J. Vis. Commun. Image R. 64 (2019) 102639

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

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

Cityscape protection using VR and eye tracking technology ☆

Le-Min Zhang [a,b,*], Ruo-Xi Zhang [a], Tay-Sheng Jeng [b], Zi-Yuan Zeng [c]

[a] School of Architecture and Civil Engineering, Xiamen University, Xiamen, China
[b] Department of Architecture, National Cheng Kung University, Tainan, Taiwan
[c] Zhihuijia Tech Co,. Ltd, Xiamen, China

ARTICLE INFO

Article history:
Received 25 May 2019
Revised 6 September 2019
Accepted 7 September 2019
Available online 7 September 2019

Keywords:
Urban renewal
Three-dimensional eye tracking
Cityscape feature identification
Human-oriented
Smart city

ABSTRACT

The traditional method for reconstructing cityscape relies greatly on the subjective judgment of designers, which makes the cityscape simple and homogenized. This paper aims to propose a new integrated approach to protect and design cityscape based on virtual reality (VR) and eye tracking technology. Through the integration and quantification of the eye tracking data and the protocol analysis data in the VR environment, this research has revealed the mechanism of identifying the cityscape features, and discovered the differences in the perception of the cityscape features by different people, thus proposing the multi-cultural integrated strategy for protecting cityscape. This research is of great significance for building a human-oriented scientific planning and protection method and promoting the application of cutting-edge digital technology in the field of smart city governance.

1. Introduction

The renovation and restoration of historical districts has gradually become an important work in the construction of major cities in China. Furthermore, under the current background of promoting urban culture, city renewal has attracted even more attention as a reflection of the cultural characteristics of a city [20]. However, in the construction of most cities, there are many common problems: the renovation of many urban historical areas has deviated from the environmental and cultural background. Specifically, many antique streets have become false bodies and lost the connotation of street culture [17]. The traditional protection and design of cityscape relies too much on the subjective judgment of planners, and lacks the technical operation method of human-oriented accurate identification of cityscape, which easily leads to the simplification and unilateralization of protection principles and renewal strategies. Therefore, it has become the most urgent challenge for the protection and renewal of historical blocks about how to systematically, objectively and scientifically identify the features of urban historical areas and maintain their comprehensive cultural values.

Virtual reality (VR) technology offers new methods for the simulation and interactive research in urban space and architectural design, thanks to the higher-quality environmental simulation and spatial interaction. The VR technology is beneficial for the analysis and protection of historical areas [7]. Space and human activities are eternal objects and topics in the research of architecture and urban planning. With the development of information technology, the data collection and analysis methods are also making progress in spatial behavior research at home and abroad, which has made the urban planning management and architectural design more intelligent and humanized.

Based on virtual reality, eye tracking technology and protocol analysis, this research integrates a set of solutions for urban design and management. Based on the interdisciplinary knowledge of urban planning, architecture, cognitive psychology and behavior, the technical route and experimental platform are constructed. Through the real-time recording, data calculation and visual quantitative analysis of the eye tracking recognition and spatial behavior in the VR environment, this research uses the protocol analysis to find out the real observed and psychological feelings of the people, sum up the influences of cityscape features on the visual psychology and recognition behavior of the people. Finally, this paper proposes more targeted mechanisms and methods for the design and protection of cityscape.

2. Related work

2.1. Eye tracking and spatial visualization

The acquisition of spatial information is mostly achieved through vision. The research on visual attention and its cognitive

☆ This paper has been recommended for acceptance by 'Maofu Liu'.

* Corresponding author at: School of Architecture and Civil Engineering, Xiamen University, Xiamen, China.

E-mail addresses: zlmzj@126.com (L.-M. Zhang), zhangruoxi@xmu.edu.cn (R.-X. Zhang), tsjeng@mail.ncku.edu.tw (T.-S. Jeng), zengziyuan@zhjvr.com (Z.-Y. Zeng).

https://doi.org/10.1016/j.jvcir.2019.102639
1047-3203/

processes has received much attention in the field of spatial cognition [6]. The research of eye tracking began over 100 years ago, and it has been greatly concerned by the academic community because it can significantly improve the objectivity and effectiveness of cognitive psychoanalysis. At present, eye tracking has been extensively applied to the fields of psychology, behavioral cognition, environmental cognition and media interaction. However, eye tracking is rarely investigated in the field of urban planning and architectural design because the research on the cities and buildings usually requires three-dimensional and dynamic perspective. Therefore, the traditional eye tracking technology is far from meeting this requirement. At the same time, traditional eye tracking analysis requires a lot of effort in data acquisition and analysis [5]. Secondly, even if a large amount of data is available, it is difficult to visually integrate the eye movement data of different users. Finally, the eye movement data used in traditional thermogram analysis is two-dimensional, making it difficult to analyze the front and back features in the 3D street space. Therefore, how to find the eye tracking data and the user's spatial behavior for rapid acquisition, as well as conduct convenient deep integration is of great significance for the research of spatial behavior of urban planning and architecture.

However, with the rapid development of digital technology in recent years, the use of eye tracking technology has gradually begun to be applied to the field of architecture and urban planning for spatial cognitive visualization research, and certain results have been achieved. At present, the main research directions of this field include: eye movement perception analysis in urban and natural landscapes [2,18], urban spatial wayfinding analysis [14,16], architectural perception and spatial evaluation [3], perceptual research on the urban space that uses different vehicles [11].

2.2. Advantages of eye tracking research in VR environment

Virtual Reality (VR) is a system simulation environment featured by multi-source information fusion, interactive 3D dynamic vision and physical behavior. VR technology provides new method for the simulation and interactive research of urban and architectural space. Some scholars use VR technology to conduct spatial cognitive research on cultural tourism experience, helping the experiencer to obtain a better urban cultural experience [4]. Besides, VR technology has also achieved good results in multi-population participation in architectural design [13].

The combined research of VR technology and 3D eye tracking has been slowed down in the early years due to VR technology [6]. With the rapid development of VR technology in 2015, most scholars believe that combining VR technology with eye tracking, i.e. the research on eye tracking in VR environment, will have a promising future, which can greatly promote the analytical research and application of cognitive behavior in 3D space [5]. As VR head-mounted display can obtain 3D positioning information, data processing can quickly and accurately superimpose the eye movement data and user's spatial activity data. The combination of the 3D model component in the VR environment and the spatial positioning of the gazed object can greatly improve the accuracy and efficiency of eye tracking analysis.

Since 2017, the combination of the VR technology and 3D eye tracking has been rapidly developing, such as wearing a portable eye tracker in the VR environment for the performance evaluation of indoor navigation system and building a new method for 3D gaze analysis and visualization [15], combining the two technologies to construct an experimental environment for shopping behavior research [12], integrating protocol analysis, eye tracking and VR technology to redesign urban space [8]. The experimental design, technical application and data analysis of these studies have laid an important foundation for future research.

3. Research method

3.1. Theoretical methods

On the basis of systematically analyzing the existing urban and rural planning and architecture as well as the cityscape recognition theory, this study integrates the theories and quantitative analysis methods of cognitive psychology, spatial behavior and sociology. Further, aiming at the visual perception and psychological process of cityscape feature recognition, this study constructs the interdisciplinary theoretical knowledge system and research methods.

3.2. Experimental methods

The data used for the study of urban regeneration not only include the existing infrastructure, but also users and other different data. The key to the research is to establish an experimental platform to quickly integrate and analyze multi-dimensional data. This research combines VR system and eye tracking technology to carry out experimental design, and independently develops VR eye tracking experimental platform. In the process of the experiment, the 3D eye movement data, spatial behavior data and protocol analysis data of the experiencers are dynamically superimposed and analyzed, in order to understand the recognition and perception of cityscape by the crowd. The main experimental methods and steps are as follows:

(1) Experimentally find out the distribution of spatial visual attention of the public in traditional streets and lanes, analyze the causes and laws, summarize the identification mechanism, and classify the features that receive much attention.
(2) Analyze the differences and tendencies of the people with different cultural backgrounds in identifying the characteristics of cityscape features, conduct comparative study on the classification of the streets, find out the weight relationship of the traditional street features in southern Fujian and build a quantitative indicator system.
(3) Assist in guiding the method for redesigning traditional streets and lanes, and finally build an auxiliary system platform for traditional street protection and renewal design based on human-oriented and multicultural values, thus realizing the auxiliary design application of VR 3D eye tracking platform.

3.3. Key technology: 3D eye tracking technology in VR environment

The technology of 3D eye tracking technology in VR environment is independently developed for this research. The developed technology can realize the combination of spatial depth perception and visual perception through the 3D eye tracking technology that quantitatively tracks the human eye's visual sensitivity. As a result, the 3D eye tracking simulation is closer to the crowd behavior in real space. The developed technology plays an important role in improving the accuracy and usability of the simulation experiment.

The technology is developed based on the theory of visual sensitivity distribution and the principle of optical reversibility. The distance between the human eyes and the visible object is different, and the visual range is different. Therefore, the visual sensitivity will change dynamically. According to the study, visual acuity will decrease by about 50% every 2.5° from the center to 30°, and it will have greater decrease beyond 30° [1]. Based on the traditional eye tracking technology, the 3D eye tracking technology in virtual environment is developed in this research. This technology defines the eyes as a cone-shaped source (Fig. 1). The visual energy value of the object is: the time of gazing multiplied by the cumulative value of the visual acuity. As shown in Fig. 1, the cone-shaped line of sight simulates the gradual change of visual acuity. A 3D visual energy thermogram is generated by accumulating the energy of visual light particle, which can be calculated by:

$$C = \sum_{i=1}^{n} T \times V, \quad i = 1, 2, 3 \ldots\ldots n \tag{1}$$

Here: the visual where C is thermal value, T is the time of gazing to the object and n is the tracking scan order.

$$V = (1/2)^{\alpha/2.5^\circ}, \quad \alpha \in 0 \sim 30^\circ \tag{2}$$

where V is the visual acuity and α is the viewing angle.

4. Experimental design

4.1. Overview of real cases

The ancient city of Zhangzhou is a typical representative of the historical and cultural cities in southern Fujian, China. The ancient city is located in the center of Zhangzhou, Fujian Province. It covers an area of about 86 ha and has a long history of more than 1,300 years. The ancient city has unique historical cityscapes and architectural features. At the same time, local residents have lived here for thousands of years, so that they have distinctive regional characteristics and rich cultural values. At present, with the rapid development of economy, the transformation and renewal of the ancient city of Zhangzhou faces many problems, and the traditional architectural features are undergoing great changes and impacts [19]. For example, the Hong Kong Road fragment is the core protection area of the ancient city. Although the Hong Kong Road has won the United Nations Habitat Environmental Protection Award, it has been damaged to varying degrees due to the aging and the life activities of the local residents. Therefore, how to interpret the characteristics of Hong Kong Road, how to extract the regional culture in the traditional street space of southern Fujian and conduct in-depth analysis from the perspective of crowd spatial behavior and perception is of great necessity for promoting human-oriented protective planning and integrating urban multiculturalism [21].

4.2. Construction of research platform

The platform is a multi-line system consisting of three parts: server, analog end and data acquisition end. The data can be quickly transmitted through networking. Through the Unreal engine rendering, the virtual scene of Hong Kong Road is established, and the virtual helmet and eye tracker are used to collect the spatial positioning data and the eye tracking data. Finally, the

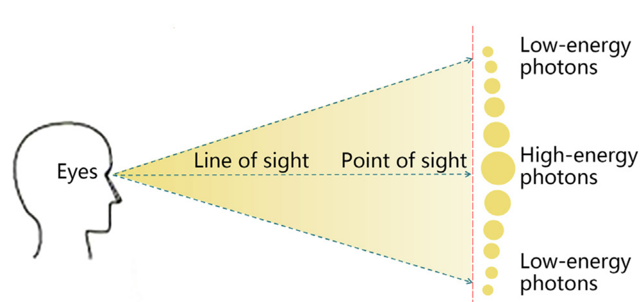

Fig. 1. The cone-shaped line that simulates the change of visual acuity using eye tracking, as well as the 3D visual gaze heat map generated by the energy of visual light particle.

visualized 3D distribution data of the crowd's attention under the virtual scene is obtained through software operation. For the analog end, the VR helmet (device: HTC-Vive) and eye tracking technology (device: 7Invensun aGlass eye tracker, programming: C++ and Unreal Engine) are used to build virtual scenes while simultaneously tracking the targeting users in the VR environment in real time. For the data acquisition end, the experimental data is recorded and summarized, and the user's space behavior is visually analyzed in the form of 3D point cloud in real time (Fig. 2).

The core of the experimental system is the acquisition and spatial localization of 3D eye movement data. In the VR eye tracking experiment, the data acquisition and processing of the 3D eye point cloud is as follows (Fig. 5). Firstly, the positioning data of the VR helmet and the eye tracking data are synthesized to obtain the spatial positioning data of the gazed object; then, the spatial positioning data is superimposed with the 3D model component in the VR environment; finally, the data acquisition program is written in C++ on the Unreal platform, which integrates the tracking data of eye movement, beating and smooth trailing and the spatial behavior data and converts it into the thermal model and distribution map in the form of 3D point cloud.

4.3. Construction of VR scene

Streets are the habitual, accidental, or potential movements of observers. As people observe the city on the move, other environmental features are also laid out along the road, and the images are enhanced by typical spatial features and special facade features [10]. Therefore, the key to the construction of the research scene is to realize the organic integration of the "human-space-interface" for analysis and design, and to avoid strengthening one form or feature in isolation.

The ancient city of Zhangzhou is a typical representative of the most important historical and cultural ancient cities in southern Fujian of China. The Hong Kong Road is a core protection area of the ancient city of Zhangzhou. The virtual Hong Kong Road is generated and simulated based on the real Hong Kong Road. The length of the street is 95.2 m. There are two stone arches in the middle, forming an important space node. The distance between the arches is 26.8 m. The width of the street is 3.2–5.8 m, and the height of the buildings on both sides is 3.4–9.2 m. In order to improve the accuracy of data acquisition and the efficiency of data processing, the experiencers are only allowed to take activities in the street between the two arches. Firstly, a 3D scanner is used to perform 3D scanning and mapping on the streets and lanes of Hong Kong Road. Then the 3Dmax software is used to modify the architectural and environmental models of the experimental sections. Finally, the simplified model is imported into the unreal engine to realize the construction of VR scene. In order to realize the data acquisition and analysis of the eye movement experiment, this experiment explores each scene feature in the model as a single module component and numbers it in groups. Finally, the components of the experimental location in the VR scene are summed up to total 3348 groups (Fig. 3).

4.4. Sample selection

This research uses random sampling and selects 100 experiencers as experimental subjects. Due to the visual limitations of the 18 participants, data from 82 (36 males and 46 females) are selected for final analysis. The age of the 82 participants is between 18 and 50 years old, and the average age is 24.5 years old (standard deviation of 5.3). Their average time of living in southern Fujian is 5.1 years (standard deviation of 6.9), of which 27 people have a good understanding of southern Fujian culture, 31 people have general understanding, and 24 people don't know much.

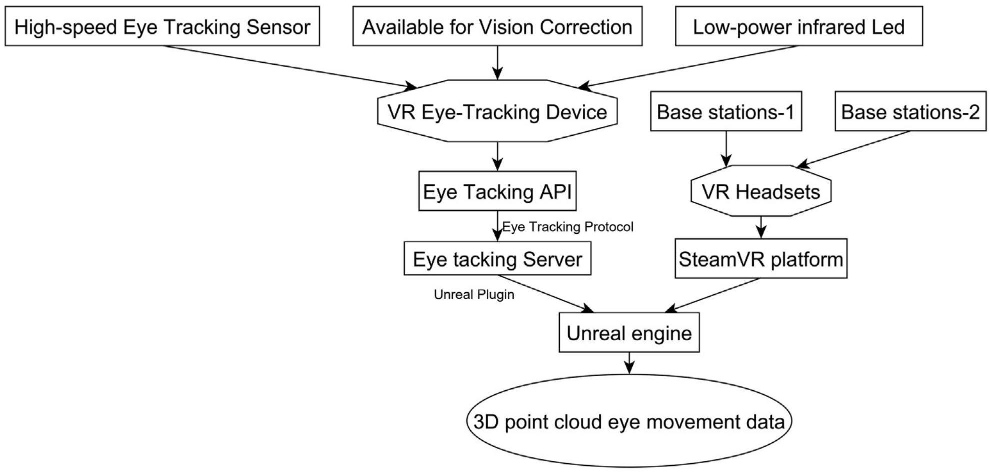

Fig. 2. System architecture diagram based on VR and eye tracking technology, and the data acquisition and integration flow chart of 3D eye movement data.

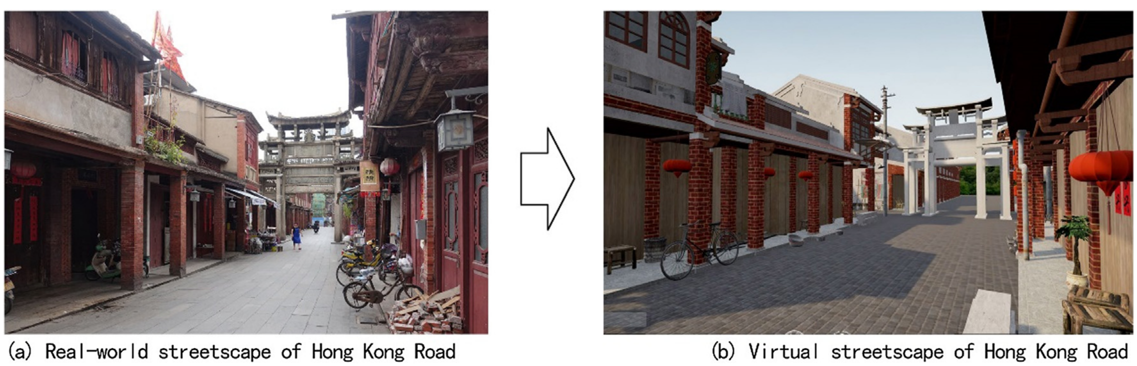

Fig. 3. Construction of the VR experimental scene of Hong Kong Road.

4.5. Experimental setup for protocol analysis

The research uses three methods of eye tracking, protocol analysis and questionnaire survey to integrate and analyze the data. This research aims to track the cognitive thinking, spatial perception and behavioral activities when people are traveling in VR environment. At the same time, the real-time eye movement data is tracked in the protocol analysis, and the data is acquired through the combined method of concurrent and retrospective verbal protocol analysis, which greatly improves the efficiency and accuracy of the protocol analysis.

Before the experiment, the basic information and cultural background of the participants is recorded through the questionnaire, and then they are allowed to get familiar with the VR environment and operation, which takes about 3 min. After wearing the experimental equipment, the participants need to first perform eyeball matching and necessary optical adjustments. Then, they are told that they are asked to answer the following three questions in the experiment:

- Please look for the most representative features of southern Fujian architectural culture in the environment you are staying in?
- What do you think is the most attractive feature of southern Fujian architectural culture that appeals to you?
- Please confirm that whether the scale of the street space is appropriate?

In the experiment, each participant has 2 min of experience, while telling their own experience and answering the questions. The system background records the eye movement data in real time and visualizes it on the monitor, while the staff records the concurrent verbal data of the experience activity in the background in real time, and compares the two parts of the data for the subsequent protocol analysis.

After the experiment, the data acquisition terminal is used to visualize the dynamic 3D eye movement thermogram of the participants during the experiment. The participants are guided to review the experience process and check the objective eye movement data and subjective psychological description. Besides, the experiment organizer can check and query some special eye tracking phenomena, for example: Why have you been here for so long? What did you look at here just now? Do you like this?

5. Analysis of experimental data

Through experiments, we can clearly grasp the user's spatial activities and concerns. The data we have acquired is in 3D form. This virtual cityscape integrates VR and eye tracking data, which allows us to check the user's attention while they are walking on the street. We can get the attention data, process it and visually reflect it on each building feature. In order to facilitate the identification process during data analysis, the entire street environment is uniformly rendered as a gray tone. Then, each building feature will display different colors according to the value of the attention. Therefore, red represents great visual focus, and yellow represents medium focus, green represents low focus, and gray represents no focus (Fig. 4).

At the same time, through the data segmentation and combined analysis, the spatial behaviors of the groups with different backgrounds can be further analyzed. The eye tracking process generates a huge amount of data. That is, a two-minute experience by a user will generate a data set of about 60,000 spatial point clouds. How to analyze the data to find out the law and relationship is very important. Through data aggregation analysis, it is found that there are 82 features with high degree of attention in the experience, accounting for only 2.4% of all 3384 model objects in the scene. This value reflects that the visual behavior of people in the street space has a very high commonality and regularity. Among the 82 features, there are different types of cityscape features that we usually think of, as well as some types of features that are not usually paid attention to. Besides, there are also large differences in the interested features and the spatial positions of people with different backgrounds.

Analyzing and mastering the similarities and differences between the various groups of people in understanding the cityscape features will help establish a bridge between the "human" and the "urban space", which is the basis for showing the respect for multicultural values and promotion in urban design. This research will make the humanistic scale control more objective and precise in urban design, which makes the street protection mechanism truly human-oriented and better achieves the integration of multiculturalism.

5.1. Comparative analysis of professional users and naive users

According to the architecture knowledge of the participants, we divide the participants into the professional group and naive group. According to the questionnaire, the professional group: 27 people have architectural background (17 local architects, 10 foreign architects), naive group: 54 people have no architectural background. Through the analysis of the eye tracking thermogram data of the participants, it is found that these two groups vary greatly in the recognition (Fig. 5). This is because the background of architectural expertise has a direct impact on the environmental aesthetics, spatial cognition and spatial behavior of the experiencer, but what is the degree of this impact? Therefore, we put forward a hypothesis: We believe that the professional group is more concerned with the characteristics of the street and the cityscape than the naive group, and the former group can see more features of the cityscape features.

In order to prove the above point of view, we conducted a multi-dimensional segmentation analysis on the experimental data of the two groups to find out the differences and laws between the two groups. Fig. X shows the scatter-regression analysis of the spatial behavioral data of the crowd, where blue represents the professional group and red represents the naive group. We find significant differences in the spatial height distribution of viewpoints between the two groups. That is, the viewpoint height of the professional group is significantly higher than that of the naive group. In this study, SQL (Structured Query Language) is used to extract the required time and viewpoint height in the MDB database. The X axis is used as the time axis and the Y axis as the height axis to perform scatter regression analysis on these eye movement data. The analysis indicates that, the average viewpoint height of professional group is 3.83 m, while that of naive group is 1.98 m. Obviously, the average viewpoint height of the former is significantly higher than that of the latter. At the same time, the trend lines of the two groups are relatively flat.

The eye tracking data of the professional group show that they focus more on the distant features, in the order of wooden bucket arches, cornices, doors and windows, colonnades, archways. The protocol analysis shows that the group gives more attention to: red bricks, roofs, windows, and archways. The viewpoint of the professional group is relatively higher, with the average viewpoint height being 3.83 m. Besides, the characteristics of the second floor are more concerned (Fig. 5a) and the number of the found features is higher in average. The local architects show more obvious features than the foreign architects. In particular, the eye movement data of 2 experts showed that more than 73% of their time was spent on the second floor, which was significantly higher than other participants. The protocol analysis shows that the professional group believes that the spatial characteristics of the southern Fujian streets are mainly distributed in the roof and the top of the column head.

The eye movement data of the naive group shows that they focused more on the nearby features, in the order of colonnades, wooden bucket arches, lanterns, archways, nostalgic small objects (Fig. 5b). Surprisingly, eye movement data show that the naive group could find most of the basic features of the building and street space, but they were not conscious or not too interested. The naive group has a lower viewpoint and an average viewpoint height of 1.98 m. As shown in Fig. 5b, the protocol analysis shows that the naïve group gives more attention to red bricks, color, arcades, and archways. Most of the non-professional participants focused on the street features distributed on the first floor, showing a strong interest in the nostalgic small objects. This group could quickly find the architectural features of the first floor, rather than

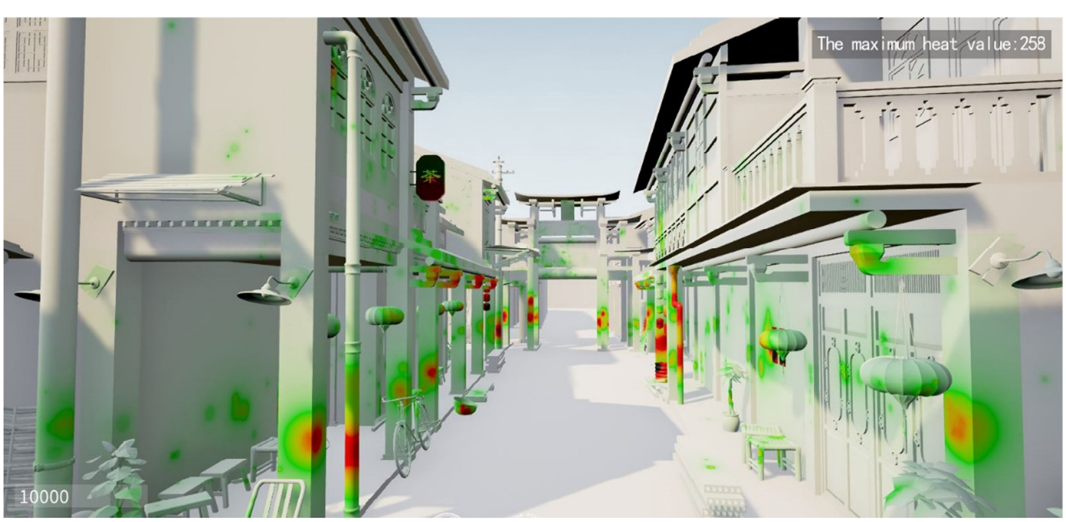

Fig. 4. Behavior heat map in 3D eye movement space.

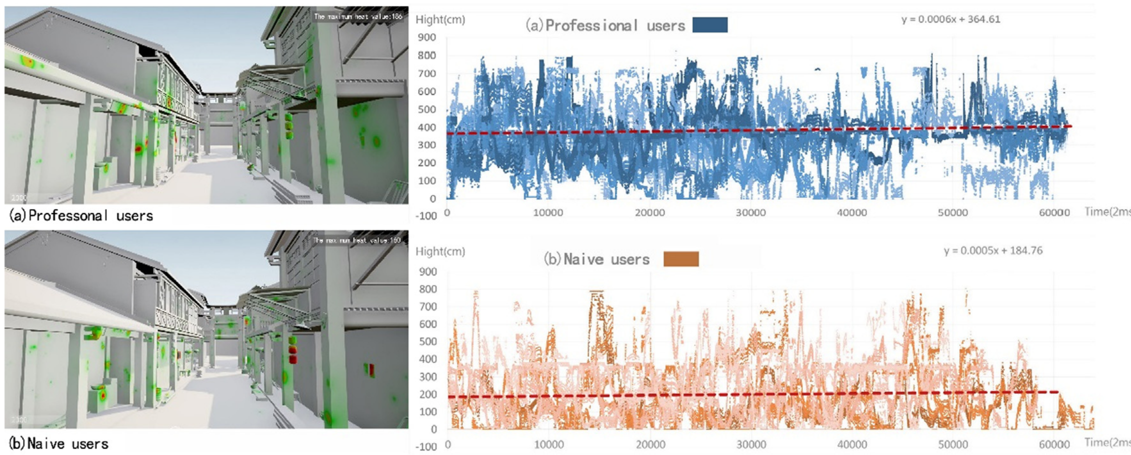

Fig. 5. Comparative analysis of the dynamic viewpoint height and scatter of architectural and non-architectural samples. Their average viewpoint heights are 3.83 m and 1.98 m, respectively, which differs by 1.85 m. The average viewpoint height of the profession group is significantly higher than that of the naïve group. The trend line of the scatter of the profession group is Y = 0.0006x + 364.61, and that of the naive group is Y = 0.0005x + 184.76.

those of the second floor and above. At the same time, it was found that the naive group was easily interfered by those non-cityscape features, such as the inappropriately-constructed canopies or water pipes.

The comparative study found that the professional group tended to focus on the farther and higher architectural features. This is because the professional group prefers to grasp the overall architecture and street space as a whole because of professional training and they like to look for the characteristic features in the spatial combination and aesthetic form of the street. For example: the consecutive sets of important grille windows on the second floor on the east side between the two arches have successfully received the attention from 75% of the participants from the professional group due to the simple shape and appropriate size. Both the local and foreign architects gave almost the same attention to these windows.

In contrast, the naive group favored attention to the architectural features near the perimeter and lower height. This is because the naive group is more inclined to pay attention to the experience of life, and those environmental features that can produce emotional resonance can stimulate their interest. For example, those nostalgic features of life have a great influence on triggering the homesickness emotion of people. Especially, those small objects on the first floor (bicycles, bamboo chairs, couplets) will cause them to stop, observe carefully, and associate these objects with their experience. At the same time, we found that the naive group without professional training had certain intuition recognition ability for the basic street features, but people rarely realized it and they would not appreciate the features. Therefore, the guidance tips may be required to achieve that purpose.

Therefore, it can be concluded that: in the identification of street space and architectural features, the professional group tended to grasp the whole scene, while the naïve group preferred the nearby features, and different groups had their own unique features. The architectural style is a straightforward voice. For the beautiful and distinctive street features, almost everyone is able to perceive it. The key to the identification is that the people and the identified objects need to resonate. In 1974, Lefebvre distinguished between "architectural space" and "space of architects" in space production. "Architectural space" benefits from the people's experience and is one of the ways to produce social space. "Space of architects" is the operation of the space by those architects through their professional practice. Lefebvre believes that the space of architects is not neutral, not innocent, because the training of the architect's drawing skills, in fact the entire training, puts the vision on top of other senses. Maintaining the hegemony of images for a long time, replacing the reality with the wonders, is out of the aesthetics of the general public [9]. Therefore, the design and protection of good cityscapes should not be limited to the subjective experience of architects, but should be more concerned with the acceptance and preferences of the public.

5.2. Comparative analysis of subjects with varying degrees of familiarity with the southern Fujian culture

The eye tracking data of the samples that have no understanding of southern Fujian culture (including the samples that have not so good understanding and have no understanding) shows that the sample paid more attention to: arcade colonnades, archways, arches, life objects, lanterns, bicycles and other objects. The protocol analysis shows that the samples had more attention to: red bricks, arches, lanterns, and bicycles. Five samples considered the small objects to be more attractive, and it would be more attractive if there were performances and business activities. The protocol analysis of eye movement data is basically consistent, showing that the samples were more concerned with some nostalgic small objects, although some of them were not unique to southern Fujian culture.

The eye tracking data of the samples that have moderate understanding of southern Fujian culture shows that the more attention is paid to: arcade colonnades, archways, wooden arches, couplets, billboards and other objects. The protocol analysis shows that the samples paid more attention to: red bricks, archways, billboards, and small pavilions. Four samples thought that the couplets on the door were very cultural, which was very attractive to them. The eye movement data shows that the samples also focused on the wooden bucket arches and water pipes, but the samples did not say that they were distinctive and why they were attracted. At the same time, the eye movement data and protocol analysis found that the samples were more concerned with the cultural features in advertising and couplets and they also liked them. Besides, three samples suggested that they were attracted to small-scale street space, and felt that it was pleasant in such a spatial scale.

The eye tracking data of the samples that have good understanding of southern Fujian culture (including those that have very good understanding and those that have good understanding) shows that the samples were more focused on: red bricks, wooden arches, archways, cornices, and windows. The protocol analysis shows that the samples paid more attention to: spatial scales, material textures, and feature structures. The samples paid more attention to the scale and details of the space as they considered that the spatial scale is closer to daily life, full of market atmosphere and human touch. Several samples suggested that the details of some objects were incorrect, or the spatial proportions and scales were different from those of their hometown. Even some samples from the Quanzhou area would question the features with no swallowtail ridges do not belong to southern Fujian culture. At the same time, it was found that about one-third of the samples had low eye-visual scanning rate, and the duration of attention to some objects became longer. This phenomenon meant that the samples were interested in the region, or there was more information to be interpreted [5], which is consistent with the protocol analysis showing that the samples were scrutinizing the details.

6. Multi-cultural integration of street protection strategies

It is necessary to integrate multiculturalism to adopt a dynamic protection strategy for different levels of street protection areas based on their actual conditions. At present, many cities have protected and managed historical areas, but most of them tend to have a strict control over the core areas of the cityscapes, and they are accustomed to apply this method to the management of all levels of protection areas, while ignoring the diverse social needs of the cityscapes of different levels. Therefore, the protection of the non-core areas lacks the necessary methods. Therefore, this research aims to establish a multicultural integration street protection strategy by identifying different people's different needs for the cityscapes.

Firstly, we investigate the similarities and differences of the visual attention of people with different degrees of understanding for southern Fujian culture. The similarities and differences of the attention patterns of different groups of people are identified by using the eye movement data and spatial behavior data. Then, according to the previous research, the clusters of basic features and the clusters of different features are generated. Finally, in the different levels of control areas, the needs of different social groups and the actual conditions are considered to propose a targeted dynamic protection strategy for cityscapes, as shown in Fig. 6.

(1) Classification of cityscape features

In the protection and renewal of historical cities, the key to building a traditional street protection mechanism for multicultural integration is to focus on human-oriented direction, scientifically strengthen the participation of multiple publics, and appropriately weaken the subjective experience of architects. Using visual attention to analyze the psychological feelings of different people, and understanding the cityscape protection from the perspective of multiple people is of great significance for expanding the scientific design of cityscape protection.

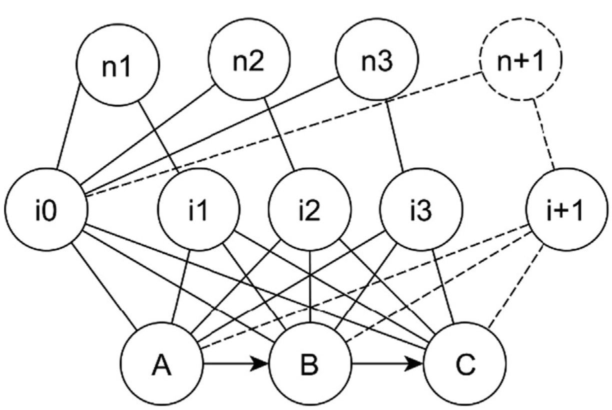

Fig. 6. n1, n2, n3...: people with different cultural backgrounds; i0: basic cityscape features, i1, i2, i3...: classification features of different cityscape features; A: core area of cityscape protection, B: buffer area of cityscape protection, C: coordinated development area of cityscape protection.

The experiment has shown that the street space understood by the architects is different from that by the public. The people in different cultural backgrounds have certain differences and weights in the street space they care about and understand. According to the data distribution trend of the characteristics of the people in south Fujian, we found that the focused features of different levels of cultural understanding showed a trend from fuzzy to specific. In order to find out the similarities and differences of people with different cultural backgrounds, we have carried out classification and superposition analysis of the experimental data, and scientifically and objectively classified the features of cityscape. Specifically, the features that received high degree of attention and showed small fluctuation with the cultural background represent that the features were widely concerned by all the experiencers and is the basis for determining the cityscape. We classify such features as basic features. At the same time, we found that from the people who are not familiar with southern Fujian culture to the people who are familiar with southern Fujian culture, the data of the objects they are concerned show more obvious distribution laws. According to the distribution laws, we classify the corresponding features of each group as preliminary cityscape features, cityscape enhancement features and original cityscape features.

In order to better understand the identification of the architectural and cityscape features by the people with different cultural backgrounds, we performed a double bubble analysis on all the experiencers and sub-items. The familiarity to southern Fujian culture is used as the Y-axis. The comprehensive questionnaire divides the experiencers into five categories (very familiar = 5, familiar = 4, moderately familiar = 3, not so familiar = 2, not familiar = 1). The cityscape features and the content of the experiencer are classified and used as the X-axis. The blue bubble represents the attention of the total crowd, the red bubble represents the attention of the five categories of people, and the larger bubble indicates higher degree of attention (Fig. 7).

- Basic cityscape features. This kind of feature is the consensus of all the people's cognition, and it is the basis for understanding the traditional street style of southern Fujian culture. In the study, we found that the features of the arcade colonnade, archway and wooden arch are generally concerned by all kinds of experiencers.
- Elementary cityscape features. People who don't know much about southern Fujian culture have less attention to the features of life or nostalgia. They tend to feel the fresh life. We found that the features such as daily necessities, lanterns, chairs and nostalgic bicycle showed strong attractiveness and nostalgia (except for basic cityscape features), and even some features of non-local culture will arouse people's interest. The spatial features of streets are very attractive to ordinary experiencers.
- Cityscape enhancement features. The experiencers who have a certain understanding of the local culture have certain preferences for the cultural characteristics of the cityscape features. This research found that they pay much attention to and show strong interest in the features such as couplet, advertisement and cultural decoration. Such cityscape features are helpful to enhance the cultural atmosphere of the street space.
- Original cityscape features. Those who have a better understanding of the local culture show higher demand for seeing the original cityscape features. The study found that the experi-

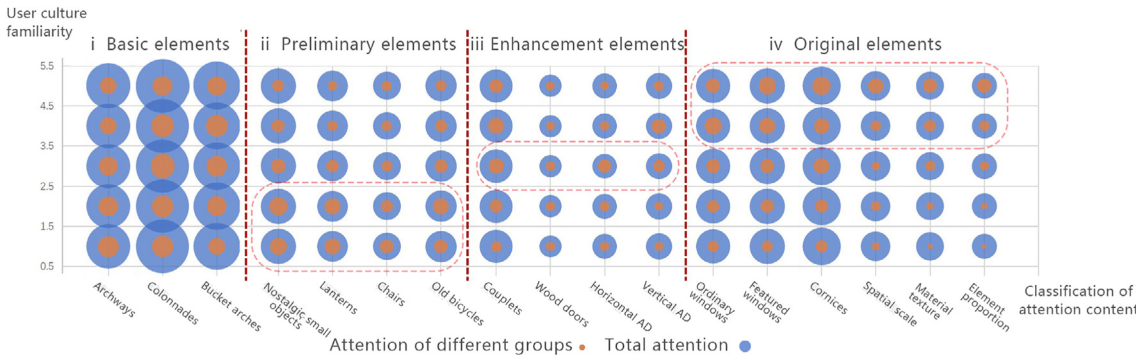

Fig. 7. Comprehensive analysis of the attention of the people with different cultural backgrounds on the architectural and cityscape features.

encers were eager to see the original cityscape features, and they were sensitive to constructing the style of cityscape, the spatial scale, the material texture and the local culture, and only a little bit of distortion would be questioned by them. This cityscape feature is greatly regional, original and representative, and is not only the cultural root of the street style, but also the source of attraction.

(2) Convergence protection control of different levels of cityscapes

In the cityscape protection of the ancient city of Zhangzhou, we classify the clusters of cityscape features based on the eye tracking data and spatial behavior data of the people with different levels of understanding for southern Fujian culture. Besides, we propose the dynamic proportions of different cityscape features for different levels of cityscape protection areas, in order to guide the design and control of cityscape protection (see Table 1).

Core area of cityscape protection. The core area of cityscape protection is located in the center of the ancient city, in order to let people feel the authentic style and charms of the ancient city's streets. The protection of the area focuses on the use of basic cityscape features and original cityscape features, ensuring that the adopted building scale, the cityscape of the components, the materials and the construction methods are consistent with the traditional architectures in southern Fujian. The key parts are designed and constructed carefully, and the architecture parts with high visual attention are recommended to be constructed by traditional craftsmen. It is also recommended to utilize some discarded traditional building materials and components to participate in the protection and repair of the core area. At the same time, we also need to excavate and upgrade the old-fashioned brands, ensuring that the core area is filled with the rich atmosphere of the ancient city of southern Fujian.

Buffer area of cityscape protection. The buffer area is an extension of the core area of cityscape protection, in order to harmonize the overall environment of the ancient city. The area focuses on the use of enhancement cityscape features, basic cityscape features and elementary cityscape features, and strict control of building volume is required in this area. On the basis of ensuring the continuation of the basic cityscape features, we attach great importance to the use of couplets, advertisements and cultural decorations to enhance the cultural atmosphere of the streets. Besides, we also use some nostalgic scenes and sketches to set off the nostalgic atmosphere of the ancient city, to ensure the continuity of the historical style and environment of the ancient city.

Coordinated development area of cityscape protection. The coordinated development area is the further expansion of the buffer area, in order to promote the coordinated development of the cityscape and economy of the surrounding urban areas of the ancient city. The protection of the area focuses on making most people feel the traditional atmosphere of the ancient cities of southern Fujian even in the outer area. It is advised to emphasize the creative use of basic cityscape features, and to use the visual images of features such as the colonnade, archway and wooden arch of the southern Fujian culture to let people realize that the southern Fujian culture has continued in the development of modern cities.

7. Discussion

The database coding structure of the BIM model is very suitable for the simulation analysis of eye tracking, which is of great help to promote the scientific evaluation and optimization of architectural design. In the architectural design process, architects use the BIM model for immersive VR experience and the visual simulation analysis of spatial eye movement to help architects make more scientific judgments on the perceived characteristics of architecture space and surface. Furthermore, the two-way transmission of BIM information enables the timely and scientific evaluation and optimization of design results.

The architectural design is closely related to the local climate. In different climate zones, different architectural forms are required to adapt to the local climatic conditions. VR technology can be used to simulate the climatic environment of different regions, analyze the real space perception of users, the changes of spatial eye movement and psychological feelings, which helps to create a more comfortable and more suitable architectural space and form.

Performing VR simulations on a large scale in the city helps to promote the interactive research between urban architecture and human settlements. In the large-scale spatial simulation of cities, VR technology is used to simulate different transportation, and the perception of people with different scales at different speeds, so as to better understand the interaction between people and urban environment, and further create more convenient and comfortable urban scales and living environment.

Table 1
Suggestions for the dynamic protection control of different levels of cityscape protection area.

	Basic cityscape features	Original cityscape features	Enhancement cityscape features	Elementary cityscape features
Core area	●●●●●	●●●●●	●●●	●●
Buffer area	●●●●	●●●	●●●●●	●●●
Coordinated development area	●●●	●	●●	●●●●

Degree of recommendation for cityscape features·

8. Conclusion

Meanwhile, the methods and technologies used in the experimental study can be extended to support the integrated design of urban space in the future.

The research on the protection of traditional cityscapes in southern Fujian culture has investigated the similarities and differences in the perceptions of cityscape features by different people in a comprehensive and systematic way, revealed the identification mechanism of cityscape features, and established a multi-cultural integrated planning and design method for cityscape protection. In this study, the efficiency and accuracy of conducting the research have been greatly improved due to the use of VR technology and eye tracking technology. The use of the combined technology can help architects quickly and scientifically find the weight relationship between the cityscape features of southern Fujian streets, and systematically understand the role of different cityscape features in protection planning and design, as well as the protection and control of the cityscapes in different levels of protection areas. This research has made important contributions to the humanization and quantitative research of cityscapes. Although the research scope and architectural style of this study are still relatively limited, the research methods and results have laid a good foundation for larger-scale research. Meanwhile, the methods and technologies used in the experimental study can be extended to support the integrated design of urban space in the future.

Declaration of Competing Interest

We argue that there is no conflict of interest.

Acknowledgement

This research is supported by the Young Scientists Fund of the National Natural Science Foundation of China (Grant No. 51808472). The authors would like to thank Prof. Sheng-Fen Chien in National Cheng Kung University for her assistance, suggestion, and discussions.

References

[1] J. Besharse, D. Bok, The Retina and its Disorders, Academic Press, 2011.
[2] L. Dupont, M. Antrop, V. Van Eetvelde, Eye-tracking analysis in cityscape perception research: Influence of photograph properties and cityscape characteristics, Cityscape Res. 39 (4) (2014) 417–432.
[3] T. Fukuda, M. Taguchi, Feature of Users' Eye Movements during a Distributed and Synchronised VR Meeting using Cloud Computing. eCAADe 2013, 18, 219, 2013.
[4] D.-I.D. Han, J. Weber, M. Bastiaansen, O. Mitas, X. Lub, Virtual and augmented reality technologies to enhance the visitor experience in cultural tourism *Augmented Reality and VR*, Springer, 2019, pp. 113–128.
[5] R.J. Jacob, K.S. Karn, Eye tracking in human-computer interaction and usability research: Ready to deliver the promises *The mind's eye*, Elsevier, 2003, pp. 573–605.
[6] P. Kiefer, I. Giannopoulos, M. Raubal, A. Duchowski, Eye tracking for spatial research: Cognition, computation, challenges, Spatial Cognition & Comput. 17 (1–2) (2017) 1–19.
[7] Chyi-Gang Kuo, Tay-Sheng Jeng, The research of augmented reality registration technology applying on architecture and urban outdoor guiding systems-using "cultural heritage guiding system" and "invisible shop signboard" as examples, Architectural J. 66 (2008) 145–166.
[8] Le-Min Zhang, Tay-Sheng Jeng, Ruo-Xi Zhang, Integration of VR, 3-D eye-tracking, and protocol analysis for re-designing street space, in: Paper presented at the CAADRIA 2018 – 23rd International Conference on Computer-Aided Architectural Design Research in Asia, 2018.
[9] H. Lefebvre, D. Nicholson-Smith, The Production of Space, Oxford Blackwell, 1991.
[10] K. Lynch, The Image of The City, MIT Press, 1960.
[11] A. Mantuano, S. Bernardi, F. Rupi, Cyclist gaze behavior in urban space: An eye-tracking experiment on the bicycle network of Bologna, Case Stud. Transport Policy 5 (2) (2017) 408–416.
[12] M. Meißner, J. Pfeiffer, T. Pfeiffer, H. Oppewal, Combining VR and mobile eye tracking to provide a naturalistic experimental environment for shopper research, J. Business Res. (2017).
[13] I. Mokas, L. Sebastien, N. Witters, R. Malina, Valuation of Urban Green Spaces using VR, 2019.
[14] A. Motamedi, Z. Wang, N. Yabuki, T. Fukuda, T. Michikawa, Signage visibility analysis and optimization system using BIM-enabled VR (VR) environments, Adv. Eng. Inf. 32 (2017) 248–262.
[15] H. Schrom-Feiertag, V. Settgast, S. Seer, Evaluation of indoor guidance systems using eye tracking in an immersive virtual environment, Spatial Cognition Comput. 17 (1–2) (2017) 163–183.
[16] S. Schwarzkopf, R. von Stülpnagel, S.J. Büchner, L. Konieczny, G. Kallert, C. Hölscher, What lab eye tracking tells us about wayfinding a comparison of stationary and mobile eye tracking in a large building scenario, in: Paper presented at the Eye Tracking for Spatial Research, Proceedings of the 1st International Workshop (in conjunction with COSIT 2013), 2013.
[17] Chang Qing, Several issues in the protection and development of terroir buildings, Times Architecture 3 (2000) 25–27.
[18] Guo Suling, Zhao Ningwei, Zhang Jianxin, Xue Ting, Xu Liu Peixue, Shuai, Xu Dandan, Visual quality assessment of cityscape based on eye movement——taking college students' eye movement experiments on Hongcun tourism cityscape pictures as an example, Resour. Sci. 39 (6) (2017) 1137–1147.
[19] Shao-Sen Wang, Ya-Min Zhao, Exploration on the organic renewal of traditional residential buildings in Luzhou Ancient City in Southern Anhui, Southern Architecture 176 (06) (2016) 75.
[20] Zhang Jie, Lu Jie, From large-scale urban design to "space of daily life", Urban Planning 27 (9) (2003) 40–45.
[21] Le-Min Zhang, Ruo-Xi Zhang, Analysis of the features of "homesickness" in the organic renewal of southern Anhui historical scenery district——taking the protection and renovation design of the ancient city of Zhangzhou as an example, Urban Architecture 20 (2015) 12–13.