Malaysia Architectural Journal

Journal homepage:
https://majournal.my/index.php/maj
e-ISSN: 2716-6139

Study on Visual Perception Patterns of Public Space in Urban High-Density Residential Areas: The Example of Xi’an

Ziheng Wang [1], Li Zhang [1*]

[1] School of Architecture, Tsinghua University, China

ARTICLE INFO

Article history:
Received: 30 April 2024
Received in revised form
Accepted: 15 October 2024
Available online: 23 June 2025

Keywords:

High-Density Residential Areas; Semantic Segmentation; Eye-Tracking

ABSTRACT

In response to the accelerated process of urbanisation, the development of high-density residential areas has emerged as an effective strategy to meet the increasing demand for housing. However, the limited public space available in these areas frequently fails to provide adequate separation between public and private domains. The proximity between public and private spaces often results in residents' sightlines intersecting with private areas during public activities. Such spatial arrangements compromise both personal privacy and the quality of public activities. This study examines the relationship between the proportions of spatial elements within the human field of view and the duration of gaze across a range of settings. The objective is to analyse visual attention patterns to inform strategies to enhance the attractiveness and engagement of public spaces.

In contrast to conventional design approaches that rely on subjective evaluations, the measurement of human physiological data provides a more objective means of quantifying spatial experiences. Given the crucial role of visual information in spatial perception, eye-tracking technology captures and quantifies attention patterns, enabling designers to provide precise and actionable recommendations for improving spatial arrangements.

This study employed a 360-degree panoramic camera to capture 120 photographs of public activity areas across 20 residential communities in Xi'an, Shaanxi. Semantic segmentation was used to extract information from the images. Furthermore, 50 participants observed the scenes with their eye-tracking data recorded. By correlating spatial data with eye-tracking metrics, this study investigates the relationship between spatial elements and visual attention in residential public areas, revealing perception patterns across different spaces.

[*] *Corresponding author.*
E-mail address: brianchang@mail.tsinghua.edu.cn

1. Introduction

1.1 Background

As urbanization continues to advance, urban development in China has made significant progress. Since the 1980s, advancements in construction technology have led to significant transformations in the design and structure of residential areas. High-density residential zones have become a crucial strategy to address mounting housing needs. While the configurations of these high-density residential areas exhibit variability, they are unified by a core principle: guaranteeing that essential performance standards, such as fire safety and adequate sunlight, are met. Furthermore, these complexes integrate well-designed public activity spaces with the objective of enhancing living quality and reinforcing community vitality.

After meeting basic residential needs, contemporary residential area design increasingly emphasizes human-centered considerations, aiming to foster social interaction and elevate quality of life. The conventional approach to design research tends to prioritize the physical dimensions of space, frequently overlooking the psychological and behavioral responses of users. Designers often face challenges in creating public spaces that balance functional requirements with user engagement. A comprehensive understanding of how individuals perceive and interact with public spaces is essential for advancing design quality [1]. In comparison to traditional design foundations that rely primarily on subjective evaluations, the measurement of physiological data on the human body can provide a more objective analysis of an individual's experience within a space.

1.2 Spatial Elements and Visual Attention

Eye tracking has been shown to effectively quantify visual attention in design science. Attention quantification tools are widely used in various urban public spaces. In the process of spatial perception, visual information predominates [2][3]. Eye-tracking data quantify residents' attention patterns in public spaces, enabling designers to optimize spaces from a user-centered perspective and provide practical recommendations [4]. This technique records the trajectory of eye movement, dwell time, and pupil dilation. These metrics reveal points of interest and the distribution of attention as individuals view different content. Researchers, including Suarez [5] and Cottet [6], have found that elements or scenes that sustain longer gaze durations are frequently linked to higher aesthetic value assessments. As a tool for measuring physiological indicators, eye-tracking technology offers concrete evidence of visual attention distribution as individuals observe their surroundings [7].

In terms of the types of spatial elements and visual perception, Li [8] discovered that urban imagery featuring trees and individuals garners higher ratings for a space. Ghulam [9] noted that artificial constructs in parks captivate over half of all visual attention, a finding echoed by Amati [10] who reported that artificial constructs in parks are deemed more attractive than natural settings. Zhu [11] provided a synthesis of the visual allure of various elements through a case study of trails in wetland parks. Additionally, Chen [12] conducted a cross-analysis combining semantic segmentation with eye-tracking technology to explore how spatial elements of different information densities influence street space design. The study revealed that the correlation between the semantic proportion of spatial elements and gaze duration varies significantly across distinct spatial types.

1.3 Research Questions and Objectives

There is considerable scope for further investigation into the embodied experiences of individuals residing in high-density urban environments. It is notable that studies which accurately measure and interpret the effects of spatial semantics on user behaviour and psychological responses are scarce, particularly about the distribution of visual attention. This study aims to discuss the impact of the semantic proportion of spatial elements on the duration of visual fixation on specific elements. It examines the correlation between spatial semantics and visual attention in public spaces within residential areas, proposing two central research questions.

The first question arises from the hypothesis that individuals allocate more attention to larger areas in an image. This investigation aims to ascertain whether there is a positive correlation between the duration of gaze on spatial elements in public spaces and the proportion of the image occupied by these elements. The second question is derived from the hypothesis that even when an element occupies a significant portion of an image, it may not correspond to an expected increase in gaze duration. This suggests that other elements in the space can influence the visual attention duration on specific elements.

2. Methodology

An experimental platform was established utilizing 120 residential area scenes, with participants recruited to freely observe these scenes. Semantic segmentation was used to quantify spatial elements, and eye-tracking technology measured gaze duration. The cross-analysis of semantic data and eye-tracking data revealed the relationship between gaze duration on buildings and the semantic composition of the scenes. Correlation analysis was used to verify the first hypothesis, while principal component analysis tested the second hypothesis regarding the influence of spatial elements on visual attention (Figure 1).

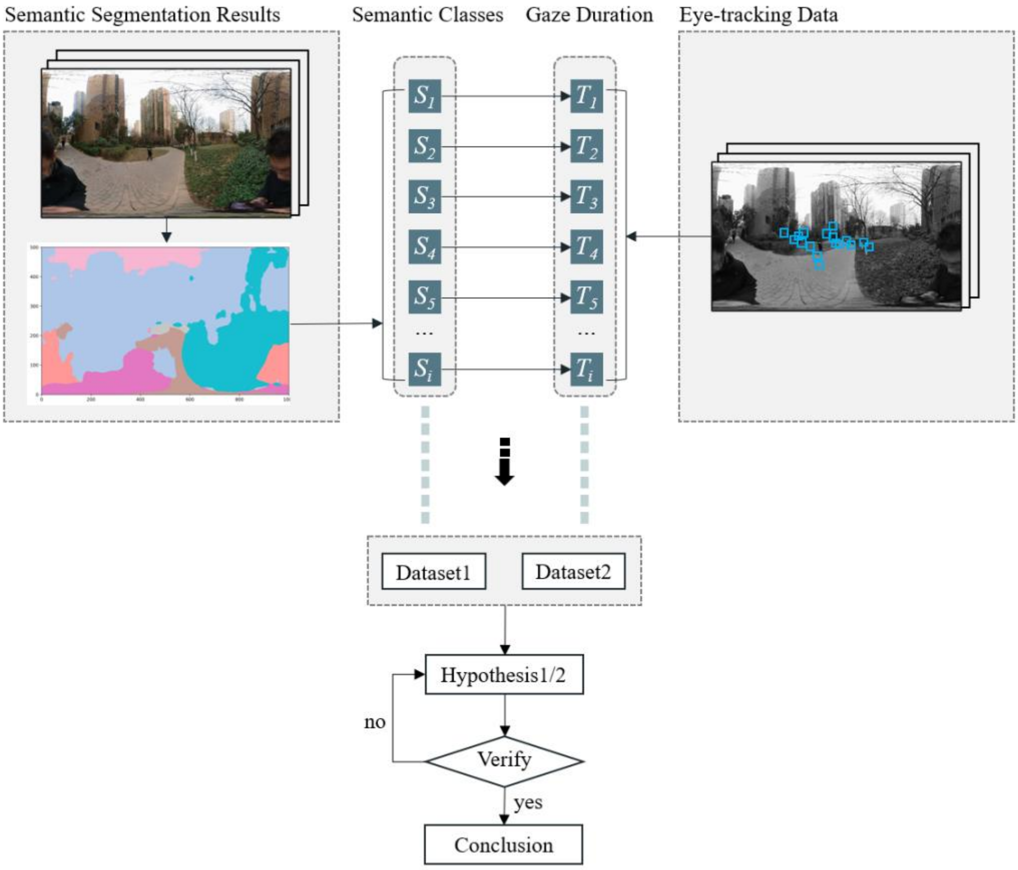


Fig. 1. Overview of research process

2.1 Collection and Categorization of Spatial Data

The study focused on 20 high-density residential areas in Xi'an, Shaanxi, encompassing both new developments and well-established neighborhoods that are widely regarded as being of public interest. To obtain a comprehensive understanding of the characteristics of each scenario, 120 typical scenes from three types of public space within these communities were meticulously documented over a three-day period using a 360-degree panoramic camera. As shown in Figure 2, the scenes are categorized into 41 main thoroughfare spaces, 43 inter-building and forecourt spaces, and 36 public activity spaces. The systematic collection of these data provides a robust foundation for the analysis of utilization patterns of public spaces and interactions among residents in residential areas.

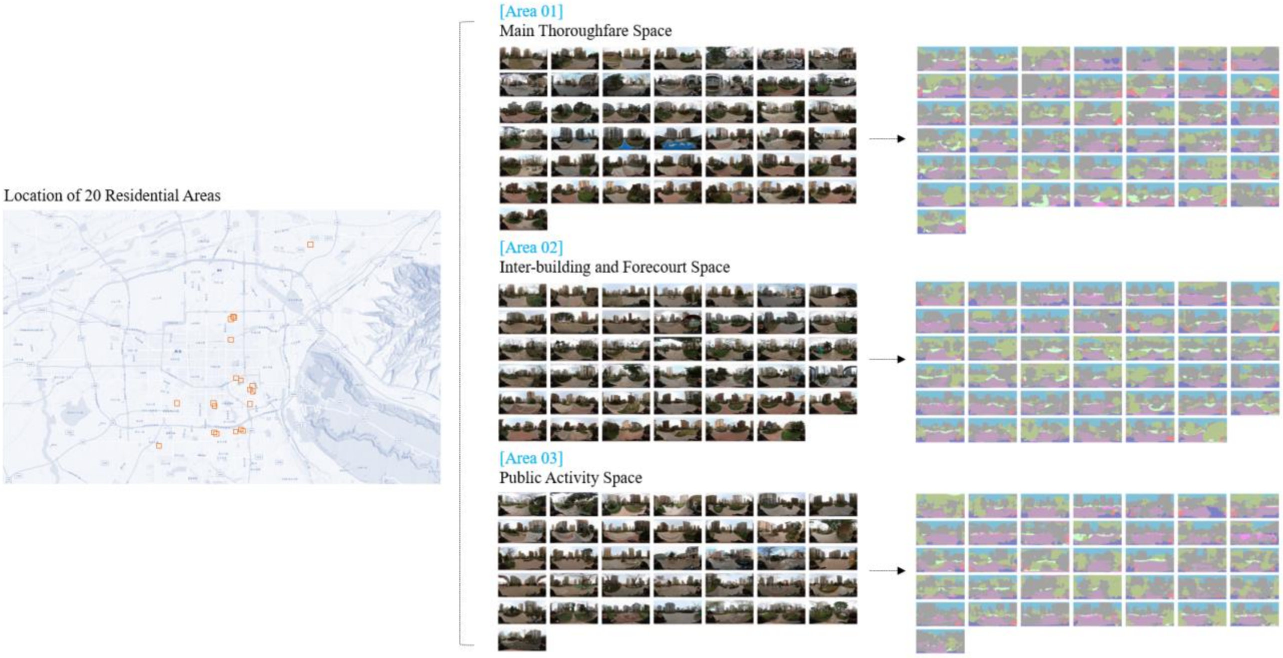


Fig. 2. Classification of 120 residential area scenes

Main thoroughfare spaces are areas located along the primary pedestrian routes within a residential community, typically bordered by the facades of surrounding buildings. Inter-building and forecourt spaces refer to the areas in front of residential buildings, which are usually enclosed on one side by other buildings, vegetation, or fences. Public activity spaces describe open areas specifically designated for various resident activities. Figure 3 illustrates these three typical space types in a high-density residential area.

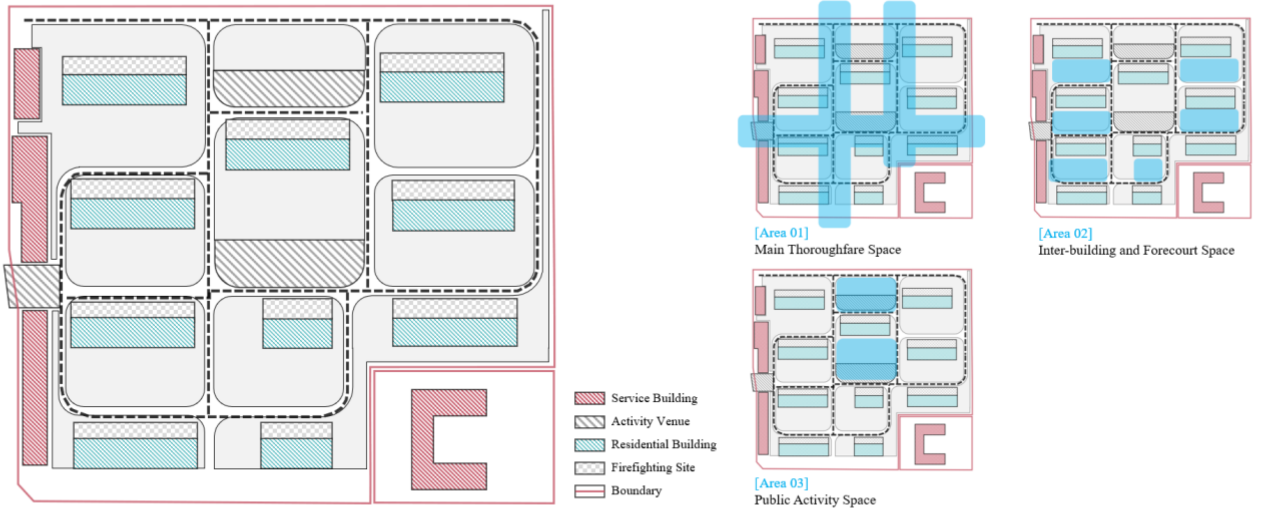


Fig. 3. Diagram of three typical spaces in high-density residential area

2.2 Semantic Segmentation of Spatial Elements

The 360-degree images were converted to equirectangular projection images with a resolution of 2048×4096 pixels to ensure consistency and precision during analysis. This projection method minimizes distortion in central areas of the image, which aligns with participants' primary gaze direction. The equirectangular projection map is a method of representing spherical images on a plane. It preserves accuracy near the equator while introducing significant distortion near the poles. This method maps meridians and parallels evenly to straight lines, causing the poles to be stretched into the top and bottom edges of the image. During experiments, participants' gaze typically focuses on the equatorial region, thereby minimising the impact of polar distortions. Consequently, this projection ensures reliable data with high central image accuracy.

Semantic segmentation technology was employed to extract key semantic information from projection-transformed images of public spaces in residential areas. A semantic segmentation model was trained on the Cityscapes dataset. The Cityscapes dataset was designed specifically for urban street scenes and provides high-quality pixel-level annotations. This dataset enables the segmentation of images into 19 distinct semantic categories, including road, sidewalk, building, wall, fence, pole, traffic light, traffic sign, vegetation, terrain, sky, person, rider, car, truck, bus, train, motorcycle, and bicycle [13]. The semantic categories in question encompass all relevant elements within public spaces of residential areas. By defining spatial elements within the scenes, this process facilitated efficient quantification of spatial scenarios. The segmentation model utilized in this study is based on the DeepLabv3 with ResNet-101 backbone, which has demonstrated excellent performance in various scene parsing tasks [14].

2.3 Eye-tracking Experiment Procedure

The experimental procedure was divided into four main phases: pre-experiment, preliminary calibration, core experiment, and post-experiment evaluation. The post-experiment phase collected subjective feedback from participants.

A total of 50 participants, aged between 20 and 35 years (mean age 25.9, standard deviation 1.66), were recruited online from a diverse range of educational backgrounds. The group included 26 males and 24 females, all with normal or corrected-to-normal vision.

During the pre-experiment preparation phase, each participant's sitting posture was adjusted to align their face orientation with the primary direction of the VR scene capture.

In the preliminary calibration phase, the head-mounted display was calibrated, and the eye-tracking system was synchronized using images analogous to those employed in the main experiment, thus ensuring the veracity of the recorded eye-tracking data.

The core experimental phase involved participants viewing 120 spatial scenes over two separate sessions. Eye-tracking data were recorded with great precision using a head-mounted eye-tracking device (HTC VIVE PRO EYE), with 30 coordinates captured per second. Each scene was presented for a duration of 10 seconds, with a 4-second interval between scenes (Figure 4). The participants were instructed to view the scenes freely, without the imposition of any specific observational goals. The 360-degree panoramic format allowed head movement, but excessive turning was discouraged to prevent neck strain. Participants viewed 60 scenes per session, with each session lasting about 17 minutes. To prevent visual fatigue, a black screen was introduced after every six scenes to provide rest. If participants needed a break or wanted to remove the VR headset, they were allowed to do so, but recalibration was required before resuming.

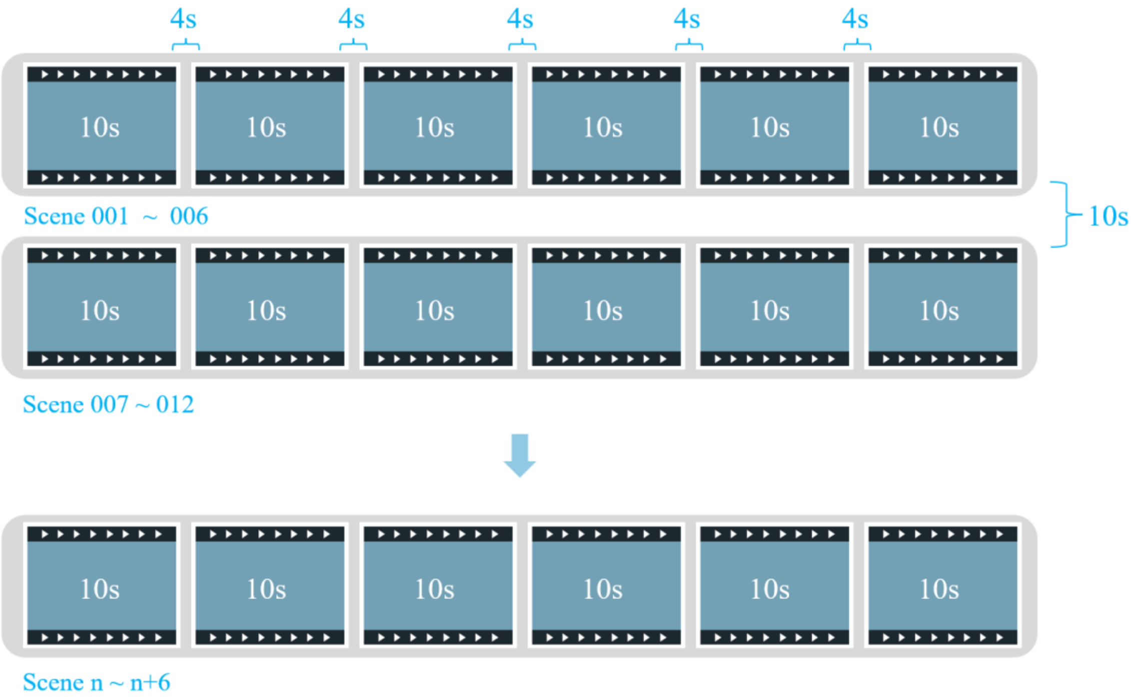


Fig. 4. Scene playback duration and interval times

The order of the scenes was randomized to avoid order effects. Furthermore, the experimental setup was verified and recalibrated between sessions to ensure data consistency and accuracy. The combination of naturalistic observation instructions, controlled viewing times and periodic rest intervals was designed to create a comfortable and immersive environment for participants, thereby capturing genuine eye-tracking behaviors within residential spaces.

In the post-experiment phase, participants completed a questionnaire to provide detailed feedback on their experience. The survey focused on identifying the most engaging spatial elements, participants' preferences for certain areas, and subjective opinions on the overall experience and specific scene preferences. This feedback was essential for understanding how spatial configurations affect user engagement and satisfaction.

2.4 Variable Definition

In this study, the independent variable is the semantic proportion of different categories of spatial elements, such as buildings, sky, roads, terrain, and vegetation. The dependent variable is the duration of time spent focusing on these categories of spatial elements during the observation of various spatial scenes.

3. Results

3.1 Dataset Construction

3.1.1 Semantic Segmentation Data of Scenes

Following projection conversion, each of the 120 typical spatial scenes was converted into an image with a resolution of 2048×4096 pixels. Each image contains 8,388,608 pixels, with the origin of the coordinate system at the bottom-left corner of the image. Semantic segmentation is employed to assign a specific class to each coordinate.

3.1.2 Participants' Eye-tracking Data

Participants' observations were systematically recorded as eye-tracking data sets, typically generating about 300 data points per scene. After collecting this raw data, the spherical coordinates were transformed into planar coordinates within a data frame that maintained the 2048×4096 resolution ratio to ensure consistency and accuracy across various measurements and align with the semantic data from the scenes.

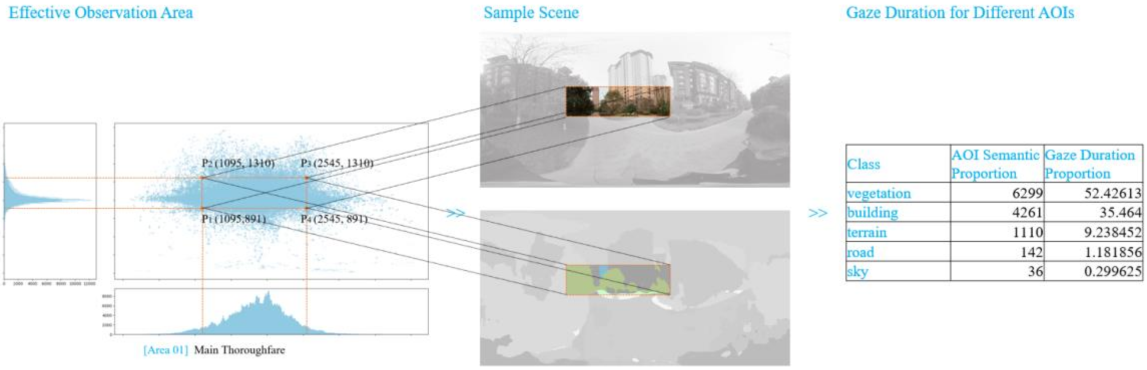
The dataset comprised eye-tracking data collected from three types of residential public space, as detailed in Table 1. To ensure the integrity of the data, any measurements potentially affected by eye blinking, closing or tearing were excluded, as these could be misinterpreted as fixed gaze due to consistent coordinates over time. Following the initial data cleaning process, a total of 1,611,573 eye-tracking data points remained for in-depth analysis.

Table 1
Inventory of scenes and eye-tracking data across different types of three residential public space

Public Space Type	Number of Scenes	Number of Eye-tracking Data Sets	Number of Eye-tracking Coordinates
Inter-building and Forecourt Space	43	2,124	576,934
Main Thoroughfare Space	41	2,029	552,530
Public Activity Space	36	1,777	482,109
Sum	120	5,930	1,611,573

3.2 Data Preprocessing and Preparation

To ensure the quality and representativeness of the data, observations within the 95% confidence interval on each axis were selected, defining the valid observation areas (Figure 5). This reduced the influence of outliers and focused the analysis on predominant gaze patterns.


Effective Observation Area
Sample Scene
Gaze Duration for Different AOIs
P2 (1095, 1310)
P3 (2545, 1310)
P1 (1095,891)
P4 (2545, 891)
>>
>>
[Area 01] Main Thoroughfare

Class	AOI Semantic Proportion	Gaze Duration Proportion
vegetation	6299	52.42613
building	4261	35.464
terrain	1110	9.238452
road	142	1.181856
sky	36	0.299625

Fig. 5. Definition of effective observation areas

3.3 Gaze Fixation Patterns

The results of the statistical analysis indicated that within the effective observation area, the horizontal range of gaze exhibited by participants was significantly broader than the vertical range. Furthermore, the most frequently observed points were situated slightly above the participants' horizontal eye level. Participants' attention within the effective observation area focused primarily on five spatial elements: buildings, vegetation, terrain, roads, and sky. These elements collectively accounted for over 97% of the observed areas (Figure 6), thereby underscoring the predominant semantic components of public spaces within residential settings.

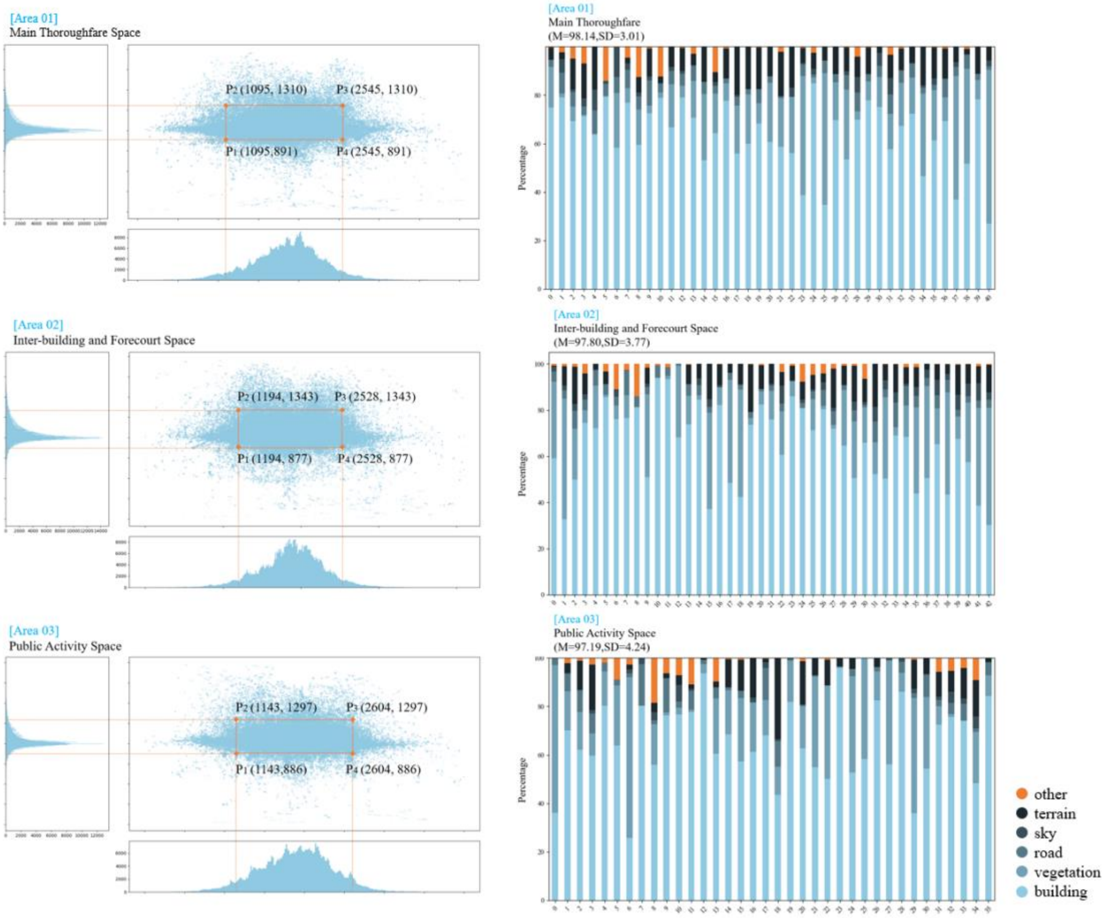


Fig. 6. Primary focus areas and proportion of spatial elements

3.4 Statistical Analysis

3.4.1 Univariate Analysis of Relationships

This section tests the hypothesis that individuals tend to allocate more attention to larger areas within an image. The analysis specifically examines whether the proportion of the area covered by various semantic elements within public spaces of three types of residential areas correlates with the duration of observers' gaze on these elements. The proportion of the area serves as the independent variable, while the gaze duration acts as the dependent variable. In this section, the semantic proportion of spatial elements is sequentially selected as the independent variable, and the gaze duration on these spatial elements is paired as the dependent variable for correlation analysis.

To quantify the relationship between the proportion of spatial elements and the duration of gaze on these elements, Pearson's correlation coefficient was used to analyse the relationship between the independent and dependent variables. This statistical method was selected because it effectively measures the strength and direction of a linear relationship between two continuous variables. Pearson's correlation was utilised to ascertain whether a larger visual area captures more attention, as hypothesised. The results are detailed in Table 2.

Table 2
Correlation of proportions of spatial elements with gaze duration in three spaces

Area Type	Elements	Correlation	P-value	Slope
Inter-building and Forecourt Space	building	0.758	3.93E-09	0.66
	road	0.467	0.002	0.84
	sky	0.627	6.80E-06	0.09
	terrain	0.571	6.45E-05	0.24
	vegetation	0.882	4.12E-14	1.04
Main Thoroughfare Space	building	0.586	5.58E-05	0.70
	road	0.351	0.024	0.45
	sky	0.045	0.775	0.00
	terrain	0.814	8.25E-11	1.91
	vegetation	0.818	6.50E-11	0.72
Public Activity Space	building	0.643	2.26E-05	0.69
	road	0.727	5.00E-07	0.67
	sky	0.571	2.73E-03	0.06
	terrain	0.881	1.36E-12	1.48
	vegetation	0.817	1.19 E-9	0.59

The correlation coefficients obtained were employed to ascertain which spatial elements exert the most significant influence on gaze duration in different public spaces. Generally, a correlation coefficient above 0.7 indicates a strong association between the variables.

Analysis of the data reveals discrepancies with the initial hypothesis, which posited that an increase in semantic elements would lead to a corresponding increase in gaze duration. However, observations across three types of public spaces show that specific elements within each space do not consistently exhibit strong correlations with gaze duration.

The results of the correlation analysis indicated that only two to three elements within each spatial category exhibited a significant relationship between gaze duration and their proportional presence in the scene. Furthermore, the majority of these elements exhibited slopes that significantly deviated from the value of "1." This finding indicates that the duration of attention allocated to specific elements within a scene does not increase in a consistent linear manner with the increase in their spatial proportion. In other words, although the hypothesis suggested that gaze duration would increase linearly with a uniform slope as the area proportion increases, the research results reveal a deviation from this pattern, implying that other potential variables might play a more critical role in influencing visual attention.

The results of this study refute the first hypothesis, demonstrating that an increase in the proportion of semantic elements does not necessarily lead to an increase in gaze duration. The gaze duration on specific spatial elements may be associated with the proportions of other elements within the scene. Therefore, subsequent research will explore the second hypothesis: even when a spatial element occupies a significant portion of an image, it may not result in the expected increase in gaze duration, indicating that other elements might influence visual attention duration.

3.4.2 Multivariate Principal Component Analysis

PCA was used to further investigate how the semantic proportions of spatial elements influence gaze duration. PCA simplifies the data structure by reducing multiple related variables into a few uncorrelated principal components, minimizing multicollinearity and improving the model's interpretability. In this analysis, the dependent variable is "gaze duration on architectural elements," which partially reflects the time individuals spend not observing other elements of the public space.

In the inter-building and forecourt space, the model demonstrates a lack of explanatory power for the dependent variable, indicating significant errors in predicting building gaze duration. As a result, this space's data has been excluded from further analysis. Conversely, in the main thoroughfare space, the model exhibits moderate explanatory power, allowing for partial interpretation of changes in building gaze duration. In the public activity space, the model shows strong explanatory power, accurately predicting building gaze duration.

The detailed Principal Component Analysis (PCA) loadings are presented in Table 3, while the Principal Component Regression (PCR) results are provided in Table 4.

Table 3
Principal component loadings and explained variance

Area Type	PC	Building	Vegetation	Terrain	Road	Sky	Explained Variance
Main	PC 1	-0.534	0.543	0.341	-0.492	-0.246	49.92%
Thoroughfare	PC 2	-0.401	0.403	-0.519	0.220	0.598	26.39%
Space	PC 3	0.085	0.248	-0.652	0.083	-0.706	13.39%
	PC 4	0.365	0.365	-0.355	-0.818	0.263	9.7%
Public	PC 1	-0.522	0.634	0.037	-0.480	-0.306	44.83%
Activity	PC 2	0.490	0.007	0.679	0.307	0.452	25.89%
Space	PC 3	0.200	-0.319	0.653	-0.189	-0.629	17.98%
	PC 4	0.268	-0.104	0.175	-0.765	0.549	11.26%

Table 4
Principal component regression

Area Type	PC1	PC2	PC3	PC4
Main Thoroughfare Space	-5.109	-2.761	1.189	4.907
Public Activity Space	-3.026	-10.602	1.716	0.549

In the main thoroughfare space, the increase in the proportion of buildings significantly enhances the duration of attention to the buildings, indicating a strong visual appeal in this environment. By contrast, terrain had a relatively limited negative impact on attention duration, with its influence being weaker than that of vegetation. The effects of the sky and road on attention duration are minor, with their overall impact being less significant. Similarly, in the Public Activity Space, the increase in the proportion of buildings enhances attention duration, demonstrating a dominant visual effect in this context. However, terrain exerts the most significant negative impact on attention duration, making it the strongest negative factor in this space, surpassing the influence of vegetation. The sky also has a relatively significant negative impact, second only to terrain, while the road exerts a comparatively weak effect on attention duration.

3.5 Participants' Feedback and Perceptions

Upon completion of the experiment, participants were invited to rate the attractiveness of various scene elements on a scale of 1 to 5, with 1 indicating no attraction and 5 signifying extreme attraction. The ratings are presented in detail in Figure 7. It is noteworthy that the categorization of scene elements for subjective ratings was more detailed than the semantic segmentation used earlier in the study. New categories such as signboards, people/animals, and artificial structures were introduced, and buildings were subdivided into residential and other types.

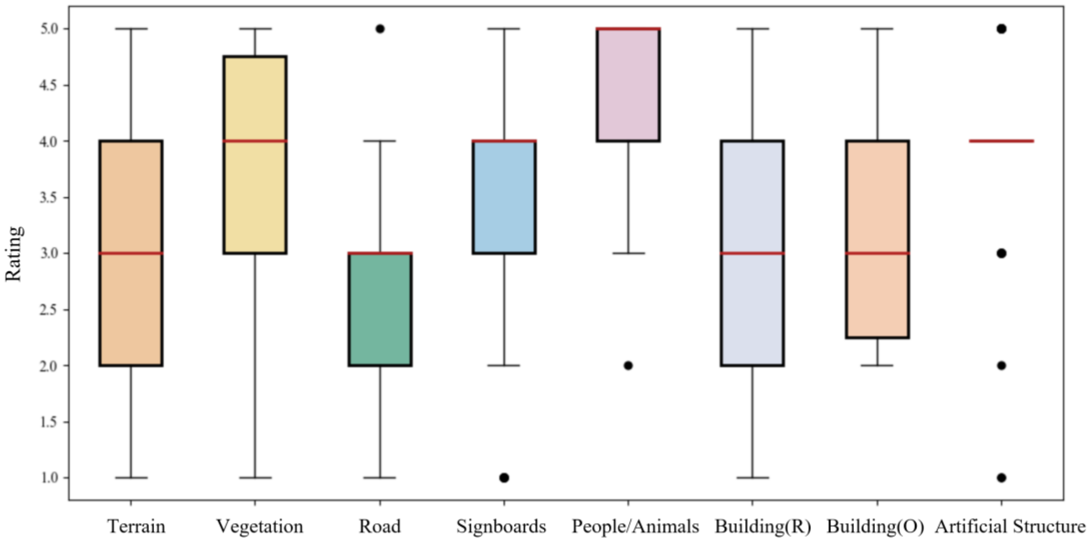


Fig. 7. Participants' ratings of attractiveness for different spatial elements

Participants expressed a strong preference for green spaces, especially tall trees, flowers, and the interplay of light and shadow, which they found visually appealing. Open green belts and carefully designed residential greenery also received high praise. Signboards were noted to ignite curiosity and a desire to explore, enhancing visual attraction, although they did not significantly extend the duration of attention. Architectural features like doorways, windows, and unique architectural forms, along with the spacing between buildings, were highly appealing. The integration of green space between buildings and optimal natural lighting conditions was also highlighted as crucial.

The subjective feedback provides detailed insights into the participants' perceptions and preferences, which complement the quantitative data analysis. While the data analysis highlights the considerable impact of green spaces and vegetation on gaze duration, subjective descriptions indicate that lighting conditions also influence the perception of these elements. The introduction of more detailed categories, such as signboards and specific building types, further elucidates elements that attract attention without necessarily increasing gaze duration. This integration of subjective and objective analyses provides a comprehensive understanding of the complex ways in which spatial design influences visual attention and participant experience.

4. Discussion

This study offers insights into the influence of spatial elements on visual attention within high-density residential public spaces. The findings suggest that visual attention is shaped by the interplay of multiple factors.

While larger elements like vegetation and buildings generally attracted more gaze, the relationship between element size and attention was not consistent across different space types. For example, the sky, despite occupying a large portion of the visual field, did not consistently lead to longer gaze durations. This indicates that sheer size alone is not the sole determinant of visual engagement.

PCA showed that factors influencing gaze duration on buildings varied across different public spaces. In main thoroughfare spaces, vegetation consistently attracted more attention, whereas in public activity spaces, terrain was the dominant factor. This highlights the importance of considering not just individual elements, but their collective interaction when designing public spaces.

It is important to note that the study has several limitations that warrant careful consideration. First, the semantic segmentation models used in this study may lack sufficient granularity for categories like buildings or vegetation, potentially affecting the precision of element identification and analysis. Second, the sample's overrepresentation of younger participants may limit its overall representativeness. Incorporating a more diverse age distribution might enhance the sample's representativeness and, subsequently, the generalisability of the study's conclusions.

5. Conclusion

In high-density residential areas, the ability of different spatial elements to capture residents' visual attention varies significantly. This study emphasizes the importance of carefully adjusting the proportion of these elements in the field of view when designing public spaces. By concentrating on enhancing elements, such as vegetation in main thoroughfare spaces and terrain in public activity spaces, the design can more effectively engage residents' attention, thereby contributing to a more enriching spatial experience.

Acknowledgement
This research was supported by the National Natural Science Foundation of China (No. 52278023), China National Key R&D Program (No. 2022YFC3801500) and Cyrus Tang Foundation.

References

[1] Zhang L., Deng H., Mei X., Pang L., Xie Q., Ye Y. (2022). Urban Ergonomics: A design science on spatial experience quality. Science Bulletin, 16, 1744-1756.

[2] Itti, L. (2000). Models of bottom-up and top-down visual attention. Dissertation Abstracts International, Volume: 61-05, Section: B, page: 2406; Adviser: Christof Koch.

[3] Koch, K., Mclean, J., Segev, R., Freed, M. A., Ii, M. J. B., & Balasubramanian, V., et al. (2006). How much the eye tells the brain. Current Biology, 16(14), 1428-1434.

[4] Zhang L., Xie Q., Deng H., Mei X., Pang L., Ye Y. (2022). An Ergonomic Analysis Approach for Spatial Experience Proof. World Architecture, 09, 42-47. https://doi.org/10.16414/j.wa.2022.09.019

[5] Suarez, L. A. D. L. F. (2020). Subjective experience and visual attention to a historic building: a real-world eyetracking study. Frontiers of Architectural Research, 2020, 9(4), 31.

[6] Cottet, M., Vaudor, L., Tronchere, H., Roux-Michollet, D., Augendre, M., & Brault, V. (2018). Using gaze behavior to gain insights into the impacts of naturalness on city dwellers' perceptions and valuation of a landscape. Journal of Environmental Psychology, 60(Dec.), 9-20.

[7] Vainio, T., Karppi, I., Jokinen, A., & Leino, H. (2019). Towards Novel Urban Planning Methods—Using Eye-tracking Systems to Understand Human Attention in Urban Environments. Human Factors in Computing Systems. ACM.

[8] Li, J., Zhang, Z., Jing, F., Gao, J., Ma, J., Shao, G., Noel, S. (2020). An evaluation of urban green space in Shanghai, China, using eye tracking. Urban Forestry & Urban Greening, 56(1).

[9] Gholami, Y., Taghvaei, S. H., Norouzian-Maleki, S., & Sepehr, R. M. Identifying the stimulus of visual perception based on eye-tracking in urban parks: case study of Mellat Park in Tehran. Journal of Forest Research.

[10] Amati, M., Sita, J., Parmehr, E., & McCarthy, C. (2018). How eye-catching are natural features when walking through a park? Eye-tracking responses to videos of walks. Urban Forestry & Urban Greening, 67-78.

[11] Zhu, X., Zhang, Y., & Zhao, W. (2020). Differences in environmental information acquisition from urban green—a case study of Qunli National Wetland Park in Harbin, China. Sustainability, 12(19), 8128. https://doi.org/10.3390/su12198128

[12] Chen Z. (2023). Positive Nudges for Urban Regeneration via Eye-Tracking and Behavioural Evidence. World Architecture, 07, 68-69. https://doi.org/10.16414/j.wa.2023.07.006

[13] Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., ... & Schiele, B. (2016). The cityscapes dataset for semantic urban scene understanding. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 3213-3223).

[14] Chen, L. C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2017). Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. IEEE transactions on pattern analysis and machine intelligence, 40(4), 834-848.