照及深度访谈）进行引导，用于理解空间体验相关重要事件的视觉认知过程。

此外，本研究注重理解自然状态下行人的真实认知和行为，因此采用小样本进行深度系统性挖掘[29][30]。这类深度分析把结论对实践的指导价值放在第一位，适当降低了对于样本和控制的考虑[31]。本研究选用此类研究常见的6~8人样本规模，并遵循“理论抽样”（theoretical sampling）①[32]和“理论饱和”（theoretical saturation）②[33]原则，旨在进行一次基于小样本的探索性质性研究。

2.2 场地选择

本研究选择上海市南京路步行街作为研究对象。南京路步行街（以下简称“步行街”）是上海公认的高活力公共空间，素有“中华第一街”的美誉[34]，具有高度丰富的信息和视觉刺激，利于进行各类自然状态下的行为实验。步行街于21世纪初完成改造，客流量显著上升[35]，取得了巨大的经济社会效益；继而在2020年九月初完成东拓，从西藏南路绵延至外滩，全长1 599m[34][36]。

结合南京路现有研究[37]~[41]，中段是步行街原本路段中的典型路段，东拓段刚刚建成且与步行街原本路段整体风格差异较大，因此各截取一段——中段（湖北路—河南中路部分路段，长225m）和东拓段（河南中路—中山东一路部分路段，长175m）——开展实验（图1）。

2.3 现场实验流程

实验流程及主要任务选点空间分布如图1、2所示。其中，考虑到参与者自由行走观察与兴趣点拍照在街道上常同时发生，因此实验中要求参与者同时进行。

① “理论抽样”是常用于“扎根理论”（grounded theory）研究进行理论建构时确定访谈样本的标准，即对访谈内容建立假设并产出结论，根据理论建构本身的完整性确定是否补充访谈、要补充哪些访谈的标准（来源：参考文献[32]）。

② “理论饱和”即当访谈中获得的信息开始重复，不再有新的、重要的主题出现时，就可以认为信息已经饱和，不再需要继续进行更多参与者的访谈（来源：参考文献[33]）。

在完成基础信息问卷填写、眼动仪佩戴校准后，参与者首先在河南中路路口进行目的地选择任务：要求参与者观察步行街中段和东拓段两个方向后，选择想先游逛的方向，接着进行自由行走观察和兴趣点拍照。自由行走观察时通过提示语引导参与者假想街道常见的生活化场景，如“假设自己于闲暇时光与朋友共同游逛步行街”“按日常习惯行走，随时可以停下观察或靠近感兴趣的人或物”等，同时行走过程中可以随时拍照记录感兴趣的内容。在此过程中，眼动仪全程记录参与者注视情况，同时搭载录音设备进行录音。参与者在中段和东拓段规定路段各行走一个来回后，最后于指定地点完成深度访谈。深度访谈就目的地选择结果及其理由、自由行走观察中对两个路段的整体印象、每张照片的拍摄理由进行提问，并要求参与者在自己所拍摄的照片中挑选3张最喜爱的并说明理由。在此过程中，确保参与者完整、清晰地回答所有问题。

研究选择步行街游客作为实验参与者，以确保样本的代表性和真实性。根据现有研究，步行街来访主体为青年（18~40周岁）[42]，占总访客人数的50%以上，且近年来占比逐渐增大[43]；同时，游客的职业类型以学生和公司职员为主[42]。基于此，本实验将参与者要求设定为：年龄18~40周岁，职业类型为学生或公司职员，裸眼或矫正视力达1.0以上，无色盲、色弱等视觉缺陷或眼部疾病，听力正常。实验最终招募并筛选了共7名参与者，3名女性、4名男性，5位学生、2位公司职员，且近半年内均常住上海。实验于2020年10月31日和11月1日完成，两日实验期间的天气情况相似；每位参与者的实验过程控制在40分钟左右，研究人员全程跟随。

2.4 数据分析方法

研究将采集到的数据分为“纯眼动数据”和“与行为相关的眼动数据”两类。此外，一些和眼动关系较弱、但也反映了感受和体验的数据被归为“纯行为数据”。研究对这三类数据进行了综合处理分析。

1）纯眼动数据的分析中，为了直观反映注视分布，研究采用眼动研究中常用的眼动兴趣区分析法[44]，即限定具体感兴趣的分析区域或要素为眼动兴趣区，并统计兴趣区内的累积注视时长和占比。

2）与行为相关的眼动数据除采用眼动兴趣区分析外，为探究行为决策相关的认知过程，也使用眼动扫视轨迹分析来分析环境要素的注视先后顺序。

3）纯行为数据包括拍摄的照片和深度访谈内容，研究对此类数据进行了整理编码分类统计，以系统挖掘主观心理感受。

以上数据在综合处理分析前均进行了有效片段的提取，包括目的地选择任务中任务指令发出后5s的眼动数据、兴趣点拍照中截取拍照前3s的眼动数据、自由行走观察截取全程眼动数据等。

本研究中，眼动兴趣区分析包括“自动语义分割与人工识别结合”和“计算环境要素的信息密度”两个关键分析步骤。

由于采集的眼动数据量巨大，数据分析需要围绕不同目的人工来回逐帧标注，耗时耗力，通常长度10分钟的数据，人工分析约需8小时。因此，我们采用周博磊等人[45]训练的ADE20k街景标签数据库进行图像语义分割，该数据集是MIT Computer Vision团队发布的世界最大的语义分割和场景解析的开源数据集，并由同济大学软件学院计算机视觉课题组协助开发了结合计算机语义分割的眼动自动兴趣区分析软件。该软件可实现对各环境要素注视情况的自动化批量识别，提高了数据分析效率（图3）。然而，由于自动语义分割的精度限制，一些与街道体验密切相关的重要要素和区域（如外挂店招、建筑入口空间、电子屏幕等）无法被精准识别出来，因此这些重要兴趣区需要人工识别和统计其累积注视时间，并将眼动仪前置摄像头记录的街道场景画面叠加统计数据以获得直观的可视化结果[46]。综上，本研究使用自动语义分割与人工识别结合的方法来统计各环境要素的注视情况。

再次，环境要素被注视的时长很大程度上受其视觉面积影响[47]，为了客观反映各要素对注意力的吸引程度，本研究引入“信息密度”（information density）这一指标——主观注视占比（注视比例）与客观画面面积占比（暴露比例）的比值③，用于描述单位面积下的信息输入效率，具体公式如下：

$$I_i=\left(\frac{D_i}{\sum_{j=1}^{n}D_j}\right)\Bigg/\left(\frac{A_i}{\sum_{j=1}^{n}A_j}\right),$$

式中，I_i表示第i个要素的信息密度；D_i是第i个要素的累积注视时间（ms）；$\sum_{j=1}^{n}D_j$是画面中n个要素全部的注视时间累积之和（ms）；A_i是第i个要素的客观画面面积；$\sum_{j=1}^{n}A_j$是画面中n个要素全部的客观画面面积之和。

其中，注视比例和暴露比例的取值范围均为0~100%。当信息密度为0~1时，表示注视比例低于暴露比例；>1时，表示注视比例高于暴露比例。如果把客观视觉面积占比的对象从单个环境要素扩大到其他对空间设计更有价值的兴趣区（如建筑底层空间、街道设施等），信息密度就可以广义地描述直觉系统的信息偏好，而无论是环境要素偏好还是兴趣区偏好的识别和系统干预，都可以帮助街道空间设计更好地引导使用者的注意力，从而诱发特定感受或行为。

③ 有文献计算过类似的指标，如参考文献［44］，但研究者未对此指标命名，仅叫做“二者比值”；本文中将该指标命名为“信息密度”。

表 1：各环境要素信息密度统计

环境要素		中段			东拓段			总体		
		注视比例	暴露比例	信息密度	注视比例	暴露比例	信息密度	注视比例	暴露比例	信息密度
低信息密度要素（0，1]	天空	1.57%	13.90%	0.11	1.86%	7.82%	0.24	1.71%	10.86%	0.16
	地面	13.04%	20.27%	0.64	9.65%	26.35%	0.37	11.41%	23.31%	0.49
	建筑立面	21.50%	26.60%	0.81	30.41%	34.85%	0.87	25.78%	30.73%	0.84
中信息密度要素（1，2]	人群	15.34%	12.52%	1.23	13.66%	10.17%	1.34	14.50%	11.35%	1.28
	绿化	9.70%	5.31%	1.83	3.30%	2.33%	1.42	6.50%	3.82%	1.70
高信息密度要素（2，∞）	品牌标识	2.86%	1.43%	2.00	2.56%	1.09%	2.35	2.72%	1.26%	2.16
	电子屏幕	5.70%	2.13%	2.68	5.30%	2.63%	2.02	5.50%	2.38%	2.31
	建筑入口空间	19.28%	9.04%	2.13	22.68%	8.08%	2.81	20.98%	8.56%	2.45
	外挂店招	9.20%	1.94%	4.74	/	/	/	4.74%	0.97%	4.89

3 研究结果

3.1 环境要素存在信息密度差异

目的地选择任务中，我们发现环境要素的信息密度存在显著差异。本研究按ADE20k街景标签数据库中的环境要素划分，确定了9类环境要素（天空、地面、建筑立面、人群、绿化、建筑入口空间、电子屏幕、品牌标识和外挂店招），统计各环境要素信息密度的平均值后，研究根据信息密度的高低将各要素划分为低（0，1]、中（1，2]、高（2，∞）三级（表1；图4，5）：外挂店招、建筑入口空间、品牌标识和电子屏幕的信息密度最高，其次是绿化、人群，建筑立面、天空和地面最低。

高信息密度要素中，外挂店招以绝对优势占据第一（总信息密度4.89），其客观视觉面积极低（0.97%），且仅出现在中段步行街中。而电子屏幕和品牌标识（总信息密度分别为2.31和2.16）也以极低的暴露比例（分别为2.38%和1.26%）获得了相当多的注视。建筑入口空间等具有透明性与开放性特征的环境要素，也以较低的暴露比例（8.56%）得到了较多的关注（20.98%）。

低信息密度要素中，建筑立面去除建筑入口空间、外挂店招等要素后的其余部分，虽然具有最高的总体注视比例（25.78%），但其总暴露比例也很高（30.73%），导致其总信息密度并不高；天空和地面的信息密度最低（总信息密度分别为0.16和0.49），尤其是天空，虽然在视野中面积也较大，却得到了极少的关注。

可见，视觉面积占比最大的要素不一定能得到最多的关注，而一些环境要素虽然视觉面积占比非常小，却可以吸引很多的关注。因此，我们有理由相信，当人们在街道中搜索信息时，倾向于关注建筑入口空间、外挂店招等反映空间活动内容和状况的要素；绿化和人群可能因为在一定程度上反映了环境的舒适度和场地的维护情况，也被认为提供了部分有效空间信息；视觉面积占比较大而信息较少的天空和地面则容易被忽略。

3.2 通过对街道高信息密度环境要素的系统设计塑造独特空间体验

自由行走观察任务中，研究发现步行街中段的外挂店招、电子屏幕等高信息密度要素集中出现在街道侧界面的中部。为了解这种设置对注视和感受有何影响，研究除统计街道侧界面、顶界面和底界面外，根据建筑楼层将侧界面进一步划分为上部（裙房顶层及以上）、中部、下部（底层）并统计其注视分布。

整体而言，街道侧界面的注视显著高于底界面和顶界面，而侧界面中东拓段的注视集中在下部，中段集中在下部和中部（表2；图6，7）。

具体来说，在步行街中段，虽然侧界面下部的注视最多（28.30%），但中部也获得了与下部相近的注视比例（24.20%）。现有研究表明，街道中人的关注主要集中在建筑底层空间[24][48]，即侧界面的下部，而中部和下部有同等被注视程度的情况并不常见。访谈中参与者对步行街中段的评价多提到“热闹繁华”“有趣丰富”“很有特色”等描述，并在

表 2：自由行走观察任务中街道各空间界面的注视比例

	街道空间界面				
	侧界面			顶界面	底界面
	下部	中部	上部		
中段	28.30%	24.20%	5.87%	1.60%	24.55%
东拓段	38.05%	16.16%	12.94%	0.15%	19.25%

被问到理由时，均提到“不常见的外挂式店招”“大型的电子广告屏幕”“有年代感的建筑装饰”等出现在侧界面中部的环境要素。此外，研究团队也注意到参与者在实验中除眼球运动外，也多次出现抬头、转头、转身等观察建筑物中高层空间的头部和肢体运动。一般人行走时水平方向的视域可达到两侧各90° 的范围，而向上和向下的垂直视域是非常有限的，为了看清楚行走路线，视轴线甚至会向下偏移10° [4]，因此行走时视线会相对集中在街道侧界面的下部。但当高信息密度要素被调整至一定高度时，人们会很快适应并对哪里可以获得更多信息产生新的判断，继而将更多注意力转向具有一定高度的区域。由于步行街中段将各类极具特色和年代感的外挂店招、电子屏幕集中设置在了街道侧界面中部，并包含了建筑山花、转角等丰富的装饰细节，创造了一个不同于一般街道的空间层次和注意集中区域，由此改变了人的基本预期，而视线仰角的变化也使得访客不断抬头、转头、转身等；在此过程中，参与者的视觉体验变得应接不暇，从而在一定程度上获得了“独特”“丰富”“鲜明”的景观风貌感受。

3.3 个人情感促进环境吸引力的形成

通过分类统计兴趣点拍照任务中参与者拍摄的照片，并结合拍照前的眼动注视轨迹和深度访谈，可以理解与环境吸引力相关的视觉认知过程。针对参与者提供的拍照理由，研究采用文本编码分析，依照环境吸引力的类型将照片分为核心印象（42张）、独特体验（32张）和个人经历（8张）三类（表3，图8）。

1）核心印象类照片反映了街道独特的场景或节点，同时参与者在深度访谈中提到“南京路特色”“别的地方没有”等描述，如“各式各样”“色彩丰富”的外挂店招林立的街道空间（7号照片），“人挤人”“浩浩荡荡”的密集的人群（13号照片）等。

2）独特体验类照片反映了参与者感知到的有趣的、与日常生活差异

表 3：参与者陈述的个人经历类照片拍摄理由

照片编号	拍摄内容	参与者陈述的拍摄理由	是否选为“最喜爱的三张照片”
3		“因为我觉得这个老凤祥的广告牌挺复古的……很像我小时候来上海见过的那种，当时电视广告里也有这种类似的，很有老上海调调。”	是
4		“路尽头可以看见世贸大厦，两个角的建筑比较特别，我印象里的南京东路就是这样，可以一直看到这个建筑。”	否
18		“因为我看前面路口是红灯嘛，就走得慢，路过这个小卖部仔细看了下，好像是卖上海雪花膏的，之前来都没注意，上回我朋友来我刚带她在别的地方买过，挺有意思的。”	是

续表见下页

表 3：参与者陈述的个人经历类照片拍摄理由

照片编号	拍摄内容	参与者陈述的拍摄理由	是否选为“最喜爱的三张照片”
19		“我刚来上海的时候，第一次来南京路就在这里和朋友合影过，当时我们几个玩到深夜，街上都没几个人了，确实很难忘啊，挺怀念当时的。”	是
与 19 号照片拍摄内容相同，因此未编号		“我读本科时候，我妈来看我，当时带她来南京路就在这拍过游客照。”	否
22		“这张我是想拍前面的外滩中心大楼……我小时候来这儿最喜欢的就是这个楼，特别是晚上，觉得顶上的灯亮起来特别好看，确实也很少楼会做这种造型吧。”	是
26		“前面就是和平饭店啊，快到外滩了，上回我和朋友在外滩坐了很久，我每次来都觉得晚上灯亮以后这个楼的绿顶还是很好看的。”	是
27		“这个建筑特别像我前阵子看的一个电视剧里的，就是主角的办公楼，几乎一摸一样，我都怀疑是不是就是在这取景的，特别这种纹理特别像。”	否

较大的内容，如独特的橱窗装饰（2号照片）、穿着熊猫玩偶服招揽客人的店员（8号照片）等。

3）个人经历类照片反映了能够激发参与者联想的场景或节点，参与者提到“曾经来过”“与熟悉的影视作品相关”等描述，如街道尽头的世贸大楼（4号照片）、外滩中心大楼（22号照片）等。

虽然个人经历类照片数量不多（仅占照片总数的10%），但相对其他两类照片，其被选为“最喜爱的3张照片”的比例最高，高达63%，远高于核心印象类照片的33%和独特体验类照片的6%。这说明当体验与个人经历产生关联时，容易形成情感连结，从而成为整个街道体验中参与者“最喜爱”的部分。

结合眼动轨迹，可以进一步理解个人经历类照片被选中的原因。19号照片（图9–1）是一名参与者在中段入口的标志性元素——“南京路步行街”大理石标识牌附近拍摄的照片，拍摄前3s的注视顺序（图9–2）说明其首先注意到标识牌前拍照的人和被拍的人，接着观察了标

识牌，从而对人的拍照行为进行理解，之后又反复观察了人和标识牌，在此过程中联想到曾经拍照的经历，最后观察了花坛以考量通行的可能，最终决定就地拍摄。

实验结果显示，当环境体验与个人经历相关并被赋予情感价值时（即产生了情感共鸣），容易导致正向的认知判断结果，促进环境吸引力的形成。

4 结论

本研究采用眼动追踪技术，直观理解从环境要素到行为决策的认知过程，为街道的空间体验研究提供了新的思路。同时，通过眼动数据分析结合深度访谈的研究方法，将定性研究中个人经历、情感记忆等非物质的认知因素与具体的物质空间建立起直接联系，为此类因素的分析提供了有效手段。此外，研究在真实场景中设置与街道体验相关的任务，并针对特定行为目的尝试理解认知加工过程中不同要素间的关系，利于更好地理解整个认知过程，拓展了现有的主流研究路径。本研究的主要结论包括以下几点。

1）以南京路为代表的商业步行街中，环境要素具有显著的信息密度差异。在南京路步行街中，外挂店招、建筑入口空间、电子屏幕、品牌标识等信息密度最高，其次是绿化、人群。高信息密度要素，尤其是外挂式店招的客观视觉面积占比极小，但对人的综合空间感受的影响较大，在相关空间设计中应当予以重点关注。

2）高信息密度要素的系统性设计或调整容易形成独特风貌的感受。在南京路步行街中，外挂店招等高信息密度要素的系统设计，大大强化了建筑物中部的受关注程度，从而形成了不同于其他街道的景观风貌感受和特质，大大提高了设计干预的效果。上海市黄浦区政府近年在对南京路进行系统改造提升时，也坚持对建筑物中部以外挂店招为主的环境要素进行保留修缮，此举有效延续了步行街的核心特色和景观风貌。在其他城市空间设计中，也应注意识别此类高信息密度的环境要素，以及它们对使用者空间认知和行为活动的影响，从而促进特定感受和行为活动的发生，提高空间环境的整体设计品质。

3）值得注意的是，在引起使用者关注的空间中，能够进一步诱发情感体验的场所更容易使人留下深刻印象，并导致正向的认知判断。在南京路步行街的体验中，环境线索不断激发多种记忆——包括感官记忆和情绪记忆——参与者通过街道空间联想到了对自己有重要情感意义的个人经历，强化了个人和场所之间的情感关联，从而促进了环境吸引力的形成。在规划设计时，我们可以采用具有识别性的环境线索（如统一、反复出现的特殊标识、街道家具等），从而提示并强化这种情感关联。

基金项目

· 国家自然科学基金“基于实景体验计算的城市街道景观风貌精准修补”（编号：51878461）

图 1. 实验路段及主要任务环节选点空间分布
图 2. 实验流程图
图 3. 环境要素的语义分割与兴趣区自动批处理
图 4. 中段方向各环境要素信息密度
图 5. 东拓段方向各环境要素信息密度
图 6. 中段街道侧界面上、中、下三部分的注视分布
图 7. 东拓段街道侧界面上、中、下三部分的注视分布
图 8. 核心印象、独特体验、个人经历三类照片拍摄情况及被挑选为“最喜爱的三张”的情况（图中未展示参与者拍摄的全部照片，去除了拍摄内容相同的部分）。
图 9. 图 9-1 为参与者拍摄的照片，图 9-2 为叠加注视（圆）和眼跳（线）的眼动轨迹画面。