Effects of the Level of Detail on the Recognition of City Landmarks in Virtual Environments

Achref Doula * Philipp Kaufmann† Alejandro Sánchez Guinea ‡ Max Mühlhäuser §

Technical University of Darmstadt

2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW) | 978-1-6654-8402-2/22/$31.00 ©2022 IEEE | DOI: 10.1109/VRW55335.2022.00281

ABSTRACT

The reconstruction of city landmarks is central to creating recognizable virtual environments representing real cities. Despite the recent advances, it is still not clear what level of detail (LOD) to adopt when reconstructing those landmarks for their correct recognition, and if particular architectural styles represent specific challenges in this respect. In this paper, we investigate the effect of LOD on landmark recognition, generally, and on some architectural styles, specifically. The results of our user study show that higher LOD lead to a better landmark identification. Particularly, Neoclassical-style buildings need more details to be individually distinguished from similar ones.

Index Terms: Human-centered computing—Empirical studies in visualization.

1 INTRODUCTION

One goal behind creating virtual twins for real cities is the generation of environments where several tasks, such as navigation and spatial decision-making, can be performed as if they were executed in the real environments [3]. To design such systems, several factors, like the level of detail (LOD), need to be specified in order to provide an optimal user experience. For cities where several architectural styles coexist, it is not clear how to determine the adequate level of detail to be adopted for the creation of the virtual model. On one hand, the design of models with a high level of detail represents a considerable cost. On the other hand, several architectural styles have a strong focus on details, which is clearly reflected in the design and identity of the buildings in question [1]. Although several works have explored different LOD in virtual urban environments to study their effect on cognitive tasks [2], one common aspect they share is that they were conducted in fictive urban environments that are rather standard and with no link to real cities. This constitutes an experimental virtual environment that lacks the architectural richness present in many real cities, and therefore, overlooks its effects on the given tasks. In this paper, we analyze the effects of varying levels of detail on the cognitive task of landmark recognition. To this end, we conduct a between-subject user study (n=20), in which we ask participants to recognize landmarks from the city of Darmstadt in Germany, after visualizing them in two virtual environments with different levels of detail. The results show that while higher LOD leads to better recognition performance, some architectural styles require more details for landmarks to be recognized.

2 VR SETUP

We implement two virtual environments with different levels of detail (high vs. low) and 8 models for real-world buildings each. The buildings represent different types of urban landmarks (churches, skyscrapers, historical towers, a shopping center, and a central train station) taken directly from the city of Darmstadt, in Germany, and they are representative of various architectural styles present throughout German towns and cities (Neoclassical, Modern, Art Nouveau, and Medieval/Gothic). All buildings were created in Houdini[1] to adjust the levels of detail and were afterwards imported to Unity[2] and put in a scene for VR visualization. A sample of the used building models is depicted in figure 1. Within the virtual environments, the buildings were presented in their real-world scales, with the possibility to scale the scene up and down. The navigation and interaction with the scene were performed using the HTC Vive HMD[3], which comes along with two wireless handheld controllers and two base stations. The integration of the headset into Unity was enabled by SteamVR[4].

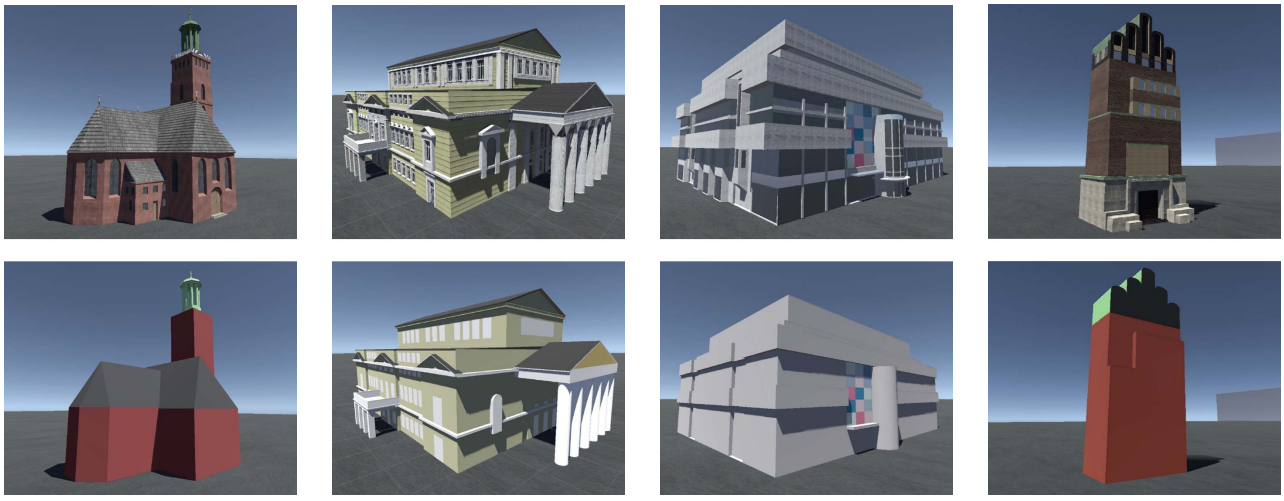
3 EXPERIMENTS AND PROTOCOL

We recruited 20 participants (12 males, 8 females) aged between 18 and 65 ($M = 29.45$, $SD = 11,22$), with an uneven former VR exposure. As we adopted a between-subject design for the study, the participants were randomly assigned to one of the two virtual environments. The task assigned to the participants was to freely roam the virtual environment and memorize the existing buildings. No time limit was given, but the elapsed time was still measured. Next, we exposed the participants to a set of real-life pictures of buildings and asked them to identify the buildings they have seen in the virtual environment. For each building in the virtual environment, we present one correct image and 5 false but similar-looking ones, resulting in a total of 48 pictures (8 buildings x 6 pictures). Figure 2 depicts the similar buildings proposed for the Neoclassical style. Within the task, these pictures were sequentially presented (one at a time) in order to avoid direct comparison between pictures. For each picture, the participants were required to answer with "yes" or "no" if they believed to have seen a corresponding 3D model in the virtual environment, and had to decide directly, i.e., skipping back to a previous picture was not possible. Consequently, participants could assign several pictures from the list to the buildings seen in the Virtual Environment. The order of the pictures within the questionnaire was initially randomly determined and remained the same for every participant. To measure the performance of the participants in the identification tasks, we used two scores as metrics. The first score is the *Identification Score*. It is calculated by accumulating the number of exact correct answers, i.e., when a participant only decided for the correct picture and none of the other five false ones. Since we had 8 buildings in our environments, the maximum identification score that can be obtained is 8. The second metric is the *Relaxed Identification Score*, which is increased by one if the participant assigned the correct building picture to the models seen in the virtual environment, independently of other false choices. The maximum relaxed identification score that can be obtained is also 8.

*e-mail: doula@tk.tu-darmstadt.de
†e-mail: ph.kaufmann@phgk.de
‡e-mail: sanchez@tk.tu-darmstadt.de
§e-mail: max@informatik.tu-darmstadt.de

[1]Houdini: www.sidefx.com
[2]Unity: www.unity.com
[3]HTC Vive: www.vive.com
[4]SteamVR: www.steamvr.com

978-1-6654-8402-2/22/$31.00 ©2022 IEEE
DOI 10.1109/VRW55335.2022.00281



(a) (b) (c) (d)

Figure 1: Example of building models with high vs low level of detail (respectively, in the upper and lower parts): (a) Stadtkirche (Medieval/Gothic) (b) Staatsarchiv (Neoclassical) (c) Luisencenter (Modern) (d) Hochzeitsturm (Art Nouveau)

Figure 2: Example of similar building proposed in our questionnaire (Neoclassical).

4 Results

We used R for the statistical analysis and set the significance level $\alpha = 0.05$. In order to identify statistically significant differences between both groups, either a Student's t-test, a Fisher's Exact Test or a linear regression was applied, depending on the factor that was tested. When the requirements for a t-test were not met, a Mann-Whitney U test was used instead.

First, we report the results of our analysis on the *overall impact* of LOD on landmark recognition, independently of their architectural style. For the identification score, participants given the high LOD environment performed significantly better (M = 4.6, SD = 2.76) than those given the low LOD environment (M = 2.4, SD = 1.78), t(18) = -2.12, p = 0.04804. In contrast to the identification score, the averages for relaxed identification score are higher due to the less strict nature of the metric (low: M = 6.0, SD = 1.49; high: M = 7.3, SD = 1.06). As the preconditions for a t-test were not fulfilled, a Mann-Whitney U test was calculated instead. For the Relaxed Identification Score, the group in the high LOD setting performed significantly better than the participants in the low LOD environment (U = 22, p = 0.03071).

To analyze the effects of changing LOD on the different *architectural styles*, we consider results for each landmark separately. Since the dependent variables are binary, a Fisher's Exact Test was carried out for each building. Only landmarks with the architectural style "Neoclassical" were significantly better recognized in the high LOD setup, and showed a significant difference in the identification score (p=0.033). For the relaxed identification score, the Fisher Exact Test did not show any significant difference.

The results of the user study confirmed that the LOD had a significant impact on building identification. Higher levels of detail lead to a better identification performance. We attribute this to the fact that higher levels of detail reduce the uncertainty of the participants about the visualized building, which makes its identification easier. When considering the architectural style, the results revealed a significant impact of the level of detail on the Neoclassical buildings only. This architectural style is known for its symmetrical, geometrically simple design. Therefore, the relative lack of variance in such buildings justifies the need for more details to establish a distinction between them.

5 Conclusion and Future Work

Our study of the impact of the level of detail on the recognition of urban landmarks in virtual environments has revealed a significant dependence between the two factors. When analyzing the impact on the different architectural styles separately, the results showed a significant impact only on some architectural styles, namely the Neoclassical. In the future, we aim to investigate this dependence further on more cognitively demanding tasks, like navigation.

Acknowledgments

This work has been funded by National Research Center for Applied Cybersecurity ATHENE and LOEWE initiative (Hesse, Germany) within the emergenCITY center. We gratefully acknowledge the support of SideFX Corporation with the donation of Houdini licenses.

References

[1] B. Barz and J. Denzler. Wikichurches: A fine-grained dataset of architectural styles with real-world challenges. *arXiv preprint arXiv:2108.06959*, 2021.

[2] J. P. Handali, J. Schneider, M. Gau, V. Holzwarth, and J. v. Brocke. Visual complexity and scene recognition: How low can you go? In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*, pp. 286–295, 2021.

[3] E. L. Newman, J. B. Caplan, M. P. Kirschen, I. O. Korolev, R. Sekuler, and M. J. Kahana. Learning your way around town: How virtual taxicab drivers learn to use both layout and landmark information. *Cognition*, 104(2):231–253, 2007.