

Article

Research on the Relationship Between Urban Visual Quality and Students' Emotional Experience Driven by Multimodal Data: A Case Study of Beijing Olympic Park Museum Group

Ruoshi Zhang ^{*}, Dingnan Chai, Zhenzhi Zhou, Rui Sun, Zekai Zhang and Chuhuan Chen

Department of Architecture, School of Landscape Architecture, Beijing Forestry University, Beijing 100083, China

* Correspondence: zhang_rs@bjfu.edu.cn

Abstract: The technological leap in the digital age has triggered a higher demand for emotional experiences. Since vision has long been recognized as the most important source of information for people to perceive the city, the visual characteristics of the urban built environment have an important impact on the emotional experience it creates. This study takes the facades and external environments of the representative urban public building category, museums, as the research objects, to explore how different types of built environment elements that make up the museum's facade and surrounding environment affect the people's emotional experience by affecting the urban visual quality. Tobii Pro Fusion devices and emotion scales are used for the collection of multimodal data. A total of 298 participants aged between 20 and 25 were recruited to participate in the experiment voluntarily and 229 valid data were finally obtained. The results show that different types of museum facades and external environmental elements have different effects on visual quality and people's emotional experience: (1) Interactive elements show the most significant impact on visual quality and emotional experience. (2) The impact of artificial elements on visual quality is related to their size and spatial composition on the museum facade. When artificial elements have uniqueness, regionality, and symbolism associated with the museum, they are more likely to trigger people's emotional experiences. (3) Visual quality is significantly correlated with people's emotional experience through both preconscious and conscious emotional cognition. The results reveal the impact of different types and characteristics of built environment elements on the visual quality of the urban environment and proves the important role of visual quality in students' emotional experience, further illustrating that designing and creating a good visual environment is conducive to promoting a humanistic urban environment. Meanwhile, this study also promotes a method of urban visual quality research and evaluation from an emotional perspective, providing a more scientific and objective way for the intervention of "emotions" in built environment research.



Received: 2 November 2024

Revised: 28 November 2024

Accepted: 25 December 2024

Published: 27 December 2024

Citation: Zhang, R.; Chai, D.; Zhou, Z.; Sun, R.; Zhang, Z.; Chen, C. Research on the Relationship Between Urban Visual Quality and Students' Emotional Experience Driven by Multimodal Data: A Case Study of Beijing Olympic Park Museum Group. *Buildings* **2025**, *15*, 57. <https://doi.org/10.3390/buildings15010057>

Copyright: © 2024 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: urban visual quality; emotional experience; museum facades; eye tracking; emotional scale; multimodal data

1. Introduction

With ever-changing digital technologies deepening and expanding in people's daily life, digital living has become a common state worldwide. Though high technologies have brought convenience to people in all aspects, long-term immersion in a virtual environment makes them increasingly demand for more emotional experiences in real urban environment [1,2]. As urban public buildings and their surrounding environment

consist of important nodes of urban environment, they have become the main space for people to carry out public activities, and naturally become an important material basis for people to gain emotional experiences and establish attachment to urban built environment [3–5]. Meanwhile, previous studies have shown that vision accounts for more than 80% of the information when people perceive the surrounding environment [6], so a high-quality urban visual environment has become an effective way to enhance people's emotional attachment to the city [7,8]. As an important carrier of a city's history and culture, urban museums make an important contribution to its surrounding urban visual environment [9,10]. It is worthwhile to study and evaluate the visual quality of museum facades and their external environment as well as their emotional influence on people.

In recent years, under the guidance of Beijing's planning to build a "museum city" from 2023 to 2035, urban designers, architects, researchers, as well as governors are concentrating more on the design and research of urban museums [11–13]. Many museums with unique appearances have gradually become landmark in the area, influencing and driving the development of the surrounding urban environment. It is precisely because these public buildings with unique cultural characteristics pay special attention to their appearance in design that they become important factors affecting the urban visual environment. As the central axis of Beijing has gradually become an important world cultural heritage, the Olympic Park museums in its core area have also become typical representatives of Beijing museums, making them an appropriate research object for this study. However, existing research rarely focuses on how their built environment characteristics affect urban visual quality, let alone related explorations from the user's emotional perspective. On the one hand, this may be due to the pluralistic and complex characteristics of emotion itself [14,15], and on the other hand, it is difficult to comprehensively and objectively describe these characteristics using traditional qualitative methods such as observation and interview [16,17]. Emotions are the basis of human behavior, and emotional perception plays a key role in the physical and mental health of individuals [18]. Visual perception of urban public space is an important way for people to obtain emotional attachment to urban environment, so quantitatively exploring the relationship between urban visual quality and people's emotional perception is of great importance. The maturity of ergonomics technology and emotion measurement methods promoted by multiple disciplines provides an opportunity to solve this problem. For the former, various ergonomic equipment that can collect and respond to human multi-dimensional perception in real time has provided a way to explore people's eye movement and reveal people's possible emotional experience in the process to a certain extent more objectively and scientifically [19–22]. For the latter, emotion measurement methods promoted by multiple disciplines to quantitatively measure the dimension, degree and intensity of specific influencing factors of people's emotion have been widely used in various emotion-related studies [14,16–18,23]. Although emotion is complex and its definition is not unified in different disciplines, relevant scholars believe that it is still possible to summarize the universal and regular emotional models [16,24–26]. In the field of built environment design and research in the past several years, the effectiveness of the above method has been confirmed in multiple studies on small-scale buildings, medium-scale urban public spaces, and natural and urban landscapes [27–31].

The purpose of this study is to explore how different types of built environment elements that make up the museum's facade and surrounding environment affect the viewer's emotional experience by affecting the urban visual quality. This study selected three representative urban museum built in past two years in Beijing central axis museum group to explore the urban visual environment they create and the emotional experience they bring to people, including The Chinese Traditional Culture Museum, the Museum of the Communist Party of China, and the China Science and Technology Museum. Due to the

limitations of experimental space and access control of colleges in the post-pandemic era, the participants finally recruited in this study were college students aged 20–25. In addition, previous eye tracking studies on visual preferences of landscape space showed that whether the participants studied spatial design-related majors would affect their attention to spatial elements [31–33]. Therefore, the major of the participants recruited in this study were also included in the latent variables of individual differences. Based on the research purpose and combined with the specific characteristics of the three selected museums, the built environment elements are divided into three categories, artificial elements, natural elements, and interactive elements, and the following research hypotheses are proposed:

- (1) Participants have different visual behaviors and preference on three different kinds of elements consisting of the museum facades and surrounding environment, which can be shown in both AOI heat map and eye-tracking indicators.
- (2) Looking at the facade and surrounding environments of the three museums composed of different elements will affect students' emotional experience, and different characteristics of elements may have different effects on both emotion state and tendency, which can be shown in the related emotional scales.
- (3) Different visual elements and their different characteristics may have different effects on participants' emotional experiences, and these effects may also be related to individual differences. Participants majoring in spatial design may exhibit different visual preferences and emotional experiences than others.

2. Literature Review

2.1. Research on the Influence of Urban Visual Quality on People's Emotional Experience

The visual quality of the built environment has long been concerned by architects, urban planners, and landscape architects intentionally or unintentionally considering the importance of vision for people perceiving surrounding environment [7,34–42]. Among them, the visual quality of the built environment is considered to directly determine people's aesthetic experience and has gradually become a recognized index for the aesthetic evaluation of the built environment [36,37,41,42]. Based on this, a series of design theories and methodologies originated from visual quality have been developed by both designers and researchers. For example, Kevin Lynch's classic study of the image of a city aims to explore the relationship between objective urban form and people's subjective place identity, and the five elements of good city form proposed emphasize the value of visual quality [43,44]. Meanwhile, the core of some postmodern architectural theorists' critique of modernist cities and architectural homogeneity such as Rem Koolhaas' *Delirious New York* also mentions its lack of visual appeal and even causes people's negative feelings [45]. With the intervention of interdisciplinary theories such as psychology and sociology in the field of built environment research, the aesthetic effect of the environment based on visual perception has a more scientific and objective theoretical basis for the emotional impact on people [26,42,46,47]. Since the early 20th century, psychologists have been using mathematical approaches and developing empirical research to advance the understanding of the influence of environmental aesthetics on people's emotions [48–50], and during the mid-20th century, these methods were widely welcomed and used in both environmental design and research [51,52].

Until now, it has been widely recognized that the visual quality of an urban environment greatly affects people's emotional experience. Among them, the attention and research on the natural elements and their characteristics of urban environment are the most in-depth and extensive [51–54]. Natural settings such as parks with trees, water, and vegetation are proved to be of great influence on people's emotion from various aspects, including emotional healing research [53–55], positive emotional triggering research [56,57],

etc. The effectiveness of using photographs as stimulus materials for visual-based emotional perception studies has also been widely confirmed by these studies [58–61]. For example, Ulrich found that looking at photographs of natural scenes had fewer negative emotions and more positive emotions than looking at urban scenes without trees [62]. Herzog et al. asked participants to evaluate photographs of different urban or forest environments and chose where they liked to rest most after mental exhaustion [63]. Meyer-Grandbastien and colleagues demonstrated the psychological recovery of visitors within urban green spaces using natural pictures as stimulus [64]. However, there are few studies focusing on the correlation between the visual quality of the built environment features of urban spaces and public buildings, especially from the perspective of people's emotional experience [31,65]. It can even be said that existing research results have not surpassed the attention of the abovementioned 20th century architects to the impact of building facade form and composition on people's perception and emotion, which contribute a lot to the visual quality of the urban built environment [66–68]. Therefore, this study takes museums, an important type of public building that constitutes the city facade, as an example to study the impact and internal connection of its facade and external environment on people's visual perception and emotional experience, thereby exploring the possible influencing factors of urban visual quality from an emotional perspective and supplementing the deficiencies of existing research. The abovementioned mature method of studying the relationship between visual quality and people's emotion of natural environment based on pictures provides a methodological reference for this study.

2.2. The Definition and Application of Eye Tracking in the Study of Urban Visual Quality and Its Emotional Traits

Since the visual system is the most important channel for human beings to obtain external information, eye movement is considered to be the most effective means in the study of visual information processing [6,69,70]. As far back as the Middle Ages, people began to use instruments to make simple observations and experiments on eye movement [71]. In 1901 Dodge and Cline developed the first accurate, non-forced eye movement tracking device [72]. The concept of eye tracking was first developed in 1930s, and was used in the fields of psychology, aerospace-related dynamic analysis, and commercial advertising in early days [73]. Later, it gradually entered the research on human perception. For example, it had been used in the study of the visual perception development of children [74], the visual attention allocation of aircraft traffic monitoring and avoidance [75], and the influence of auditory information on the audience's visual cognition during the viewing of TV commercials [76]. At the beginning of the 21st century, the development of optical sensor and the improvement of computer technology to process complex image information in real time have promoted the rapid development of eye tracking technology, which is then widely used in multi-field research to explore people's perception, cognition, and behavior related to emotion [59,71,77–82]. Most of the visual stimuli used in these research studies were static two-dimensional images, and their effectiveness has therefore been widely confirmed.

Specifically, eye tracking devices track eye movements by measuring the position of the fixation point or the movement of the eye relative to the head. The indicators that characterize eye movements are divided into two categories: time and space. The former includes fixation time, saccade time, saccade latency, follow-up movement time, and retrospective time, while the latter includes fixation point, number of eye movements, eye movement trajectory, heat map, etc. [80]. Changes in pupil diameter can also be recorded and measured simultaneously. As these indicators can quantify the visual attention and cognitive process of the human eye, researchers can study people's psychological process by recording and analyzing eye movement data to obtain relatively reliable results. It is more

objective than the traditional built environment research methods and has been widely used in the study of natural environment relating to perception, restoration, and aesthetic evaluation in recent years [51,62,63,77,83–85]. Among them, the most common use of eye tracking is to assess the restorative components of urban parks, such as cognitive load preference for different vegetation, water space and plants, and eye movement behavior in natural scenarios with different degrees of recovery [85,86]. There are also some studies based on the natural environment that combine eye tracking data with other modal data to explore the characteristics of human perception of the environment more thoroughly. For example, research on landscape aesthetic effect, visual quality, and soundscape space evaluation combining subjective questionnaires and objective eye movement data [83], or tourism map design research using both wayfinding and eye tracking experiments to mining users' perception needs for the spatial sequence of the park [85]. These studies provide a mature reference for the use of eye tracking technology in environmental perception and environmental evaluation. In addition to the natural environment, eye tracking technology has also been used to study more complicated urban environments to evaluate their visual quality in recent years. For example, several studies explore the differences in visual appeal of different street components with the help of the Tobii eye tracking platform. The results show that people usually notice the infrastructure such as buildings, roads, and sidewalks in the streetscape first, then the cultural landscape of the streetscape such as shops, billboards, and parks, and finally the people and animals [87]. Some studies use eye tracking to assess the urban visual safety and visual pollution of different kinds of street facades [88]. Meanwhile, some studies combine eye tracking and subjective questionnaires to evaluate the aesthetic quality of a city skyline [89]. To a certain extent, these studies have shown that different artificial environments can affect urban visual quality and people's emotional perception. However, the application of eye tracking technology in the study of visual quality of urban artificial environment is still less than that of the natural environment, especially the part that focuses on emotional experience. Moreover, most of the existing studies on the urban artificial environment focus on large-scale urban areas or streets, and there is a lack of discussion on the characteristics of single building facades and smaller-scale building components. At the same time, emotion, as the highest human need, is an important part of the aesthetic effect of evaluating the visual quality of urban environment, and relevant quantitative research is urgently needed.

Meanwhile, since different eye-tracking indicators are thought to reveal people's emotional tendencies to the environment and the degree of attraction to the environment, and neuroscience studies have shown that emotional arousal causes pupil changes by activating the autonomic nervous system [82], eye tracking has become one of the important modalities for quantifying people's emotional perception. Some researchers have coupled eye tracking data with questionnaires, scales, and other physiological data according to different research questions to more comprehensively explore people's emotional perception of the environment [31,80,83,85,87,88]. In this study, Tobii Pro Fusion and Tobii Pro Lab were used to collect and analyze people's eye movement and its emotional traits when viewing the three museum facades, and the PAD scale was used to simultaneously collect the intensity and dimension of people's emotional perception. Coupled analysis of these two modalities was conducted to explore the relationship between urban visual quality and emotional perception. Hopefully, the results will provide useful enlightenment for museum facade design and further urban visual quality improvement.

3. Materials and Methods

With the help of Tobii Pro Fusion eye tracking devices and emotional scales, multi-modal data that can reflect people's visual preferences and emotional experience character-

istics of the three museum facades were collected. Tobii Pro Lab was used to record and process eye tracking data and generate series of visual AOI heat maps based on the AOI drawn according to the research purpose. SPSS was used to perform descriptive statistics and difference analysis on the data collected from the emotional scales and eye-tracking experiment, as well as correlation analysis between specific environmental characteristics on the emotional scale. By comprehensively analyzing the above eye tracking data and emotional scale data, the relationship between the visual quality of the museum facade and surrounding environment and people's emotional experience can be further explored. It is hoped that the results can improve the understanding of the relationship between visual preference and human emotional experience from a theoretical perspective, thereby providing a reference for built environment design and evaluation methods for improving urban visual quality concerning people's emotion and truly promoting the creation of a humanized urban environment.

3.1. Research Objects: Three Landmark Museums in Beijing Olympic Park Museum Group

Located in Chaoyang District, Beijing, the Beijing Olympic Park Museum Group along Beijing central axis is an important area for Beijing to promote the construction of a "City of Museums", reconstructing the northern end of Beijing's central axis. The museum group consists of three distinct museum buildings: the Chinese Traditional Culture Museum, the Museum of the Communist Party of China, and the China Science and Technology Museum (Figure 1). The three museums selected in this study form the Beijing Olympic Park Museum Group in terms of location, so viewing the pictures of the three museums at the same time is pasted into the real city visual experience. At the same time, the specific characteristics of the three museums are different, so it is possible to better study the different effects of real environmental elements on the visual quality and human emotional experience.

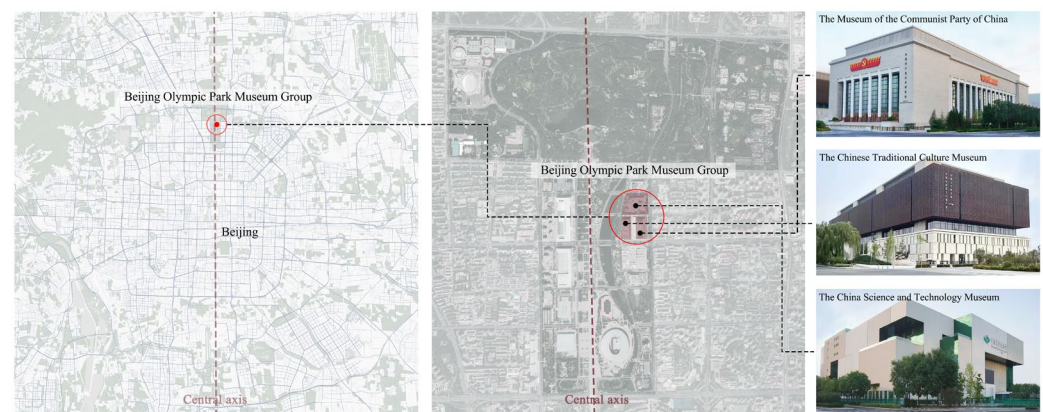


Figure 1. Location of three museums taken as research objects.

The Chinese Traditional Culture Museum was designed by the architects von Gerkan, Marg, and Partners (GMP) and opened to public in 2022. It is designed for the display of China's cultural heritage and is believed to be an important component of the new cultural quartier in the Olympic Green in Beijing. In line with the theme of the museum, the architectural design of the museum is inspired by traditional Chinese cultural elements in terms of structure, materials, and form. The building comprises three distinct dimensionally and structurally coordinated parts: the plinth built with light-colored Xiu Shi granite, the glazed intermediate area, and the copper-colored main body floating above. The structural pattern employed in the curtain wall facade of the main body relates to the patterns and structures found in Chinese stained windows, which is woven from aluminum profiles

with a matte copper coating. The light and shadow space created by the curtain wall in the interior further reflects the translation of the ancient Chinese architectural window and the atmosphere of the Chinese interior space.

The Museum of the Communist Party of China (CPC) is a museum specially devoted to permanent and comprehensive exhibitions of the Party's history. It was officially completed on 5 May 2021. Its architectural design responds to the guiding spirit of General Secretary Xi Jinping, to create a sacred hall, a spiritual home for Party members, an educational center to guide the public to follow the Party's instructions, and an authoritative institute for Party history research and publicity. Therefore, compared with the other two museums in this study, the museum building of the CPC is full of historic connotations. For the overall design of the space, the museum imitates the Chinese character form of "gong", which looks like a horizontal H. It means worker or labor in Chinese to imply that the Chinese Communist Party is the vanguard of the Chinese working class, Chinese people, and the Chinese nation. For the structural design, the museum adopts the traditional colonnade structure to reflect the essence of Chinese architecture: upright and square, simple and grand. The facade is paved in a traditional way with off-white masonry.

The China Science and Technology Museum is the only comprehensive museum of science and technology at national level in China. It was initially built in 2006 and opened to the public in 2009. As a museum adjacent to the venue area of the Beijing 2008 Olympic Games and built at the same time, its design was not only expected to reflect the characteristics of a museum of science and technology but should also follow the Olympic motto "Green Olympics, Scientific Olympics and Humanity Olympics". Therefore, composed of a slew of building block-like concrete blocks that articulate with each other, the structure of China Science and Technology Museum is featured like a Lu Ban lock, or a magic cube, symbolizing that science is an endless process of "unlocking" and "discovering secrets". The materials used were more modern than the other two museums in the group, including stainless steel cladding panels with different textures and green glass curtain walls.

It could be said that the shapes and facades of the three museums selected in this study have their specific external characteristics, and they all consciously use different symbolic techniques to express the characteristics of the museums in the design and construction, trying to attract people's attention at first glance. In this way, they provide a suitable case study to study the relationship between the urban visual quality created by museum facades and human emotions.

Photographs of the facades of three museums were used as experimental materials in this study. The photographs of the three museums were all taken on 5 June 2023 from 5 a.m. to 7 a.m. using a Sony ZVE10 digital camera, and all photos are of the same resolution of 2560×1440 pixels and show the main façade of each museum, with the proportions of sky, land, and buildings as consistent as possible. To keep people's attention on the facade, Adobe Photoshop 2019 was then used to manipulate the three facade photos, erasing traffic lights, pedestrians, and adjusting the sky conditioner to the same color tone. The experimental stimulus materials therefore consist of three processed photos, with Sample 1 being the Museum of the Communist Party of China, Sample 2 being the Chinese Traditional Culture Museum, and Sample 3 being the China Science and Technology Museum. In the three samples, the building occupies the main part, and the stimulus material contains the visual elements of the building and the visual elements of the surrounding environment. By analyzing the H (hue), B (brightness), and S (saturation) of the processed images, the natural elements of the three samples are similar in these three indexes, which ensures that the impact of ambient differences on the experiment is minimized.

3.2. Participants

Since the laboratory where the data are collected is located in Beijing Forestry University, off-campus participants need to register and make an appointment in advance, which affects the recruitment process of diverse participants to a certain extent. In addition, due to the limited time available for eye tracking devices, in order to ensure the smooth progress of the experiment, participants were recruited from the school. A total of 298 participants were recruited to participate in the experiment voluntarily, including 249 females and 49 males, aged between 20 and 25. All the participants had normal vision and were not informed concerning the purpose of this study. Information on the level of education and major was obtained from participants, as previous studies have shown that people who study environmental design related majors (urban planning, landscape architecture) may have a different visual focus on the pictures of the built environment than others. Informed consent was provided to the participants prior to the experiment. Participants were informed of the experimental procedures and all agreed to provide their eye tracking data for further analysis. Participants were allowed to stop the experiment at any time. Meanwhile, according to previous eye tracking experiment, the numerical composition of the group was sufficient, and the results could be considered representative. Combined with the criteria of Tobii Pro Fusion and the emotion scale to determine whether the collected data were qualified, 229 valid data were finally obtained by removing the test data with an eye tracking sampling rate of less than 80%, unqualified eye movement calibration, and blank abnormalities in the eye movement and scale. The final number of participants with valid data used for analysis was therefore 229, with 193 females and 36 males. The specific distribution of genders and majors is shown in Table 1.

Table 1. Participants' distribution according to gender and major difference.

	Environmental Design Related Majors (Urban Planning, Landscape Architecture)		Other Majors (Economics, Chemistry, Mathematics, Forestry)	
	Original Participants	Participants with Valid Data	Original Participants	Participants with Valid Data
Male	22	16	27	20
Female	115	89	134	104

3.3. Experimental Protocol

The experiment was conducted in a laboratory specially arranged for eye-tracking at Beijing Forestry University from August to September 2023. The experimental area is surrounded by five white baffles that were 0.8 m wide and 1.8 m high, creating a relatively private and relaxing experimental environment for the participants. There was a height-adjustable chair, and a computer monitor equipped with a Tobii Pro Fusion eye tracker placed on the table in the area (Figure 2). The chair allowed for participants of different heights to flexibly adjust their sitting height, ensuring that the eye tracker could accurately capture eye movement behavior.

For the test, the Tobii Pro Fusion (250 Hz) eye tracker attached to a 21.5-inch monitor with a resolution of 1920×1080 pixels was used to record the participants' eye movement data. The eye tracker used in this study was a Tobii Pro Fusion (Tobii AB, Danderyd, Sweden) with dimensions of $374 \text{ mm} \times 18 \text{ mm} \times 13.7 \text{ mm}$ and a capture speed of 250 Hz. The eye tracker was mounted on the underside of the laptop screen and linked to the computer via USB. Tobii Pro Lab (Tobii AB, Danderyd, Sweden, version 1.207) software was also installed on the computer for device management, data calibration, and subsequent data analysis. The specific experimental process is as follows (Figure 3).

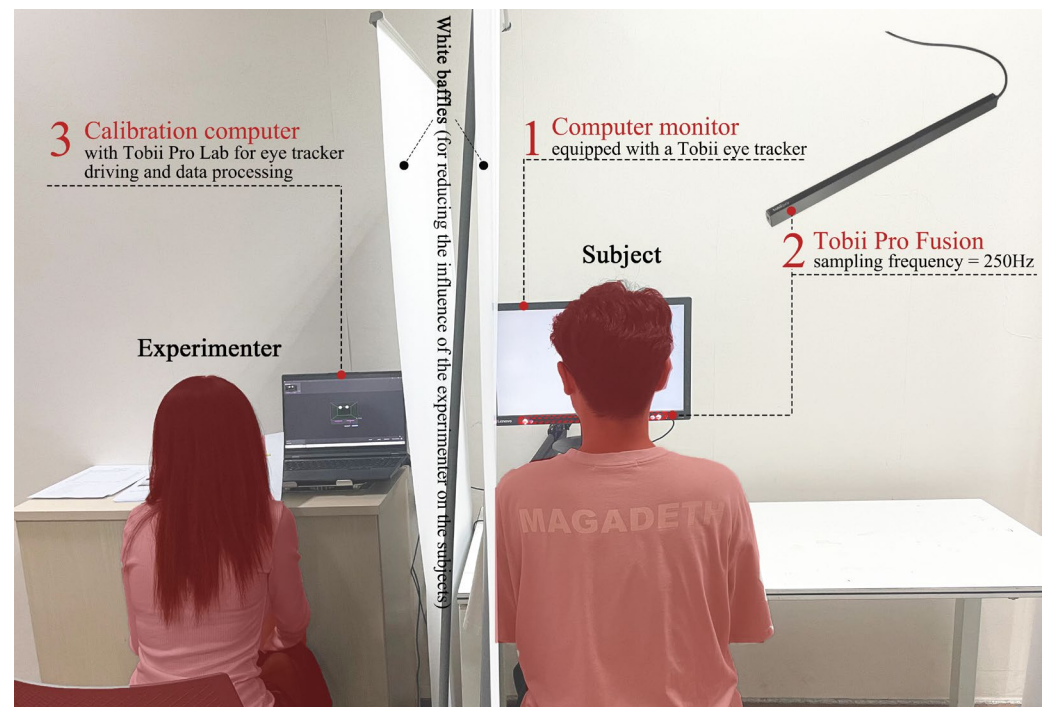


Figure 2. Experimental environment and instruments.

Preparation Phase				Experiment Phase								
Participants adapt to the experimental environment and sign the consent form.	Participants sit 60–65cm in front of the computer screen.	Eye tracking calibration.	Participants familiar with the picture switching method in free time through practice of the trials.	Photos of the three museum facade are played in random order.	Photo 1 Free time	Blank 2s	Photo 2 Free time	Blank 2s	Photo 3 Free time	Blank 2s	End 1s	Participants fill out the emotional scales after viewing the photos.
5 mins				3–5 mins								
				5 mins								

Figure 3. Experimental process.

After signing the informed consent form, 298 participants underwent the eye tracking experiment and emotional scale measurement one by one. Participants were guided to sit in the arranged enclosed experimental area. An examiner sat outside the area on the participant's left hand side to give participants necessary guidance before and after the test. At the same time, the examiner observed the process of eye tracking through the second screen to prevent various complications in the data collection process.

The protocol began with the eye tracking experiment. After the confirmation of participants' personal information, participants were guided to adjust the height of the chair and their sitting position to ensure the distance between their eyes and the screen was about 60 cm. Following the guidance of the examiner, the eye tracker calibration was performed for each participant. The calibration program selected in Tobii Pro Lab was a five-point calibration mode with the best trade-off between data quality and time cost, until the calibration results "Calibration accuracy" and "Calibration precision (SD)" were less than 0.4. If one of the two eyes was not detected by the eye tracker, the participant was excluded. Therefore, the participant was asked to look at the points indicated by the examiner during the calibration, as detecting each of the eyes during the whole test is of vital importance to guarantee the validity of data.

The test was performed under normal lighting conditions after the calibration. In order to make the experiment clear to the participants, three grayscale pictures that do not

cause strong mood swings were selected before the formal experiment and participants were asked to simulate the operation of switching images of the experiment (Figure 4). The instructions for both practice images and formal experiment images were written in white background images, and Tobii Pro Lab was used to arrange them in the right place to guide participants to conduct simulation operation and subsequent formal experiment. In addition, to reduce the impact of additional variables, a fixation point was presented after the instruction image and before each stimulus image, so that the center of the eye movement behavior of each participant was in the same fixed position (the center of the screen) when the participant started looking at the stimulus. To ensure this, the “advance on” in Tobii Pro Lab was modified to 1 s so that the screen would automatically jump from the fixation to the experimental stimulus image when the participant fixated continuously on the central fixation for 1 s.

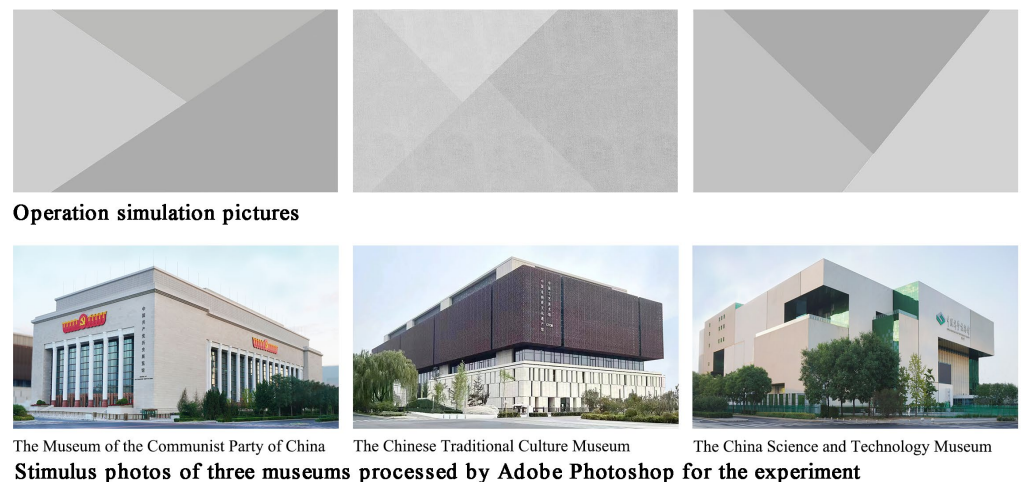


Figure 4. Operation simulation pictures and stimulus photos used in the experiment.

When participants felt they had a complete understanding of how to conduct a formal experiment, they were asked to move on to the formal experiment. Once the experiment began, the aforementioned three processed images of the museum facades were randomly played as stimulus materials (Figure 4). Participants were asked to switch to the next image only after they felt they had carefully observed and understood the exterior features of the three museums depicted in the three images, including the forms, materials, colors, structures, and surroundings of the museum building. Participants were guided to conclude the eye tracking experiment with a closing sentence written on a white background. The data were recorded objectively by the eye tracker.

After the eye tracking experiment, participants were asked to fill in emotional scales consisting of the PAD emotional scale and emotional attachment scale of specific environmental characteristics based on the feelings of the three pictures they had just observed.

3.4. Data Collection and Analysis

3.4.1. Eye Tracking Data Collection and Analysis

The eye tracking data collection and analysis process were performed in Tobii Pro Lab and consist of three parts: AOI drawing according to the research question, real-time recording and export of eye tracking data, and analysis of eye tracking data.

As the research aims to explore the visual quality promoted by building facades consisting of different kinds of elements and their influence on people’s emotion, combined with the characteristics of the stimulus material, the elements contained in the image were divided into three categories to illustrate the areas of interest: artificial elements, natural elements, and interactive elements. Artificial elements included architectural visual

elements (e.g., wall without texture, wall with special texture, glass), and street environment visual elements (e.g., roads, fences). Natural elements included vegetation and sky, and interactive elements included text and logo. Based on this, the AOI of the three images are shown in Figure 5. The heat map generated by Tobii Pro Lab could then be combined with this AOI partition to analyze the characteristics of specific eye tracking metrics on different elements.

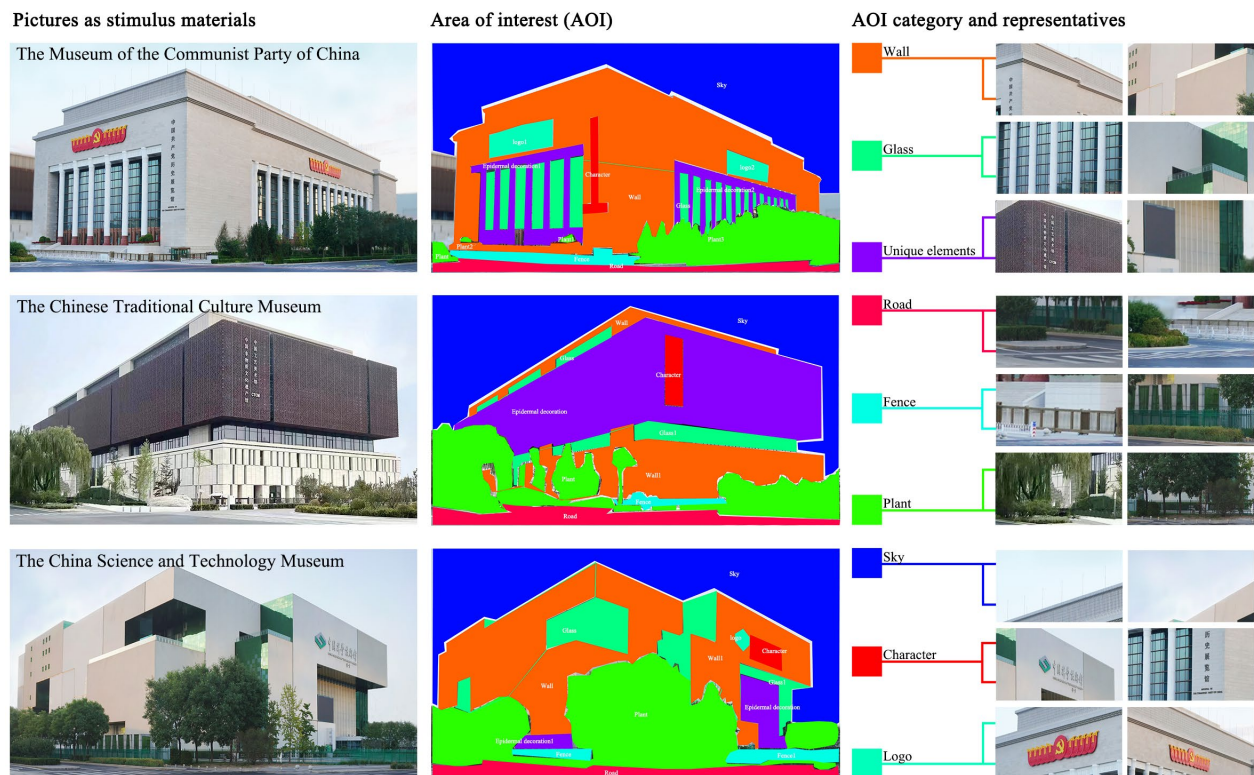


Figure 5. AOI and related representatives.

Once the experiment had been performed on all participants, all recordings could be reviewed in Tobii Pro Lab. Tobii Pro Lab recorded the overall characteristics of eye movements, such as the duration and trajectory of each participant's eye movements when they viewed three images of the museum's exterior, as well as specific eye movement indicators selected especially for exploring the relationship between the visual quality created by the chosen city museum and people's emotion. To achieve this, this study selected four eye tracking indicators that are thought to be able to reflect or reveal a person's emotional experience. Its connotation and its relationship with emotions are shown in the Table 2. At the same time, the software automatically records the sampling rate of each participant's eye movement. In this study, data with a sampling rate greater than or equal to 80% were selected for follow-up analysis to ensure the reliability of the data.

Then, the data for each participant were automatically exported from Tobii Pro Lab to Excel (Microsoft® Office Excel 2011, Microsoft Corporation, Redmond, WA, USA). The data with a sampling rate of less than 80%, the calibration failure data, and the blank anomaly data were cleared, and finally 229 valid data were obtained. Social science statistical software packages (SPSS 20, SPSS Inc., IBM Corporation, Somers, NY, USA) were used to analyze the overall characteristics, specific differences, and correlation analysis results of the data. The mean and standard deviation of the four eye tracking indicators were used to describe the participants' overall eye movement feedback to different environmental elements. As the results of the Kolmogorov–Smirnov test showed that none of the four eye tracking indicators were normally distributed, Mann–Whitney U was selected for the

nonparametric test to assess whether there were differences in the eye movement indicators of different participants to different environmental elements, as well as whether there was any influence from the participants' major relating to environment design or not. A p -value < 0.05 is considered statistically significant. Spearman correlation analysis was used to analyze the correlation between different factors, different study subjects, and eye tracking indicators.

Table 2. Eye tracking metrics chosen in this study.

Abbreviation	Unit	Full Name and Explanation
TFD	s	Total duration of fixations; total duration of fixations inside the AOI/picture; reflects cognitive attention
NF	n	Number of fixations; times of participants fixation on the AOI; reflects visual information processing
FFD	s	Duration of first fixation; the duration of the initial gaze at this AOI; reflects visual initial processing
TFF	s	Time to first fixation; time taken to perceive the AOI (area of interest); reflects speed of attraction

3.4.2. Application and Analysis of Emotional Scale

As the generation of human emotion is a complex cognitive process that is difficult to generalize simply, the emotion model is mostly divided into discrete emotion model and dimensional emotion model [24]. The discrete emotion model divides emotions into several independent elements while the dimensional emotion model pays more attention to the interaction and association between the above elements, which is more suitable for this study to explore the characteristics of interaction between people and the built environment. Among the dimensional emotion models, the PAD model, consisting of three dimensions, pleasure, arousal, and dominance, has proven effective in measuring and interpreting a person's emotional state. It is suitable for different cultural backgrounds, and has the advantages of being fast and intuitive, as well as having good fault tolerance and a wide range of applicable types, which meets the research purpose. Based on this model, Mehrabian et al. developed the PAD Scale to measure emotional values [90]. Subsequently, the Institute of Psychology of the Chinese Academy of Sciences has summarized a set of Chinese simplified PAD scale suitable for Chinese environments, which have been proven to have good structural validity and applicability and are widely used in many fields and disciplines in China [91]. In addition, in order to further explore the influence of specific characteristics in the artificial, natural, and interactive elements contained in the research object on people's emotional experience, this study also constructed a scale of emotional experience degree of specific environmental characteristics.

- PAD scale

This study used the PAD scale (Chinese version) to quantitatively evaluate the impact of visual quality on people's emotional perception on the facade of urban museums and correlated it with the results of the aforementioned eye tracking data. At present, the application of PAD scale in the fields of architecture, urban design, and landscape architecture is relatively rare, which is closely related to the lack of systematic research on people's emotional perception in the study of the built environment. However, the value of the PAD scale has been widely demonstrated in the adjacent field of multidisciplinary emotional design research [92–94]. Therefore, this study also hopes to verify the effectiveness as well as shortcomings of the PAD scale in quantifying people's emotional experience in the study of the built environment and promote the development of methods for the study of

emotional attachment between people and the built environment. The specific application of the PAD scale is as follows:

As the participants recruited were all Chinese, the PAD scale used in the experiment was a version of the simplified PAD scale suitable for the Chinese context, which was summarized by the Institute of Psychology of the Chinese Academy of Sciences. The Cronbach's α of P , A , and D are 0.95, 0.83, and 0.78, respectively, according to the aforementioned research, which validates its reliability. Developed from the PAD emotion model, the scale included three dimensions, pleasure–displeasure, arousal–non-arousal, and dominance–submissiveness, each of which had 4 items and was a 9-point semantic difference scale. Each item was composed of pairs of adjectives that represented different emotional states, and the emotions expressed by each pair of words were opposite in their dimension, while the other two dimensions were basically the same. Participants were asked to rate their emotions according to which emotion was stronger and how intense it was. The score on the item from left-to-right is scored as “−4” to “4”, and when marked in the middle, it was scored as “0”. The final dimension score was the average of the 4 item scores that measure that dimension. Since there are half of the inverted items (negative descriptions) in each dimension, the score for each dimension is calculated as follows:

$$\begin{aligned} P &= (P1 - P2 + P3 - P4) / 4 \\ A &= (-A1 + A2 - A3 + A4) / 4 \\ D &= (D1 - D2 + D3 - D4) / 4 \end{aligned} \quad (1)$$

After further experimental research, the Institute of Psychology of the Chinese Academy of Sciences obtained a reference table of 14 basic emotional PAD values (Table 3). With the basic affective PAD value, it is possible to assess the emotional tendency and the degree of inclination of the person.

Table 3. PAD scale.

Serial Number	Emotion	−4	−3	−2	−1	0	1	2	3	4	Emotion
P1	Angry										Interested
P2	Pejorative										Friendly
P3	Painful										Delighted
P4	Infuriated										Excited
A1	Sleepy										Sober
A2	Calm										Excited
A3	Relaxing										Interested
A4	Restrained										Surprised
D1	Controlled										Mastered
D2	Compliant										Dominative
D3	Humility										Proud
D4	Affected										Influential

Specifically, PAD emotional tendency referred to the degree of sparseness between the measured emotional state and the 14 basic emotions, which could be spatially expressed as the distance relationship between the measured emotional state and the spatial coordinates of the 14 basic emotional states. The coordinate distance of the emotional space can be calculated by the Euclidean distance algorithm, and the minimum distance is the biased basic emotion.

$$L_n = \sqrt{(P - p_n)^2 + (A - a_n)^2 + (D - d_n)^2}, n = [1, 14], n \in Z \quad (2)$$

L_n is the distance between the 14 basic emotion coordinates and the measured emotion coordinates; P , A , and D are the values of the measured emotion coordinates; and p_n , a_n ,

and d_n are the coordinates of the 14 basic emotion coordinates (Table 4). The magnitude of the distance value reflects the degree to which people's emotion is inclined to the 14 basic emotions. Among them, the basic emotional type with the smallest emotional distance from the participants is considered to be the emotional tendency of PAD after watching stimulus pictures.

Table 4. P , A , and D values of 14 basic emotion coordinates.

Serial Number	Basic Emotional Type	P Value	A Value	D Value
1	Joy	2.77	1.21	1.42
2	Optimism	2.48	1.05	1.75
3	Relaxation	2.19	−0.66	1.05
4	Surprise	1.72	1.71	0.22
5	Gentleness	1.57	−0.79	0.38
6	Dependence	0.39	−0.81	−1.48
7	Boredom	−0.53	−1.25	−0.84
8	Sadness	−0.89	0.17	−0.7
9	Fear	−0.93	1.3	−0.64
10	Anxiety	−0.95	0.32	1.02
11	Contempt	−1.58	0.32	−0.63
12	Disgust	−1.8	0.4	0.67
13	Resentment	−1.98	1.1	0.6
14	Hostility	−2.09	1.00	1.12

In addition, the values of P , A , and D can be used to further analyze people's emotional state (Table 5). Among them, pleasure (P) reflects the availability and friendliness of the object studied. When the state of pleasure is high, people may have fluent thinking and judgment and be able to choose the environment that they truly prefer. When people are in a certain state of arousal (A), it is conducive to concentration, and people may also be more easily attracted to and interested in the environment, but if the arousal degree is too high and the duration is too long, people will be prone to fatigue due to long-term high mental concentration, so the positive and lower level of arousal is more conducive to improving people's emotional experience. For example, the emotion of loss is negative, with a lower level of arousal; anger, on the other hand, is negative, with a higher level of arousal. Dominance (D) refers to whether the observed environment is controllable for people, and the higher the dominance, the lower the people's cognitive load and learning cost, that is, the higher the degree of access to or interaction with the environment. For example, fear is a low level of dominance because it is difficult for people to control their fear.

Table 5. Emotions in the three-dimensional pleasure–arousal–dominance (PAD) space.

Exuberant (+P+A+D)	Bored (−P−A−D)
Dependent (+P+A−D)	Disdainful (−P−A+D)
Relaxed (+P−A+D)	Anxious (−P+A−D)
Docile (+P−A−D)	Hostile (−P+A+D)

- Emotional attachment scale

Combined with the research question and the characteristics of the studied museums, the emotional attachment scale consists of 11 items, including material, color, natural elements, form and structure, privacy, diversity, sociability, territoriality, playability, uniqueness, and changeability, to describe the material, social, and variable characteristics of the studied museum and the surrounding environment as comprehensively as possible (Table 6). The scale is measured based on a Likert scale method, where 1 point indicates no emotional experience of the specific characteristic while 7 points indicate very strong emotional experience. The Cronbach's α of the scale calculated is 0.95, which proves its

reliability. Meanwhile, its validity and utility for the study of people's emotional experience of the built environment have been confirmed in previous studies [26,27,93]. The average value of each characteristic is used as a proxy for the intensity of the participant's emotional experience. Combined with the median score of this scale, when the value of the result is greater than 3.5, the emotional experience is considered to be obvious. Meanwhile, the higher the value, the stronger the emotional experience. SPSS is used to assist in data collation and calculation in this process.

Table 6. Emotional attachment scale of specific environmental characteristics.

	1	2	3	4	5	6	7
Material							
Color							
Natural elements							
Form and structure							
Privacy							
Diversity							
Sociability							
Territoriality							
Playability							
Uniqueness							
Changeability							

3.5. Research Framework

This study is an exploratory quantitative study based on the coupling analysis of multimodal data. Based on the research objectives, the logical framework of this research is shown in Figure 6.

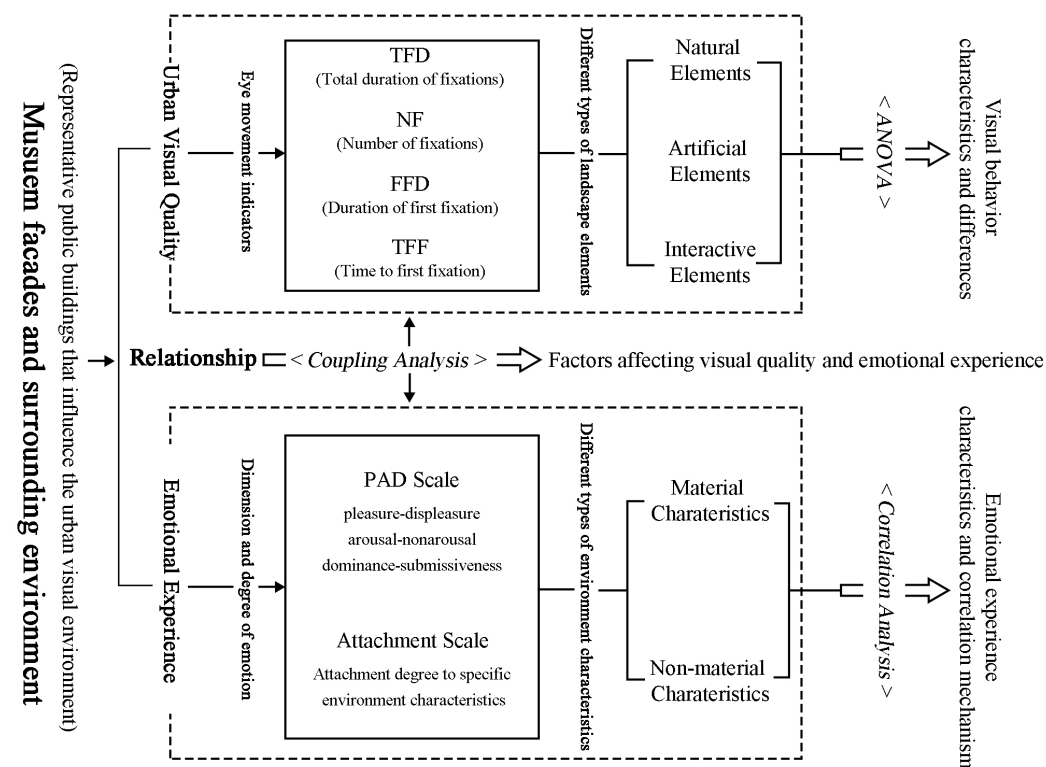


Figure 6. Research framework.

4. Results and Discussion

4.1. Eye-Tracking Results Analysis: Visual Quality Promoted by Urban Museum Facades

4.1.1. AOI Heat Map Analysis: Overall Visual Quality of Museum Facades Based on Fixation Characteristics of Different Elements

The AOI heat maps are used to visualize the position and distribution of fixation points in the three stimulus images representing the three museum facades (Figure 7). Previous studies have shown that areas with denser fixations reflect higher attention and interest of participants to a certain extent, thus contributing more to the visual quality of the urban environment they represent [31,70,85,87]. Therefore, in the results of the AOI heat map, the differences in the visual behavior of the participants on different built environment elements of different museum facades intuitively reveal the impact of related elements on urban visual quality.



Figure 7. AOI heat map results.

Specifically, the heat map results show that the elements that participants pay the most attention to in the three museum facades are all interactive elements including text and logo, as the heat map of such elements on the facade of the three museums is dark red and orange-red. This shows that when people look at the facades of the museum, the elements that can directly convey information on the exterior of the building attract their attention and may raise more interests. This result further supports the conclusion of previous studies that the identification elements such as text have a significant impact on visual quality, and their value in the urban environment seems to be more prominent than their role in visual interaction with books and industrial products [31,70,80,85].

Compared with the interactive elements, the artificial and natural elements of the three museum facade pictures do not have very concentrated fixation points. For artificial features, the heat map shows that participants are attracted to unique materials, colors, or relatively small and more prominent facade components. Meanwhile, the parts with uneven changes or spatial transitions are also more likely to be noticed by participants. These are shown in green on the heat map. On the one hand, this confirms the conclusion that the material differences in the design of facades are conducive to promoting their attractiveness [41,42,87], and on the other hand, it further emphasizes the important role of the spatial characteristics of building form on visual quality.

As for the natural elements, although they are not the elements contained in the museum facade itself, they have become an important component of the urban public environment together with the facade, and thus also affect the role of the museum facade in the visual quality of the city. A large number of theoretical and empirical studies on the visual quality of cities have confirmed the emotional healing effect of natural elements on people [34,62,77,85,86]. The heat maps reveal that the proportion of natural elements may directly affect people's attention to this element; for example, the stimulus picture of the Science and Technology Museum gathered more fixation points than the Party History Museum and the Arts and Crafts Museum, which may be closely related to the denser and more proportional plants in the picture. The important influence of natural element area on visual quality has also been mentioned in previous studies of landscape perception [77].

The AOI heat map provides an overall attraction state of the museum facade to participants, and the impact of specific elements on visual quality and emotional perception needs to be further analyzed in combination with the specific data of eye movement indicators.

4.1.2. Eye-Tracking Index Analysis: Effect of Different Facade Elements on Visual Quality

With the help of SPSS software, descriptive statistical analysis and an ANOVA test were performed on four selected eye movement indicators that were more relevant to participants' emotional experience, so as to explore the influence of different museum facade elements and surrounding environment components on people's visual preferences, as well as the commonalities and differences between the three chosen museums. Based on this, their impact on the visual quality of the city was further explored. At the same time, the influence of participants' gender and major on visual preferences was also discussed.

The descriptive statistical results (Table 7) and difference analysis results (Table 8) of the four eye tracking indicators together provide a detailed description of the differences in the visual characteristics of different urban environments and their specific components on people's visual experience.

Table 7. Descriptive statistical results of eye movement indicators of different visual elements.

Visual Elements		Specific AOI Category	Eye Movement Indicators							
			TTFF (s)		FFD (s)		NF (n)		TFD (s)	
			Mean	SD	Mean	SD	Mean	SD	Mean	SD
Artificial elements	Elements of the building	Walls	1.95	3.574	0.194	0.144	21.665	14.324	5.899	4.486
		Glass	5.13	4.464	0.208	0.121	9.539	10.218	2.505	2.787
		Unique elements	4.333	4.949	0.207	0.098	10.354	9.201	2.754	2.742
	Elements of the street	Road	4.091	9.018	0.06	0.119	0.576	1.434	0.153	0.465
		Fence	7.099	7.676	0.165	0.16	2.084	2.797	0.593	0.923
Natural elements		Plant	3.799	4.006	0.196	0.135	10.122	10.119	2.716	2.911
		Sky	4.813	8.896	0.082	0.136	0.876	2.097	0.227	0.665
Interactive elements		Character	1.184	2.363	0.263	0.269	10.943	8.308	3.387	2.74
		Logo	1.939	4.638	0.115	0.156	3.298	5.263	0.937	1.593

Specifically, the indicator TTFF (time to first fixation) described the length of time it took for people to notice the stimulus and was thought to be an effective reflection of the visual attraction of the stimulus to a person. In general, the shorter the TTFF, the faster the person paid attention to the stimulus, which proved to a certain extent that it was more visually appealing [79]. The results show that there are significant differences between TTFFs of different elements (Tables 7 and 8). In general, the interaction elements had the shortest TTFF, meaning strongest primary visual attraction to people, and the attraction of text was particularly obvious. This is consistent with the results of AOI heat map, and also supports previous research [31,69,70]. At the same time, the standard deviation of TTFF of text was also the smallest, indicating that there was little difference in the degree of attraction to different participants, that is, its influence on urban visual quality may be more universal and important. This showed that text might significantly affect people's emotional experience in the creation of urban visual environment, which has not been emphasized in previous studies relating to urban visual quality, and was necessary to be paid attention to in design. The second smallest category of TTFF was the museum facade wall, which corresponded to the maximum area of each museum facade. Therefore, it was likely that the reason why it was more noticeable was related to its largest area. This

also corresponded to the result that the TTFF value of the fence element was the largest, which may be difficult to pay attention to because it occupied the smallest proportion of the area of the picture and was relatively fragmented. To a certain extent, the influence of the building and its surrounding environment on the urban visual quality was closely related to the visual area occupied by the built environment. Combined with previous studies, this conclusion is valid for different types of elements, and at the same time quantitatively confirms the rationality of classical design theory. At the same time, the results of difference analysis reveal that the three museums with different external environment characteristics composed of different elements also have significant differences in the TTFF results, which further confirmed findings from previous studies showing that the spatial combination of built environment elements may significantly affect people's visual experience [65]. From the perspective of individual differences, the results of difference analysis showed that gender and major both had a significant impact on the results of TTFF, but the influence of major was relatively small, which is different from the conclusion of some researchers that environmental design students may show different visual attention to spatial elements than others [33]. The results may indicate that the first visual attraction of the external space of the museum to people is more derived from the physiological feedback of the preconscious. But due to the imbalance in the proportion of male and female participants in this study, the influence of gender needs further discussion.

Table 8. Difference analysis results of eye tracking indicators.

	TFD (s)	FFD (s)	NF (n)	TFF (s)
Overall Difference	0.000 **	0.000 **	0.000 **	0.000 **
Difference among different visual elements				
Artificial elements (A)	0.000 **	0.000 **	0.000 **	0.000 **
Natural elements (N)	0.000 **	0.000 **	0.000 **	0.000 **
Interactive elements (I)	0.000 **	0.000 **	0.000 **	0.000 **
Difference among different museums				
The Chinese Traditional Culture Museum (CTCM)	0.000 **	0.000 **	0.000 **	0.008 **
The Museum of the Communist Party of China (MCPC)	0.000 **	0.000 **	0.000 **	0.008 **
The China Science and Technology Museum (CSTM)	0.000 **	0.000 **	0.000 **	0.008 **
Difference among participants of different major				
Participants in design related majors (D)	0.047 *	0.059	0.478	0.827
Participants in non-design related majors (ND)	0.047 *	0.059	0.478	0.827
Difference among participants of different gender				
Male (M)	0.003 **	0.894	0.000 **	0.000 **
Female (F)	0.003 **	0.894	0.000 **	0.000 **
n	229	229	229	229

* Significant difference with $p < 0.05$. ** Significant difference with $p < 0.01$.

The indicator FFD (first fixation duration) described the duration of the initial gaze at the studied stimulus and represented the initial visual processing, therefore serving as an indicator of how much an aspect of the stimulus initially attracted attention. It was believed to be particularly valuable when used alongside TTFF. According to previous research, the shorter the TTFF, the longer the FFD, meaning that the stimulus is more eye-catching and might contribute more to urban visual quality. Based on this, character was clearly the most visually appealing, which is consistent with the conclusions of previous studies [87]. The

FFD of glass and unique elements of museum facade was also longer, but its TTFF value was larger at same time, indicating that people might have been interested in the glass of the building facade and the unique facade elements when they saw it, but the speed and degree might not be as strong as the interactive elements such as character. People seem to have spent more time paying attention to the decoration of the building's skin, probably because the decoration provided a wealth of visual information, as mentioned in relating research [41,42,87]. The shortest urban environmental elements of FFD were roads and sky, and their TTFF data were also larger, indicating that the visual attraction was relatively weak, and the impact on urban visual quality was relatively small. Meanwhile, the standard deviation of the FFD was relatively small, indicating that the FFD values for each element were more consistent. Combined with the results of difference analysis, different elements and their combining of different museums' environment have obvious effects on the FFD results, but the gender and major differences in participants do not show significant differences in FFD; this is a discrepancy with the results of previous eye tracking studies conducted on the design major students. To a certain extent, this shows that students generally showed similar cognitive patterns and information processing speeds when dealing with urban visual elements. Meanwhile, gender influence needs further discussion because of the imbalance proportion of participants in this study.

The indicator NF (number of fixations) described times of people fixation on the AOI. In the case of consistent AOI, a greater NF was believed to indicate a stronger interest in the corresponding AOI, which might reflect a stronger visual attraction and emotional experience. The results show that there are significant differences between NFs on different elements and different museums. Specifically, the NF of the wall was significantly larger than that of the other elements, which was more likely due to the fact that the AOI of wall occupies a significantly larger area in the stimulus picture than the other elements. The NF of characters, unique elements, and plants was also larger, while characters account for the smallest AOI area, indicating that it had the strongest attraction, which was consistent with the results of its TTFF and FFD and previous research. Combined with the relatively large SD values, it showed that different people might have different visual experiences of these three types of elements. The results proved the empirical design theory that the unique consideration of facade materials and decoration in the architectural design and the configuration of surrounding natural plants can effectively enhance its contribution to the urban visual quality [42]. In addition, the results of difference analysis showed that the NF value of different genders was significantly different, but different majors did not have a significant effect on the NF value, which may indicate that people's interest in different visual elements of the environment outside the museum is also more affected by preconscious physiological characteristics than acquired experience or education. This complements what has not been mentioned in previous studies.

The indicator TFD (total duration of fixations) describes total duration of fixations inside the AOI and reflected people's cognitive attention. According to previous researches, when the AOI was the same, the longer the TFD, the higher the attention to the corresponding AOI, which might reflect that the elements corresponding to the stimulus have a stronger effect on visual quality and may trigger stronger emotional experience. Therefore, excluding the effect of area, the characters still showed the strongest visual appeal, followed by plants and unique elements. The results of TFD had similar trend to NF, which correlation was consistent with previous eye tracking studies [77,80,81,83].

4.2. Emotional Scale Analysis: Emotional Experience Triggered by Observing Museum Facades

4.2.1. Analysis of PAD Emotional Scale

According to the formula, the calculated *P*, *A*, and *D* values are shown in the Table 9. Combined with the table data and the Euclidean distance algorithm shown in the formula, the emotional tendencies of the participants after viewing three pictures of the museum's exterior and environment can be determined.

Table 9. PAD values of all and different kinds of participants.

		P		A		D	
		Mean	SD	Mean	SD	Mean	SD
All participants (n = 229)		1.713	1.193	0.072	1.191	−0.264	1.129
Gender difference	Male (n = 36)	1.688	0.976	−0.215	1.049	−0.326	1.119
	Female (n = 193)	1.718	1.229	0.126	1.208	−0.253	1.130
Major difference	Design-related (n = 105)	1.742	1.254	0.077	1.187	−0.268	1.135
	Non-design-related (n = 124)	1.679	1.117	0.067	1.195	−0.260	1.122

As is shown in Table 9, the overall PAD values for all participants were 1.713; gender differences and major differences had no significant impact on the positive or negative value and size of PAD values. However, due to the large differences in the sex ratio of participants, this result needs to be further explored. Specifically, the *P* value is positive, indicating that the content of the experimental pictures brings positive emotional experience to the participants; the *A* value is positive, indicating that the participants are in an excited state when watching the pictures during the experiment; and the *D* value is negative, indicating that the participants have a weak sense of control over the scene and are in an affected state. According to Tables 3 and 4 in Section 3.4.2, the emotional model of the results corresponded to “gentleness” and the emotional tendency corresponded to “dependent”. The former reflects the participants’ stable emotions and attitudes towards their current situation and emotional state, while the latter reflects positive emotional experience, higher neurophysiological activation levels, and lower individual control to a certain extent, indicating that the participants have developed trust in the environment they are viewing. These emotional reactions reveal to some extent that the museum’s facades and surrounding environment successfully evokes participants’ positive emotions, which not only provides a new application scope of the PAD scale in the study of the built environment, but also to some extent provides relatively objective and scientific evidence for the empirical conclusion that museums and other cultural buildings have important emotional value for people in classical research [52,67,92].

According to the Formula (2), the spatial coordinate distance relationship between the emotional state of males, females, design-related students, and non-design-related students was calculated (Table 10), respectively, and the affective tendency of the students was “surprise”. Combined with the values of PAD in the above Table 9, it can be found that the mean value of *P* of different genders and majors is greater than 0, which indicates to a certain extent that males, females, design-related students, and non-design-related students may feel happy because of the novelty, uniqueness or grandeur after viewing the facade and the external environment of the museum. According to previous studies, this pleasure may not be directly caused by the “surprise” itself, but by the cognitive or aesthetic experience triggered by the emotion of surprise [93]. For the value of *A*, the mean value of male subjects was abnormally less than 0, indicating that the surprised emotion obtained from watching the pictures of stimulus materials did not trigger the physiological and psychological activation of the male group, which may be related to

the relative emotional stability of men on the one hand, and may also reveal the gender difference in visual preference on the other hand. This is similar to the conclusion of previous studies that gender differences significantly affect the emotional experience of the same subject [1,3,16,17]. However, since the ratio of male to female subjects in this study is affected by the overall male and female ratio of recruiting university, the validity of this conclusion needs to be further explored. The distribution of D value was similar to that of P value, which was less than zero among participants of different majors, indicating that the emotion of surprise was related to the lack of control or predictability of the observed situation. Meanwhile, the specific values show that men's emotions are more affected than women's, which may indicate that men have a higher degree of trust in the environment of stimulus performance, but because men are a minority among the participants, this result needs to be further explored in follow-up studies.

Table 10. Results of the spatial coordinate distance relationship.

Serial Number	Ln					Emotional Type
	All	Male	Female	Design-Related	Non-Design-Related	
1	2.290	1.740	1.410	1.466	1.453	Joy
2	2.365	2.097	1.771	1.827	1.813	Optimism
3	1.575	2.692	2.564	2.603	2.550	Relaxation
4	1.709	0.550	0.266	0.306	0.302	Surprise
5	1.081	2.573	2.535	2.562	2.503	Gentleness
6	2.001	2.819	3.009	3.001	2.945	Dependence
7	2.664	3.023	3.134	3.140	3.079	Boredom
8	2.773	1.628	1.776	1.774	1.715	Sadness
9	2.937	0.661	0.912	0.886	0.846	Fear
10	2.697	1.468	1.612	1.610	1.552	Anxiety
11	3.539	1.877	1.678	1.724	1.679	Contempt
12	3.646	1.602	1.448	1.489	1.437	Disgust
13	3.926	1.063	0.817	0.868	0.828	Resentment
14	4.139	1.543	1.249	1.305	1.278	Hostility

4.2.2. Analysis of the Emotional Attachment Scale of Specific Environmental Characteristics

If the above PAD scale results provide the participants' overall emotional tendencies after viewing the three photos of the museum's facade and surrounding environment and the impact of major differences on them, then the environmental characteristics emotional attachment scale specifically reveals the extent to which the material and non-material characteristics that make up the museum's exterior environment affect people's emotional attachment and their internal mechanisms of action from the perspective of the built environment.

The results of correlation analysis (Table 11) show that both the material and non-material characteristics of the museum's exterior environment significantly promote the overall emotional attachment. Non-material characteristics have a stronger effect in promoting emotional attachment, as the correlation factors of four of the characteristics are greater than 0.550, which partly confirms the conclusion that the non-physical social characteristics and activities carried by architecture may have a greater impact on people's emotions than materiality [67]. Among them, the diversity and playability have the most significant positive effect on emotional attachment among all characteristics (the correlation factors are 0.632 ** and 0.649 **, respectively, both greater than 0.600), which to a certain extent shows that the diverse characteristics and possible playability of the museum's exterior environment perceived by the subjects through vision can effectively trigger their emotional attachment to the environment. This conclusion is rare in previous studies based on the natural environment. There is also a significant positive correlation between

diversity and playability, which to a certain extent indicates that diverse external environment characteristics may promote users' expectations of its fun. Meanwhile, according to the results, uniqueness also significantly promotes emotional attachment, and correlation analysis shows that it is significantly positively correlated with the form and structure of material characteristics, which to a certain extent shows that the external space, surface structure and spatial form of the museum are important reasons for its uniqueness and thus trigger the emotional experience of the subjects. In comparison, among the non-material characteristics, changeability has the smallest effect in promoting emotional attachment, and the correlation analysis results also show that changeability is not strongly correlated with other characteristics. On the one hand, this may indicate that the visual characteristics of the three museum exterior environments do not significantly reveal their potential for change. To a certain extent, this proves the value of physical perception to emotion in the experience of architectural space as mentioned in embodied design researches [17,18]. On the other hand, it may also be related to the fact that the stimulus materials of the three museums in this study are static two-dimensional pictures.

Table 11. The results of correlation analysis of attachment to specific environmental characteristics.

	1	2	3	4	5	6	7	8	9	10	11	12
1 Overall attachment	1											
2 Material	0.548 **	1										
3 Color	0.526 **	0.347 **	1									
4 Natural elements	0.510 **	0.198 **	0.389 **	1								
5 Form and structure	0.587 **	0.440 **	0.358 **	0.274 **	1							
6 Privacy	0.406 **	0.242 **	0.067	0.092	0.129	1						
7 Diversity	0.632 **	0.273 **	0.232 **	0.209 **	0.371 **	0.224 **	1					
8 Sociability	0.468 **	0.125	0.130 *	0.147 *	0.140 *	0.123	0.266 **	1				
9 Territoriality	0.454 **	0.229 **	0.146 *	0.202 **	0.191 **	0.019	0.125	0.342 **	1			
10 Playability	0.649 **	0.241 **	0.287 **	0.236 **	0.310 **	0.151 *	0.425 **	0.271 **	0.197 **	1		
11 Uniqueness	0.590 **	0.239 **	0.225 **	0.276 **	0.407 **	0.126	0.386 **	0.019	0.356 **	0.340 **	1	
12 Changeability	0.578 **	0.198 **	0.199 **	0.185 **	0.212 **	0.279 **	0.318 **	0.244 **	0.093	0.424 **	0.221 **	1

* $p < 0.05$ (2-tailed), ** $p < 0.01$ (2-tailed).

As for material characteristics, the results show that the four types of characteristics significantly promoted the establishment of emotional attachment in the three museums (the correlation factors were all greater than 0.500 **), which to a certain extent confirmed the high visual quality of the external environment of the studied museums. Among them, form and structure have the strongest driving effect on emotional attachment, which once again proves the value of the architectural features of the museum's facade in triggering emotions through vision as mentioned in series of previous studies [42,87]. At the same time, form and structure show a significant correlation with materials (0.440 **), further indicating that materials have a major influence on form and structure. That is, the facade forms of the three museums reflect the mechanical properties of the materials to a certain extent, and the architectural exterior design that reveals the characteristics of the materials is conducive to the viewer's emotional experience of it. This has been repeatedly mentioned in a large number of architectural theories and practices, and this study objectively supports it in a quantitative way [66,67].

4.3. Coupling Analysis of Multimodal Data: Relationship Between Visual Quality and Emotional Experience

By comprehensively analyzing the above eye tracking data and emotional scale data, the relationship between the visual quality of the museum's outdoor environment and people's emotional experience can be further explored.

In general, both the PAD scale and the emotional attachment scale of specific environmental characteristics indicate that people had significant positive emotional experiences after viewing the three museum exterior pictures, and the emotions were in a comfortable and stable state. This proves to a certain extent that the facades and the external urban space of the three museums have created a good environmental visual quality. On the basis of the results of previous studies on the visual quality of building facades, it provides a more specific building type supplement.

Specifically, the environmental features that have the most obvious impact on visual quality are interactive material elements such as text and logos, which is consistent with previous studies. Although they occupy a very small area in the pictures of the museum's external environment, the eye tracking data and AOI heat map distribution both prove their strongest visual appeal. And as visual elements that trigger people's thinking and consciousness, interactive elements usually define the identity of the museum, thereby strengthening the degree of emotional attachment between people and the museum by highlighting museum's non-material features such as diversity and uniqueness. Combined with the significant role of non-material features in the establishment of attachment according to the results of the emotional scales, it can be further inferred that the stronger the visual appeal of environmental features, the stronger the degree of emotional experience may be, and the greater the impact on the overall visual quality they may have. This also complements the possible bias in the conclusions of previous studies with single-modal data.

The materials, structures, and the way they are assembled in the facade of the museum building also show a significant impact on the visual quality. In general, the FD in the eye tracking data and the significant correlation between material and structure and attachment in the emotional scale data both prove that the facade of the museum building studied has improved the overall visual quality of the city streets where they are located and thus brought people a positive emotional experience. Among them, the AOI heat map, TTFF, and FFD all show that the unique decorations in the facade and the parts with spatial concave and convex changes always attract more visual attention. Combined with the emotional scale data, it can be inferred that the use of these unique materials and the three-dimensional volume changes in the facade may promote people's attachment by triggering their expectations for diversity and changeability, thereby improving the urban visual quality contributed by the museum facade. In addition, the eye-tracking results also show that the area of facade materials has an important impact on visual quality. This is consistent with previous studies showing that just because a single material used in a large area has many visual focus points, this does not necessarily make it attractive [31]. Subjects' overall stable and positive emotional state revealed by the above PAD scale illustrates the role of the visual integrity of the facade in creating visual quality from another perspective. These multi-dimensional data together show that the area ratio of the color, material, and spatial structure of the building facade significantly affect the viewer's emotional experience, thereby affecting the visual quality of the urban environment. Although this has been confirmed in previous studies based on the natural environment, it has more practical and methodological value for the design of the built environment.

In addition to the abovementioned characteristics of the museum building facade, the research results also show that the natural elements around the building also have a significant effect on the emotional experience it triggers, because people always take the

building and the surrounding environment as a whole into their field of vision through visual perception. The eye-tracking results show that the trees and the sky, which occupy a relatively small area in the museum pictures, still attract more gaze points, and the results of the emotional attachment scale also reveal a significant correlation between natural elements and colors, which once again confirms the role of natural colors in promoting emotional experience shown in previous studies, as well as the value of the city's blue sky rate that people can perceive to the visual quality of the urban environment [54–56].

4.4. Limitations and Future Prospects

In addition to the above results, this study still has the following limitations that need further exploration:

First, two-dimensional images were used as stimulus materials in the study of the visual quality of the three museums, which was limited by the way Tobii Pro Fusion is used to collect eye tracking data, needing to be fixed at the bottom of the screen; although the researchers controlled for variables such as angle, light, and weather as much as possible during the shooting process, the real space of the museum facade and the surrounding environment may have had different effects on visual perception than two-dimensional images, so the representativeness of the conclusions for the real built environment needs further discussion. With the maturity of portable wearable eye trackers, future research can be extended to real-world built environments.

Second, the subjects participating in this study were all college students, and the number of males was significantly smaller due to the imbalance in the ratio of men to women in the school itself. Because of this, the results of this study cannot be generalized to the entire population. The difference in emotional feedback on visual quality among different groups and individuals needs to be paid attention to in the future, both in gender difference and age difference.

Third, in the experimental design of this study, visual perception is measured by eye tracking devices first, and then emotional attachment is measured by scale, and the logic is unidirectional. However, the process of people viewing environmental elements in a real urban environment and gaining emotions may be cyclical or mutually influential, so it is necessary to further explore the possible impact of emotional perception on visual quality. At the same time, the diversity in data modalities and the depth of coupling analysis need to be further strengthened.

Finally, the case of this study comprised three city museums in Beijing. In the future, the relevant methods can be applied to more types of urban buildings and spaces to expand the depth and breadth of the research.

5. Conclusions

In summary, this study reveals the impact of different types and characteristics of built environment elements on the visual quality of the urban environment using three newly built urban museums in Beijing as an example, and proves the important role of visual quality in people's perception, cognition, and emotional experience, further illustrating that designing and creating a good visual environment is conducive to promoting people's emotional attachment to space, thereby creating a humanistic urban environment. Tobii Pro Fusion and semi-structured emotional scales were used for multimodal data collection and analysis, which at the same time led to the establishment of a method of urban visual quality research and evaluation from an emotional perspective, providing a more scientific and objective way for the intervention of subjective and complex "emotions" in built environment research. The results of the eye-tracking AOI heat map, eye-tracking index, and emotional scale comprehensively confirmed the original hypothesis of this study, that

is, different types of museum facades and external environmental elements have different effects on visual quality and trigger people's emotional experience to varying degrees.

Specifically, from the perspective of the built environment, the analysis results based on the feature type can be condensed to guide the visual quality improvement and design of future museum buildings. Among them, interactive elements show the most significant impact on visual quality and emotional perception and are not affected by their smaller visual perception area. This reveals that urban visual elements that trigger people's thinking and cognitive processes are not only conducive to enhancing the visual appeal of building facades but can help promote people's emotional attachment by emphasizing the diversity and uniqueness of built environment. The degree of influence of artificial elements on visual quality is related to their size and spatial composition on the museum facade as well as their non-material characteristics such as uniqueness, variability, and diversity. Meanwhile, the increase in visual points does not mean an increase in attraction or a deepening of emotional experience; when the artificial elements consisting of the museum facade have a concave–convex relationship or have a certain symbolic meaning to the museum content, it is more likely to trigger the viewer's emotional experience. As for natural elements, although the vegetation around the museum does not occupy a central position or a large area in the three museum photos, it has a relatively obvious impact on the overall visual quality of the environment. Meanwhile, nature-related colors are proved to significantly promote people's emotional experience.

Meanwhile, from the perspective of people's emotional experience, the above different elements exert different effects on the visual quality created by the museum facade and its external environment, which is significantly correlated with people's emotional experience after viewing relating pictures. Combining the multimodal data results, it can be found that this correlation is established in this study in two ways: preconscious emotional perception and conscious emotional cognition. The former is mostly caused by material elements and characteristics, and the degree of influence is related to the area of the element and whether it comes from nature; generally speaking, the larger the area and the closer the element is to nature, the more obvious the impact on emotional perception. Among them, the impact of area is directly related to the size of the field of view, and the impact of nature confirms the emotional value of Biophilia experience, which is consistent with the conclusion of previous studies that even viewing natural pictures can achieve emotional healing. The latter is derived from human cognitive processing, so the signs and texts that can trigger people's thinking process are the most obvious visual elements that affect this kind of emotional experience. In this case, the generation of emotions is often not affected by the visual area occupied by the elements. In addition, material characteristics such as materials, structures, and forms may also trigger people's emotional experience through non-material characteristics such as the uniqueness, territoriality, or potential changeability they express. This experience also comes from the role of thinking, such as the sense of place identity brought by the use of material forms that conform to the identity of the museum, or the novelty brought by the unique structural design. The conclusion not only provides a reference path for constructing an emotionally oriented architectural and urban space design method but also serves as a logical framework for evaluating the emotional attachment strength of the existing built environment.

Author Contributions: Conceptualization, methodology, R.Z.; software, D.C., Z.Z. (Zhenzhi Zhou), R.S., Z.Z. (Zekai Zhang), C.C., and R.Z.; validation, R.Z.; formal analysis, R.Z., D.C., and Z.Z. (Zhenzhi Zhou); investigation, D.C., Z.Z. (Zhenzhi Zhou), R.S., Z.Z. (Zekai Zhang), and C.C.; resources, D.C., Z.Z. (Zhenzhi Zhou), R.S., Z.Z. (Zekai Zhang), and C.C.; data curation, R.Z., D.C., Z.Z. (Zhenzhi Zhou), R.S., Z.Z. (Zekai Zhang), and C.C.; writing—original draft preparation, writing—review and editing, R.Z.; visualization, R.Z., D.C., and Z.Z. (Zhenzhi Zhou); supervision, R.Z.; project

administration, R.Z.; funding acquisition, R.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Natural Science Foundation of China under Grant No. 52208005; Beijing Municipal Social Science Foundation under Grant No. 22GLC063.

Institutional Review Board Statement: The human study protocol was approved by the Human Study Ethics Committee of Beijing Forestry University (protocol code: BJFUPSY-2024-066, approved date: 18 August 2023).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data are unavailable due to privacy restrictions.

Acknowledgments: The authors would like to thank the Tobii China Innovation Program for the instruments, knowledge, and technical support for this study.

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

References

1. Ellis, D.; Tucker, I. *Emotion in the Digital Age: Technologies, Data and Psychosocial Life*; Routledge: London, UK, 2020.
2. Ogburn, W.F. Cultural lag as theory. *Sociol. Soc. Res.* **1957**, *41*, 161–167.
3. Ansaloni, F.; Tedeschi, M. Understanding space ethically through affect and emotion: From uneasiness to fear and rage in the city. *Emot. Space Soc.* **2016**, *21*, 15–22. [\[CrossRef\]](#)
4. Kaklauskas, A.; Bardauskiene, D.; Cerkauskienė, R.; Ubarte, I.; Raslanas, S.; Radvilė, E.; Kaklauskaitė, U.; Kaklauskienė, L. Emotions analysis in public spaces for urban planning. *Land Use Policy* **2021**, *107*, 105458. [\[CrossRef\]](#)
5. Weijs-Perrée, M.; Dane, G.; van den Berg, P. Analyzing the relationships between citizens' emotions and their momentary satisfaction in urban public spaces. *Sustainability* **2020**, *12*, 7921. [\[CrossRef\]](#)
6. Snowden, R.J.; Snowden, R.; Thompson, P.; Troscianko, T. *Basic Vision: An Introduction to Visual Perception*; Oxford University Press: Oxford, UK, 2012.
7. Jin, T.; Lu, J.; Shao, Y. Exploring the Impact of Visual and Aural Elements in Urban Parks on Human Behavior and Emotional Responses. *Land* **2024**, *13*, 1468. [\[CrossRef\]](#)
8. Yang, S.; Dane, G.; van den Berg, P.; Arentze, T. Influences of cognitive appraisal and individual characteristics on citizens' perception and emotion in urban environment: Model development and virtual reality experiment. *J. Environ. Psychol.* **2024**, *96*, 102309. [\[CrossRef\]](#)
9. Kochergina, E. Urban planning aspects of museum quarters as an architectural medium for creative cities. In Proceedings of the IOP Conference Series: Materials Science and Engineering, Busan, Republic of Korea, 25–27 August 2017; p. 052031.
10. Mihaila, M. Museum side of the city—from the theory to inquiry. *Procedia-Soc. Behav. Sci.* **2014**, *149*, 570–574. [\[CrossRef\]](#)
11. Liu, Y. The city of museums helps Beijing's central axis apply for World Heritage status. *Outlook New Era*, 24 April 2023; p. 002.
12. Hu, X.; Tao, S. Agglomeration and Integration: A Preliminary Study on the Construction Path of City of Museums (Beijing) from the Perspective of Space. *Chin. Mus.* **2023**, *4*, 4–13/135.
13. Zhang, J. Promotion of International Communication Function of Beijing by Museums. *Collect. Essays About Cap. Mus. China* **2023**, *2023*, 103–108.
14. Deonna, J.A. Emotion, perception and perspective. *Dialectica* **2006**, *60*, 29–46. [\[CrossRef\]](#)
15. Russell, J.A. Measures of emotion. In *The Measurement of Emotions*; Elsevier: Amsterdam, The Netherlands, 1989; pp. 83–111.
16. Meiselman, H.L. *Emotion Measurement*; Woodhead Publishing: Sawston, UK, 2016.
17. Mauss, I.B.; Robinson, M.D. Measures of emotion: A review. *Cogn. Emot.* **2009**, *23*, 209–237. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Park, C.L.; Kubzansky, L.D.; Chafouleas, S.M.; Davidson, R.J.; Keltner, D.; Parsafar, P.; Conwell, Y.; Martin, M.Y.; Hanmer, J.; Wang, K.H. Emotional well-being: What it is and why it matters. *Affect. Sci.* **2023**, *4*, 10–20. [\[CrossRef\]](#) [\[PubMed\]](#)
19. Pütz, S.; Mertens, A.; Chuang, L.; Nitsch, V. Physiological measures of operators' mental state in supervisory process control tasks: A scoping review. *Ergonomics* **2024**, *67*, 801–830. [\[CrossRef\]](#)
20. Kitagawa, A.; Murata, K.; Uraoka, Y.; Furuta, M.; Munaka, T. Human Metrics Explorer System for Multi-Device Physiological Measurements in Emotion Estimation. In Proceedings of the 2024 IEEE/SICE International Symposium on System Integration (SII), Ha Long, Vietnam, 8–11 January 2024; pp. 184–189.
21. Yang, M.; Sun, B.; Yang, J. The Influence of Emotion Expression on Ergonomics in Interface Design. In Proceedings of the 2020 IEEE International Conference on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking (ISPA/BDCloud/SocialCom/SustainCom), Exeter, UK, 17–19 December 2020; pp. 1156–1162.

22. Demirbilek, O.; Sener, B. Product design, semantics and emotional response. *Ergonomics* **2003**, *46*, 1346–1360. [\[CrossRef\]](#)
23. Baber, C. *Embodying Design: An Applied Science of Radical Embodied Cognition*; MIT Press: Cambridge, MA, USA, 2022.
24. PS, S.; Mahalakshmi, G. Emotion models: A review. *Int. J. Control Theory Appl.* **2017**, *10*, 651–657.
25. Sander, D. Models of emotion. In *The Cambridge Handbook of Human Affective Neuroscience*; Cambridge University Press: New York, NY, USA, 2013; pp. 5–56.
26. Sander, D.; Scherer, K. *Oxford Companion to Emotion and the Affective Sciences*; OUP Oxford: Oxford, UK, 2014.
27. Kaldany, A.; Patel, H.V.; Shaw, N.M.; Jones, C.P.; Breyer, B.N. Ergonomics in Urology: Current Landscape and Future Directions. *Urology* **2023**, *184*, 235–243. [\[CrossRef\]](#)
28. Javohir, S.; Omina, K.; Gulnoz, X.; Azamat, S.; Faxrifdin, Y.; Asaloy, A. Exploring the Intersection of Science and Art: The Scientific Foundations of Architecture. *Hold. Reason* **2024**, *4*, 79–83.
29. Eilouti, B. A framework for integrating ergonomics into architectural design. *Ergon. Des.* **2023**, *31*, 4–12. [\[CrossRef\]](#)
30. Zhang, R. Evaluation of emotional attachment characteristics of small-scale urban vitality space based on technique for order preference by similarity to ideal solution, integrating entropy weight method and grey relation analysis. *Land* **2023**, *12*, 613. [\[CrossRef\]](#)
31. Zhang, R.; Duan, W.; Zheng, Z. Multimodal Quantitative Research on the Emotional Attachment Characteristics between People and the Built Environment Based on the Immersive VR Eye-Tracking Experiment. *Land* **2024**, *13*, 52. [\[CrossRef\]](#)
32. Wang, P.; Yang, W.; Wang, D.; He, Y. Insights into public visual behaviors through eye-tracking tests: A study based on national park system pilot area landscapes. *Land* **2021**, *10*, 497. [\[CrossRef\]](#)
33. Lan, Y.; Liu, Q.; Zhu, Z. Exploring Landscape Design Intensity Effects on Visual Preferences and Eye Fixations in Urban Forests: Insights from Eye Tracking Technology. *Forests* **2023**, *14*, 1628. [\[CrossRef\]](#)
34. Qi, Z.; Jin, X.; Liang, X.; Han, B.; Zhou, Y. A new framework for dynamic visual landscape quality evaluation of greenways based on perceived sensory dimensions: A case study of Chengdu urban greenway. *J. Environ. Manag.* **2024**, *370*, 122950. [\[CrossRef\]](#) [\[PubMed\]](#)
35. He, N.; Li, G. Urban neighbourhood environment assessment based on street view image processing: A review of research trends. *Environ. Chall.* **2021**, *4*, 100090. [\[CrossRef\]](#)
36. Liang, X.; Zhao, T.; Biljecki, F. Revealing spatio-temporal evolution of urban visual environments with street view imagery. *Landsc. Urban Plan.* **2023**, *237*, 104802. [\[CrossRef\]](#)
37. Mundher, R.; Abu Bakar, S.; Al-Helli, M.; Gao, H.; Al-Sharaa, A.; Mohd Yusof, M.J.; Maulan, S.; Aziz, A. Visual Aesthetic Quality Assessment of Urban Forests: A Conceptual Framework. *Urban Sci.* **2022**, *6*, 79. [\[CrossRef\]](#)
38. Wang, Z.; Ito, K.; Biljecki, F. Assessing the equity and evolution of urban visual perceptual quality with time series street view imagery. *Cities* **2024**, *145*, 104704. [\[CrossRef\]](#)
39. Sullivan, W.C.; Lovell, S.T. Improving the visual quality of commercial development at the rural–urban fringe. *Landsc. Urban Plan.* **2006**, *77*, 152–166. [\[CrossRef\]](#)
40. Taylor, N. Legibility and aesthetics in urban design. *J. Urban Des.* **2009**, *14*, 189–202. [\[CrossRef\]](#)
41. Gjerde, M. Visual aesthetic perception and judgement of urban streetscapes. In Proceedings of the Paper for Building a Better World: CIB World Congress, Salford Quays, UK, 10–13 May 2010; pp. 12–22.
42. Nasar, J.L. Urban design aesthetics: The evaluative qualities of building exteriors. *Environ. Behav.* **1994**, *26*, 377–401. [\[CrossRef\]](#)
43. Lynch, K. *Good City Form*; MIT Press: Cambridge, MA, USA, 1984.
44. Lynch, K. *The Image of the City*; MIT Press: Cambridge, MA, USA, 1964.
45. Koolhaas, R. *Delirious New York: A Retroactive Manifesto for Manhattan*; The Monacelli Press, LLC: New York, NY, USA, 2014.
46. Maiocchi, M. *The Neuroscientific Basis of Successful Design: How Emotions and Perceptions Matter*; Springer: Berlin/Heidelberg, Germany, 2014.
47. Döveling, K.; Döveling, K.; von Scheve, C.; Konijn, E.A. *The Routledge Handbook of Emotions and Mass Media*; Routledge: London, UK, 2010.
48. Gulliksen, H. Mathematical solutions for psychological problems. *Am. Sci.* **1959**, *47*, 178–201.
49. Busemeyer, J.R.; Wang, Z.; Townsend, J.T.; Eidels, A. *The Oxford Handbook of Computational and Mathematical Psychology*; Oxford University Press: Oxford, UK, 2015.
50. Hadamard, J. *The Mathematician's Mind: The Psychology of Invention in the Mathematical Field*; Princeton University Press: Princeton, NJ, USA, 2020; Volume 18.
51. Roe, J.; McCay, L. *Restorative Cities: Urban Design for Mental Health and Wellbeing*; Bloomsbury Publishing: London, UK, 2021.
52. Carmona, M. *Public Places Urban Spaces: The Dimensions of Urban Design*; Routledge: London, UK, 2021.
53. Sieff, D.F. *Understanding and Healing Emotional Trauma*; Taylor & Francis: Oxford, UK, 2014.
54. Buzzell, L.; Chalquist, C. *Ecotherapy: Healing with Nature in Mind*; Catapult: Tokyo, Japan, 2010.
55. Burkhardt, M.A. Healing relationships with nature. *Complement. Ther. Nurs. Midwifery* **2000**, *6*, 35–40. [\[CrossRef\]](#) [\[PubMed\]](#)

56. White, M.P.; Hartig, T.; Martin, L.; Pahl, S.; van den Berg, A.E.; Wells, N.M.; Costongs, C.; Dzhambov, A.M.; Elliott, L.R.; Godfrey, A. Nature-based biopsychosocial resilience: An integrative theoretical framework for research on nature and health. *Environ. Int.* **2023**, *181*, 108234. [[CrossRef](#)] [[PubMed](#)]
57. Adams, W.W. *A Wild and Sacred Call: Nature–Psyche–Spirit*; State University of New York Press: New York, NY, USA, 2023.
58. Zhang, K.; Zhang, J.; Yang, J. The influence of human elements in photographs on tourists' destination perceptions and intentions. *Tour. Manag.* **2023**, *95*, 104684. [[CrossRef](#)]
59. Castelo, S.V.M.A. *EmoPhoto: Identification of Emotions in Photos*; University of Lisbon: Lisbon, Portugal, 2014.
60. Ma, Z.; Yang, G. Show me a photo of the character: Exploring the interaction between text and visuals in narrative persuasion. *J. Health Commun.* **2022**, *27*, 125–133. [[CrossRef](#)]
61. Becker, M.W. Negative emotional photographs are identified more slowly than positive photographs. *Atten. Percept. Psychophys.* **2012**, *74*, 1241–1251. [[CrossRef](#)]
62. Ulrich, R.S. Visual landscapes and psychological well-being. *Landsc. Res.* **1979**, *4*, 17–23. [[CrossRef](#)]
63. Herzog, T.R.; Maguire, P.; Nebel, M.B. Assessing the restorative components of environments. *J. Environ. Psychol.* **2003**, *23*, 159–170. [[CrossRef](#)]
64. Meyer-Grandbastien, A.; Burel, F.; Hellier, E.; Bergerot, B. A step towards understanding the relationship between species diversity and psychological restoration of visitors in urban green spaces using landscape heterogeneity. *Landsc. Urban Plan.* **2020**, *195*, 103728. [[CrossRef](#)]
65. Tang, J.; Long, Y. Measuring visual quality of street space and its temporal variation: Methodology and its application in the Hutong area in Beijing. *Landsc. Urban Plan.* **2019**, *191*, 103436. [[CrossRef](#)]
66. Frampton, K. *Le Corbusier*; Thames & Hudson: London, UK, 2024.
67. Frampton, K. *Modern Architecture: A Critical History (Fifth)(World of Art)*; Thames & Hudson: London, UK, 2020.
68. Joye, Y.; Kahn, P.H.; Hasbach, P.H. *Can Architecture Become Second Nature?: An Emotion-Based Approach to Nature-Oriented Architecture*; MIT Press: Cambridge, MA, USA, 2012.
69. Rayner, K.; McConkie, G.W.; Zola, D. Integrating information across eye movements. *Cogn. Psychol.* **1980**, *12*, 206–226. [[CrossRef](#)] [[PubMed](#)]
70. Rayner, K. Eye movements in reading and information processing. *Psychol. Bull.* **1978**, *85*, 618. [[CrossRef](#)] [[PubMed](#)]
71. Morimoto, C.H.; Mimica, M.R. Eye gaze tracking techniques for interactive applications. *Comput. Vis. Image Underst.* **2005**, *98*, 4–24. [[CrossRef](#)]
72. Dodge, R.; Cline, T.S. The angle velocity of eye movements. *Psychol. Rev.* **1901**, *8*, 145. [[CrossRef](#)]
73. Tinker, M.A. Recent studies of eye movements in reading. *Psychol. Bull.* **1958**, *55*, 215. [[CrossRef](#)]
74. Johnson, S.P.; Amso, D.; Slemmer, J.A. Development of object concepts in infancy: Evidence for early learning in an eye-tracking paradigm. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 10568–10573. [[CrossRef](#)]
75. Wickens, C.D.; Gempfer, K.; Morphew, M.E. Workload and reliability of predictor displays in aircraft traffic avoidance. *Transp. Hum. Factors* **2000**, *2*, 99–126. [[CrossRef](#)]
76. Aoki, H.; Itoh, K. Analysis of influences of aural information on viewers' visual cognition during viewing of television commercials by use of eye tracking technique. *Jpn. J. Ergon.* **2001**, *37*, 246–247.
77. Dupont, L.; Antrop, M.; Van Eetvelde, V. Eye-tracking analysis in landscape perception research: Influence of photograph properties and landscape characteristics. *Landsc. Res.* **2014**, *39*, 417–432. [[CrossRef](#)]
78. Popelka, S.; Brychtova, A. Eye-tracking study on different perception of 2D and 3D terrain visualisation. *Cartogr. J.* **2013**, *50*, 240–246. [[CrossRef](#)]
79. Eckstein, M.K.; Guerra-Carrillo, B.; Singley, A.T.M.; Bunge, S.A. Beyond eye gaze: What else can eyetracking reveal about cognition and cognitive development? *Dev. Cogn. Neurosci.* **2017**, *25*, 69–91. [[CrossRef](#)] [[PubMed](#)]
80. Skaramagkas, V.; Giannakakis, G.; Ktistakis, E.; Manousos, D.; Karatzanis, I.; Tachos, N.S.; Tripoliti, E.; Marias, K.; Fotiadis, D.I.; Tsiknakis, M. Review of eye tracking metrics involved in emotional and cognitive processes. *IEEE Rev. Biomed. Eng.* **2021**, *16*, 260–277. [[CrossRef](#)] [[PubMed](#)]
81. Kiefer, P.; Giannopoulos, I.; Raubal, M.; Duchowski, A. Eye tracking for spatial research: Cognition, computation, challenges. *Spat. Cogn. Comput.* **2017**, *17*, 1–19. [[CrossRef](#)]
82. Critchley, H.D.; Eccles, J.; Garfinkel, S.N. Interaction between cognition, emotion, and the autonomic nervous system. In *Handbook of Clinical Neurology*; Elsevier: Amsterdam, The Netherlands, 2013; Volume 117, pp. 59–77.
83. Liu, Y.; Hu, M.; Zhao, B. Audio-visual interactive evaluation of the forest landscape based on eye-tracking experiments. *Urban For. Urban Green.* **2019**, *46*, 126476. [[CrossRef](#)]
84. Scott, N.; Le, D.; Becken, S.; Connolly, R.M. Measuring perceived beauty of the Great Barrier Reef using eye-tracking technology. *Curr. Issues Tour.* **2020**, *23*, 2492–2502. [[CrossRef](#)]
85. Zhou, X.; Cen, Q.; Qiu, H. Effects of urban waterfront park landscape elements on visual behavior and public preference: Evidence from eye-tracking experiments. *Urban For. Urban Green.* **2023**, *82*, 127889. [[CrossRef](#)]

86. Nordh, H. Quantitative methods of measuring restorative components in urban public parks. *J. Landsc. Archit.* **2012**, *7*, 46–53. [[CrossRef](#)]
87. Fu, H.; Wang, P.; Zhou, J.; Zhang, S.; Li, Y. Investigating Influence of Visual Elements of Arcade Buildings and Streetscapes on Place Identity Using Eye-Tracking and Semantic Differential Methods. *Buildings* **2023**, *13*, 1580. [[CrossRef](#)]
88. Wakil, K.; Naeem, M.A.; Anjum, G.A.; Waheed, A.; Thaheem, M.J.; Hussnain, M.Q.U.; Nawaz, R. A hybrid tool for visual pollution Assessment in urban environments. *Sustainability* **2019**, *11*, 2211. [[CrossRef](#)]
89. Zhao, X.; Wu, Y.; Cui, W.; Du, X.; Chen, Y.; Wang, Y.; Lee, D.L.; Qu, H. Skylens: Visual analysis of skyline on multi-dimensional data. *IEEE Trans. Vis. Comput. Graph.* **2017**, *24*, 246–255. [[CrossRef](#)]
90. Mehrabian, A. Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament. *Curr. Psychol.* **1996**, *14*, 261–292. [[CrossRef](#)]
91. Li, X.; Fu, X.; Deng, G. Preliminary Application of the Abbreviated PAD Emotion Scale to Chinese Undergraduates. *Chin. Ment. Health J.* **2008**, *2008*, 327–329.
92. Divers, E. Theory to practice: Pleasure-Arousal-Dominance (PAD) theory for architectural color design. *Color Res. Appl.* **2023**, *48*, 445–452. [[CrossRef](#)]
93. Kim, J.-H.; Lee, J.-S. Reliability analysis of emotion evaluation EPA. PAD model in each design field. *Sci. Emot. Sensib.* **2014**, *17*, 79–92. [[CrossRef](#)]
94. Yang, K.; Kim, H.M.; Zimmerman, J. Emotional branding on fashion brand websites: Harnessing the Pleasure-Arousal-Dominance (PAD) model. *J. Fash. Mark. Manag. Int. J.* **2020**, *24*, 555–570. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.