Relationship between Resolution and Spatial Awareness in Virtual Reality Space

Target historical buildings archived with 3D scan

Joonyoung Kim[1] and Satoshi Yamada[2]
[1,2]*Ritsumeikan University*
[1]*ru0102if@ed.rutsumei.ac.jp, ORCID: 0009-0006-6109-0320*
[2]*sy@fc.ritsumei.ac.jp, ORCID: 0009-0009-3653-7196*

Abstract. This study examines the impact of spatial resolution on spatial cognition within virtual reality (VR) environments, focusing on Japanese traditional wooden and modern reinforced concrete (RC) architectures. Advances in 3D scanning and VR technologies increasingly facilitate the digital archiving of cultural heritage; however, resolution challenges affect architectural-detail representation. This study aims to identify cognitive thresholds in spatial resolution—defined through polygon and texture pixel counts—to better understand the perception of architectural elements. Participants viewed VR models of two buildings, a traditional wooden and a modern RC building, across varied polygon and pixel resolutions using head-mounted displays. Both continuous and segmented models allowed for controlled resolution adjustments. The analysis incorporated gaze tracking and questionnaires to evaluate spatial quality and characteristics. Results indicate that texture resolution is universally important across architectural styles, particularly in traditional wooden buildings, while polygon resolution thresholds primarily impact modern RC structures. Understanding spatial-resolution thresholds in VR can enhance digital heritage preservation, supporting both visual fidelity and cognitive engagement across diverse architectural forms. Future research will expand these findings to additional architectural styles and user demographics.

Keywords. Digital Heritage, VR, Resolution, Spatial Awareness, 3D Archive, Historic Buildings

1. Introduction

1.1. UTILISATION OF DIGITAL HERITAGE

In recent years, digital archives using 3D scanning data of spaces have been produced actively. Advances in technologies that support the improvement and popularisation of visual reproducibility have been remarkable. The production of digital heritage using these technologies is progressing. A well-known example is the use of these

ARCHITECTURAL INFORMATICS, Proceedings of the 30th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2025, Volume 4, 295-304. © 2025 and published by the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA), Hong Kong.

technologies in Notre-Dame Cathedral, which became a global topic after its destruction by fire. The use of digital heritage to preserve and educate visitors about cultural properties is also expanding in museums since it effectively enhances their understanding and interest in cultural heritage. Exhibits combining small-scale real objects with VR content that can be experienced in actual size are becoming popular (Rushton and Schnabel, 2020). Initially, VR content was primarily viewed using dedicated Head Mounted Displays(HMD); however, more viewing methods using augmented reality (AR) apps on gyroscope sensors of smartphones are emerging (Dezen-Kempter et al., 2020). High-precision web-based viewing is also possible, as exemplified by Shoei Yoh (Gardner et al., 2022). In addition to visual-spatial preservation, it is also effective in recording and passing on culture and customs. A digital archive of traditional funeral ceremonies, for example, can contribute to the social recording and transmission of culture and customs (Pillaca and Escobar, 2023).

1.2. RELATIONSHIP BETWEEN SPATIAL RESOLUTION AND SPATIAL COGNITION: TWO OBJECTIVES

Along with technological advances and practical efforts, research on spatial cognition is also crucial. For example, the relationship between spatial resolution (hereafter referred to as spatial resolution) and spatial recognition, which this study focuses on, is important. If the spatial resolution is too high, perceptible rendering delays may occur in humans. If it is too low, the reproducibility may significantly deteriorate. In practice, when the authors rendered a 3D scan of a building with a volume of approximately 160,000 m^3 on a game engine using an RTX 4070Ti, rendering delays were observed. This research addresses this issue by presenting spatial-resolution threshold values that influence cognitive perception. While advances in storage and playback devices, communication environments, and other technologies are rapid, this knowledge contributes to technological development and creates content as a reference for a target value. This study aims to highlight the importance of the relationship between spatial resolution and spatial cognition and provide insights, contributing to the development of digital heritage from both technological and human science perspectives.

In addition to the threshold values of spatial resolution, this study also focuses on characteristic elements (regions) in spatial cognition. The subject of this study is Japanese traditional wooden architecture and modern reinforced concrete (RC) architecture. The threshold values may differ between the two, and the significance of detail changes caused by adjustments in spatial resolution may also vary. In other words, the importance of the detail level of spatial components (such as pillars and beams) may differ between the two architectural styles. Elements that are easily perceived or observed by people during the process of reducing (or increasing) design detail can indicate the characteristic elements of the architecture or style. This study aims to understand the characteristic elements of traditional wooden architecture and modern RC architecture using the visually immersive experience made possible by advanced spatial-resolution technology.

2. Positioning of the Study

2.1. PREVIOUS STUDIES ON VR ENVIRONMENT QUALITY AND USER EXPERIENCE

Zimmons and Panter (2003) analysed the quality of textures and lighting in VR environments and examined how this quality impacts sense of presence, task performance, and object memory of the participants. They measured physiological responses and task performance accuracy, finding that physiological responses (heart rate) increased under all rendering conditions. However, rendering quality did not produce significant differences in task performance accuracy or object memory. Hvass et al. (2017) investigated polygon count and texture resolution in VR games, analysing their impact on presence. They compared high- and low-resolution conditions and found that a higher polygon count and texture resolution enhanced presence. However, because both variables were changed simultaneously, it was challenging to assess individual effects. Volkmann et al. (2020) explored the impact of polygon count on physical and self-presence in virtual environments, hypothesising that higher polygon counts improve presence. However, results showed that increasing polygon count did not significantly affect physical or self-presence, indicating that polygon count is not a decisive factor in establishing presence.

These studies address visual realism, task performance, and user perception in VR environments. However, few studies focus on spatial cognition in user experiences within digital archives. Specifically, no research has presented the relationship between spatial resolution and spatial cognition in VR spaces for digital heritage.

2.2. RESEARCH OBJECTIVES

Based on the above background and previous research, this study defines spatial resolution as polygon count and texture pixel count (hereafter referred to as pixel count) and has three research objectives. The first is to determine the threshold spatial resolution for cognitive impact. The second is to use interviews and gaze tracking to determine the spatial components that influence cognition. The third is to compare the results for traditional wooden architecture and modern RC architecture of Japan to identify their differences. These three objectives represent the novelty of this study.

3. Research Overview and Methods

3.1. RESEARCH OVERVIEW

This study involved an experiment where participants experienced virtual buildings in a VR environment using HMDs. The experimental buildings consisted of two architectural styles: traditional wooden architecture and modern RC architecture (hereafter referred to as traditional wood and modern RC, respectively). The base data was prepared using 3D scans of these buildings and photogrammetry technology. Two types of viewing spaces were then created by manipulating the spatial resolution of this base data: continuous and segmented models. The continuous model was designed to identify the spatial resolution that serves as a cognitive threshold, allowing participants to continuously adjust polygon and pixel counts separately—referred to as the polygon continuous model and pixel continuous model. The segmented model was designed to identify the spatial components that influence cognition. It consisted of six viewing spaces with three levels of spatial resolution—referred to as the polygon segmented model and pixel segmented model. A psychological-evaluation questionnaire was completed after each experience, with open-ended feedback gathered at the end. In the analysis, descriptive statistics are first presented. For the segmented model, a two-factor analysis of variance was conducted on the key questions. For the continuous model, participants gaze points were recorded to determine the characteristic spatial components when recognising spatial-resolution changes, enabling the comparison of visibility characteristics of traditional wooden and modern RC of Japan.

3.2. BUILDINGS USED FOR THE EXPERIMENT

The traditional wood building selected for the experiment was the KannonDo of Miidera Temple in Otsu City, Shiga Prefecture; for the modern RC building, the Kyoto University Gymnasium on the Yoshida Campus in Sakyo Ward, Kyoto City, was selected. The photogrammetry was performed using a Canon EDS60D camera, 17-mm lens, and ISO 100 setting processed using reality capture. Approximately 3,000 KannonDo and 7,000 Kyoto University Gymnasium photos were taken.

3.3. POLYGON SEGMENTED MODEL AND CONTINUOUS MODEL

The polygon segmented model was created by processing the base data in Blender. The unit representing the level of detail was defined by the minimum segment length between points (point distance in the point cloud), referred to as polygon resolution. Three levels were set at "5 mm (High), 50 mm (Middle), and 100mm (Low). The 5 mm was based on the maximum resolution of 3D scanners used by multiple companies domestically. The 50 mm and 100 mm referred to the acceptable measurement error range for unmanned aerial vehicle laser scanners (Tanaka, 2020). The polygon continuous model system was implemented in the Unity game engine. The resolution range was set from 5 mm to 125 mm based on the three levels of the polygon segmented model, divided into 25 stages with a step size of 5 mm. The manipulation was performed using keyboard controls. In both cases, the pixel count was set to High for the pixel segmented model. The segmented model and details are shown on the left side of Figure 1.

3.4. PIXEL SEGMENTED MODEL AND CONTINUOUS MODEL

This model was created by adjusting the pixel count of each texture. The unit representing the level of detail was defined by the pixel count per square meter, referred to as pixel resolution. The three levels were set at approximately 50,000 (High), 3,000 (Middle), and 200 (Low). These can be considered as tens of thousands (High), thousands (Middle), and hundreds (Low). In terms of pixel count of each texture, they were 8192×8192 (High), 2048×2048 (Middle), and 512×512 (Low). The segmentation width was set at 1/16 based on the polygon segmentation levels and preliminary experiments. The pixel continuous model system was also implemented in Unity. The pixel range was set from approximately 50,000 (tens of thousands, 8192×8192) to approximately 50 (hundreds, 256×256) based on the segmented model and preliminary experiments. The segmentation width was set as an area ratio of 9/16, divided into 10 levels. The polygon count was set to High for the polygon segmented model. The segmented model and details are shown on the right side of Figure 1.

The experimental system was constructed by integrating the above viewing spaces into the Unity game engine. The HMD used was a VIVE Pro EYE(FoV: 2880 x 1600 pixels), and the graphics card was an RTX 4070Ti. No significant rendering lag was observed. The viewing position was fixed to avoid it being a confounding factor in the analysis. The HMD was synchronised with gaze tracking, and a custom system was implemented to record gaze points. Gaze was defined as remaining within a 3D area of a sphere with a radius of 0.1 m for more than 0.2 s (Penkar et al., 2012).

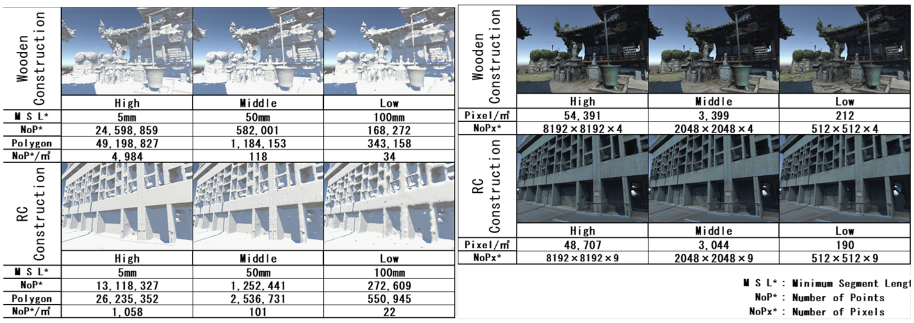


Figure 1. Left: polygon number split model data Right: pixel split model data

3.5. QUESTIONNAIRE ITEMS

For psychological evaluation, 50 items were initially selected based on previous studies (Yi and Ha, 2012; Takeda, 2012). Thereafter, based on preliminary experiment results with seven participants, 17 adjective pairs were selected, with seven items evaluating spatial quality and 10 items evaluating spatial characteristics (Figure 2. Left). A 7-point scale was used for evaluation, with higher scores indicating more positive assessments of spatial quality.

No.	*Characteristic of Space*	*No.*	*Quality of Space*
1	Comfort--Uncomfort	1	Weak--Strong
2	Stable--Unstable	2	Narrow--Wide
3	Not dizzy--Dizzy	3	Natural--Artificial
4	Clean--Dirty	4	Warm--Cool
5	Satisfied--Unsatisfied	5	Faint--Clear
6	Concentrated--Distracted	6	Modest--Fancy
7	Like--Dislike	7	Mild--Irritating
		8	Dynamic--Static
		9	Soft--Hard
		10	Dark--Bright

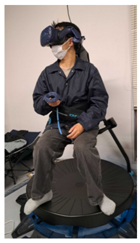


Figure 2. Left: Questionnaire Items Right: A subject undergoing an experiment

3.6. EXPERIMENTAL PROCEDURE

First, participants wore the HMD and viewed unrelated VR content to familiarise themselves with the VR experience while wearing the HMD. After observing the viewing spaces, the participants were instructed to conduct an interview related only to the experience, and the experiment began. The experience order of the segmented model was randomised. The viewing time was set to a minimum of one minute, and the viewing ended when the participant declared their observation complete. Immediately afterwards, VR sickness was checked, and the questionnaire was completed. Breaks of at least one minute were provided. This procedure was repeated 12 times, covering the traditional wood and modern RC architectural styles, polygon and pixel count spatial resolutions, and the three segmentation levels. The experience order of the continuous model was also randomised. Participants were asked to explore the stage where they felt the greatest change and report it, and the slider position at that point was recorded. After viewing, VR sickness was checked, and the reason for the slider position was recorded. This procedure was performed for both architectural styles. The number of valid responses is 30 people in their 20s.

4. Analysis Results

4.1. POLYGON CONTINUOUS MODEL

Figure 3 shows the histogram of the stages where changes were recognised. In the results of the polygon continuous model at the top, no clear high-frequency interval exists in traditional wooden. In modern RC, the 75-80 mm interval shows high frequency.

4.2. PIXEL CONTINUOUS MODEL

In the right section of Figure 3, which depicts the results of the pixel continuous model, there is a tendency for traditional wooden to concentrate around the pixel interval of 1024-819. Modern RC also shows high-frequency intervals, primarily in the pixel intervals of 1024-819 and 512-256. These results present the spatial-resolution thresholds that affect cognition, meeting the first research objective and illustrating the differences between traditional wooden and modern RC, aligning with the second research objective.

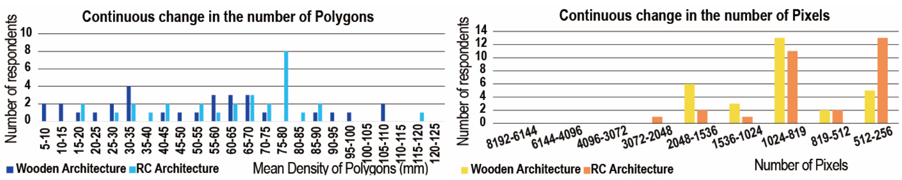


Figure 3. Left: polygon continuous model results Right: pixel continuous model results

4.3. POLYGON SEGMENTED MODEL

Figure 4 shows the average response values for spatial quality, with different line types representing each model. Focusing on the mean differences between models, a general decline in the quality assessment can be observed. However, no clear trends in quality differences between items exist for both traditional wooden and modern RC.

Figure 5 shows the average values for spatial characteristics. In traditional wooden, relatively large cognitive differences can be observed in items such as "Weak-Strong," "Faint-Clear," "Modest-Fancy," "Mild-Irritating," and "Soft-Hard," with particularly clear distinctions in "Faint-Clear," "Modest-Fancy," and "Mild-Irritating." In modern RC, although less pronounced, differences are noted in items like "Weak-Strong," "Narrow-Wide," "Soft-Hard," and "Dark-Bright."

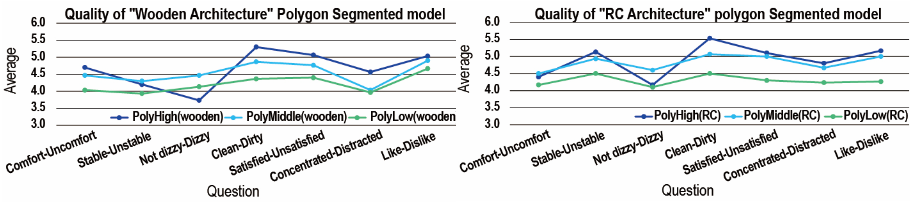


Figure 4. Quality of polygon segmented model results

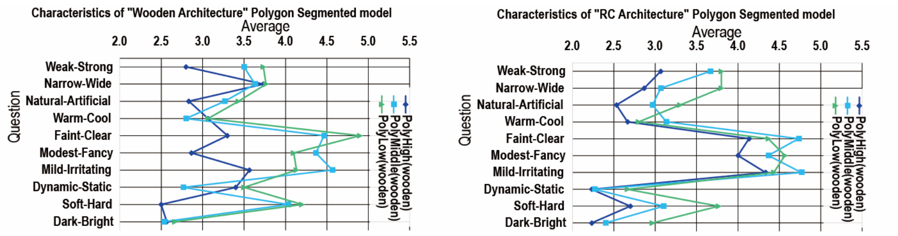


Figure 5. Characteristics of polygon segmented model result

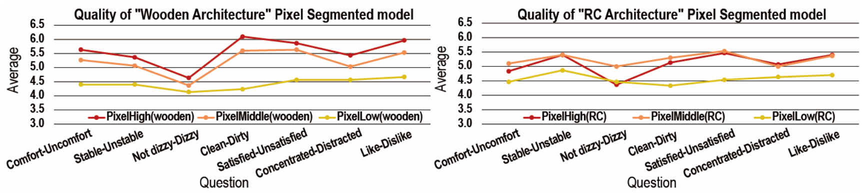


Figure 6. Quality of pixels segmented model results

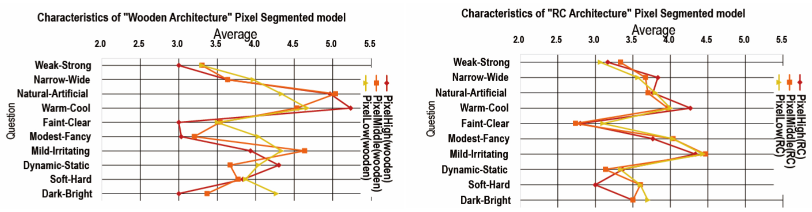


Figure 7. Characteristics of pixel segmented model results

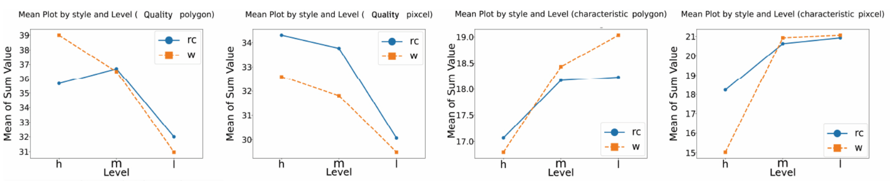
4.4. PIXEL SEGMENTED MODEL

Figure 6 shows the average plot for spatial quality, indicating a generally clearer overall decrease in evaluations than in the polygon segmented model. This trend is apparent in traditional wooden, as lines do not intersect. Looking at trends across items, both traditional wooden and modern RC show relatively large differences in the "Clean-Dirty" and "Satisfied-Unsatisfied" categories. Traditional wooden also displays clearer differences in "Dark-Bright."

Figure 7 presents the average plot for spatial characteristics. There are no significant cognitive differences in either the overall scores or specific items between the "Traditional Wooden" and "Modern RC" categories.

4.5. TWO-WAY ANALYSIS OF VARIANCE

To verify the generality of the analysis, a two-way ANOVA was conducted, with architectural style and resolution as factors (Figure 8). Separate analyses were performed for polygon count and pixel count. Due to correlations among spatial-quality items, the total score across all items was used. For spatial characteristics, the total score of items interpreted as spatial cognition clarity was used. Figure 8 shows the ANOVA table and mean plots. Results affirm significance at the 5% level for the analyses on overall spatial-quality trend and spatial cognition clarity. However, while descriptive statistics suggest that the pixel threshold has a clearer, more comprehensive effect on spatial quality in traditional wooden than in modern RC, the effect size was small, indicating that statistical significance could not be concluded.



variance table Quality polygon

	sum_sq	df	F	PR(>F)
C(style)	89.61	1	1.5039	0.2217
C(Level)	457.78	2	3.8417	0.0233
C(style):C(Level)	16.58	2	0.1391	0.8702
Residual	10367.03	174		

variance table Quality pixcel

	sum_sq	df	F	PR(>F)
C(style)	22.05	1	0.4046	0.5256
C(Level)	1218.81	2	11.1812	2.70E-05
C(style):C(Level)	161.23	2	1.4791	0.2307
Residual	9483.43	174		

variance table characteristic polygon

	sum_sq	df	F	PR(>F)
C(style)	38.27	1	1.4499	0.2302
C(Level)	726.74	2	13.7662	2.82E-06
C(style):C(Level)	116.94	2	2.2152	0.1122
Residual	4592.90	174		

variance table characteristic pixcel

	sum_sq	df	F	PR(>F)
C(style)	3.20	1	0.1244	0.7248
C(Level)	97.38	2	1.8925	0.1538
C(style):C(Level)	8.53	2	0.1658	0.8473
Residual	4476.53	174		

Figure 8. two-way ANOVA graph graph and plot (From left: polygon quality, pixel quality, polygon characteristics, pixel characteristics)

4.6. GAZE AREAS

Figure 9 shows the point cloud for gaze areas. In traditional wooden, elements such as tiles, columns, Buddha statues, and signboards were prominently gazed at, with stone-paved floors also drawing attention in the pixel continuous model. For modern RC, grid-patterned facades were the focus, especially prominent in the polygon continuous model, with particular emphasis on grid edges. In the pixel continuous model, gaze tended to be directed lower, focusing on elements like lower beams and the ground, potentially due to the larger visible concrete surface. The signboard was also noted as a focal point, consistent with interview results.

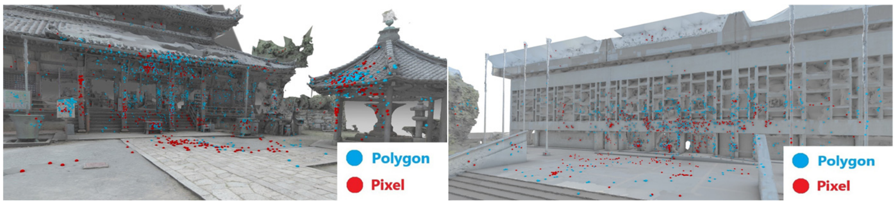


Figure 9. Gaze area. Left: traditional wooden architecture. Right: modern RC architecture

5. Discussion

First, the polygon threshold for modern RC is identified in Section 4.1, which corresponds to the distinction between the Middle and Low levels in the segmented model. Considering the impact of this threshold from the results in Section 4.3, it is suggested that it has a comprehensive influence on the quality of space. Spatial characteristic cognition seems to affect the two items, "Narrow-Wide" and "Dark-Bright," more in modern RC than traditional wooden. Based on the findings in Section 4.5, the spatial components that influence this cognition are likely related to the perception of changes in grid edges, especially in facades. Next, the pixel threshold identified for modern RC in Section 4.2 also corresponds to the Middle–Low difference in the segmented model. From the results in Section 4.4, this threshold appears to have a comprehensive impact on the quality of space. According to Section 4.5, the spatial components influencing this cognition are considered to be the edges of forms and the material properties of building surfaces and the ground. These findings related to modern RC indicate important elements within various levels of spatial detail.

Next, as interpreted in Section 4.2, the clarity of the polygon threshold influencing cognition in traditional wooden is considered low. However, based on the referenceable quantitative reasons indicating a correlation, the previously mentioned five items interpreted as spatial cognition clarity are thought to exhibit more distinct stepwise changes compared to modern RC. The pixel threshold from Section 4.4 is considered to impact the quality of space more clearly and comprehensively in traditional wood than in modern RC." Regarding "spatial characteristic cognition," the results in Section 4.4 indicate that it influences "Dark-Bright" more in traditional wooden than in modern RC. From the results in Section 4.5, the spatial components that influence this cognition are thought to include the form of columns and roofs and the material properties of the ground. These findings related to traditional wooden suggest particularly significant elements within the space.

The understanding of both polygon and pixel thresholds and their cognitive impact aligns with similar results from the study on polygon count by Volkmann et al. (2020). In contrast, Zimmons and Panter (2003) examined the effects of lighting and pixel count on simple tasks and found a minimal impact of pixel count. However, this study identified an impact of pixel count on cognitive perception, likely owing to its focus on spatial cognition.

6. Conclusion

This study's findings on thresholds of spatial resolution, cognition, and gaze provide significant insights for digital archival preservation and device development. Additionally, the observed differences in trends between traditional wooden and modern RC architectures underscore the unique characteristics of each architectural style. Although the novel results of the study hold strong potential, they have limitations. While they may apply to buildings like the ones used for the experiment, caution is needed for cases with low similarity. Future efforts will aim to clarify these trends further and enhance generalisability by including a wider variety of subjects and architectural styles.

Acknowledgment

This work was supported by JSPS KAKENHI Grant Number JP23H01579.

References

Dezen-Kempter, E., Mezencio, D. L., Dematosmiranda, E., Desá, D. P., & Dias, U. (2020). *Towards a digital twin for heritage interpretation: From HBIM to AR visualization*. In Proceedings of CAADRIA 2020, 2, pp. 181–191.

Gardner, N., Haeusler, M. H., Yu, D., Barton, J., Dunn, K., & Huang, T. (2022). *Revisiting SHOEI YOH: Developing a workflow for a browser-based 3D model environment to create an immersive digital archive*. In Proceedings of CAADRIA 2022, Vol. 1, pp. 687–696.

Hvass, J. S., Larsen, O., Vendelbo, K. B., Nilsson, N. C., Nordahl, R., & Serafin, S. (2017). *The effect of geometric realism on presence in a virtual reality game*. In Proceedings of the 2017 IEEE Virtual Reality (VR) (pp. 18–22). IEEE.

Penkar, A. M., Lutteroth, C., & Weber, G. (2012). *Designing for the eye: Design parameters for dwell in gaze interaction*. In Proceedings of the 24th Australian Computer-Human Interaction Conference (pp. 479–488).

Pillaca, G., & Escobar, D. (2023). *El retablo digital (The Digital altarpiece): Restoration of mortuary rituals with deep learning*. In Proceedings of CAADRIA 2023, 1, pp. 757–766.

Rushton, H., & Schnabel, M. A. (2020). *Exhibiting digital heritage: The curation of un-mediated experiences in museums*. In Proc. of CAADRIA 2020, 2, pp. 193–202.

Takeda, N. (2012). *Impression evaluation of indoor visual environment elements*. Doctoral dissertation, Nagoya Institute of Technology.

Tanaka, S., Imai, R., Nakamura, K., & Kawano, K. (2011). *A study on automatic generation of 3D models using point cloud coordinate data*. Journal of the Japan Society for Construction Information Modeling, 23(4), 572–590.

Volkmann, T., Wessel, D., Caliebe, T. O., & Jochems, N. (2020). *What you see isn't necessarily what you get: Testing the influence of polygon count on physical and self-presence in virtual environments*. In Proceedings of Mensch und Computer 2020 (pp. 119–128). https://doi.org/10.1145/3404983.3405590

Yi, Y. S., & Ha, M. K. (2012). *A study on the psychological responses by the change of illuminance and color temperature in LED*. Journal of the Korean Institute of Interior Design, 21(6), 129–136.

Zimmons, P., & Panter, A. (2003). *The influence of rendering quality on presence and task performance in a virtual environment*. In Proceedings of IEEE Virtual Reality 2003 (pp. 22–26). IEEE. https://doi.org/10.1109/VR.2003.1191170