Effects of Object Complexity in Occlusion, Structure, and Texture on 3D Virtual Object Observation in Virtual Reality

Kexiang Shuai
School of Advanced Technology
Xi'an Jiaotong-Liverpool University
Suzhou, China
kexiang.shuai20@student.xjtlu.edu.cn

Yue Li*
Department of Computing
School of Advanced Technology
Xi'an Jiaotong-Liverpool University
Suzhou, China
yue.li@xjtlu.edu.cn

Hai-Ning Liang*
Department of Computing
School of Advanced Technology
Xi'an Jiaotong-Liverpool University
Suzhou, China
haining.liang@xjtlu.edu.cn

***Abstract*—Virtual Reality (VR) environments involve users in 3D virtual object interactions and manipulation tasks. Many of these are for the purpose of 3D virtual object observation, such as viewing a reconstructed museum artifact in a virtual museum. In this paper, we present a study that investigated the effects of object complexity in occlusion, structure, and texture on 3D virtual object observation in VR. We implemented a direct manipulation technique that allows users to grab, move, rotate, and scale an object for close-up observations. Twenty participants used the technique to manipulate virtual objects of various levels of complexity in occlusion, structure, and texture, to complete observation tasks (search and classify marks). The results showed that among the three dimensions of object complexity, occlusion and texture have significant impacts on users' observation task completion time, but structure showed no significant impact. Our work contributes to the understanding of object complexity for 3D object observation in VR environments.**

***Index Terms*—virtual reality, interaction techniques, selection and manipulation, virtual objects**

I. Introduction

Virtual Reality (VR) presents digital environments in which the users could feel fully immersed and interact with. VR transmits information not only through texts or pictures, but through comprehensive sensory feedback, which supports user interaction and experience that is natural, immersive, interactive, and can be easily acquired by non-expert users [1]. With these advantages, VR technology has developed tremendously in recent decades, and its applications have been adopted in various domains and industries, such as education [2], gameplay [3], vocational training [4], product design [5], tourism [6], museums [7], and so on.

In pursuit of study findings that can implicate application areas that involve 3D virtual object observation in VR, we designed a study and based on the direct manipulation technique that most users are familiar with. It allows users to grab, move, rotate, and scale virtual objects. Object observation is a critical task in various virtual environments and scenarios, such as interior design, manufacturing assembly, and virtual museum visits. The virtual objects that users interact with are often of various complexity in occlusion, structure, texture, and sizes. Thus, we designed a study for three object properties: occlusion, structure, and texture. We collected user performance data on object observation tasks (search and classify marks) and evaluated the effect of object complexity on users' task completion time. Our study results showed that among the three dimensions of object complexity, occlusion and texture have significant impacts on users' observation task completion time, but structure showed no significant impact.

Our paper presents the following contributions. (1) We present an empirical study to evaluate the effects of various levels of object complexity in occlusion, structure, and texture on the observation of 3D virtual objects. (2) We found a significant linear correlation between users' perceived optimal size (0.2 to 0.4 meters) and distance (0.6 to 0.7 meters) for direct manipulation in VR.

II. Related Work

The selection and manipulation of virtual objects are important interactions in VR. Back in the 1990s, Bowman et al. [8] compared different techniques for manipulating 3D virtual objects, including the arm-extension techniques and raycasting techniques. One typical example of the arm-extension technique is the go-go technique, which takes advantage of users' arm motions and allows users to reach objects at a distance. They found that the manipulation is easy to use, but it supports a finite distance and may cause imprecision in the selection of distant objects. On the other hand, raycasting allows ease of selection but is difficult to manipulate. The authors argued that the combined raycasting selection and hand-centered manipulation could maximize the ease of use and efficiency.

Recent works tend to separate selection and manipulation as two distinct processes: selection requires the indication of the object and confirmation of the selection; manipulation involves attaching, positioning, and orienting an object. Selection faces issues such as tracking accuracy and jittery [9]. It can be

*Corresponding authors

979-8-3503-0696-5/23/$31.00 ©2023 IEEE
DOI 10.1109/CEII60565.2023.00013

challenging when there is occlusion, and objects being too small or far away. On the other hand, the manipulation of virtual objects faces different challenges, largely influenced by the distance and orientation to the target, translation distance, amount of rotation, and the required precision of the positioning and rotation [10]. Previous works have attempted to use different metaphors such as grasping (e.g., virtual hand [11]), pointing (e.g., raycasting [12]), indirect (e.g., Voo-doo dolls [13]), bimanual (e.g., iSith [14]), and hybrid (e.g., HOMER [13]), where a comprehensive overview of these techniques can be seen in [10]. Among these metaphors, grasping is one of the most commonly used metaphors [15], which is a natural technique that maps users' hand movements with the position and orientation of the virtual object.

Researchers also endeavored to find effective evaluation methods for manipulation techniques. Bowman and colleagues [16] proposed a testbed evaluation method, aiming for standardized evaluation of VR interaction techniques. They identified four categories of factors beyond the interaction technique that may influence object selection and manipulation performance, including the characteristics of the task (e.g., the required accuracy), the environment (e.g., the number of objects), the user (e.g., their spatial ability), and the system (e.g., the display). A recent review work [17] summarized two decades of research on the evaluation of selection and manipulation techniques. The authors summarized ten recommendations for evaluating selection and manipulation techniques, which provide valuable guidelines for evaluation research on interaction techniques. This work follows the suggested evaluation guidelines and presents an evaluation study on direct manipulation technique for 3D virtual object observation in VR.

III. System Implementation

We built the prototype system and implemented the technique using Unity (version 2019.2.0f1) on a computer with a Core i7-9750H CPU @ 2.60GHz, 8GB RAM, and NVIDIA GeForce GTX 300 graphics card with 8GB RAM. The VR system was developed based on the SteamVR package, which is available in the Unity Asset Store and supported on most mainstream VR headsets. Our system was deployed on the Meta Quest 2 VR HMD, with a 1920 × 1832 resolution for each eye and a 72 Hz refresh rate. Interaction techniques were based on the two hand-held controllers (see Fig. 1). We set up C# scripts to record user interaction data and exported them to CSV files for further analysis.

A. Implemented Techniques

We implemented a direct manipulation technique and realized three basic functions, allowing users to 1) grab an object at a distance and 2) change the object's position and orientation, and 3) scale the size of the object. The technique is based on the grasping metaphor and the virtual hand manipulation. We simplified the selection process in the techniques as we focus on the object observation task, which is more concerned with manipulation, namely, moving, rotating, and scaling an object to obtain an overview and delve into the details. To grab a remote object, the user needs to keep pressing the `GRIP` button of either controller to activate the magnetic function, and then press the `TRIGGER` button to grab the object. Similar to the virtual hand manipulation technique, the virtual object's translation, and rotation will follow the user's hand like people grabbing objects in the real world. The scaling of the object in hand is achieved by pressing the two `TRIGGER` buttons on both controllers. An object is scaled up when the distance between two controllers increases and is scaled down when the distance decreases.

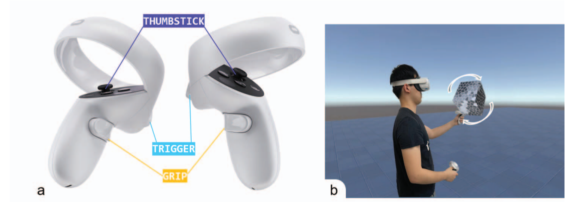

Fig. 1. (a) Meta Quest 2 VR controller and the main control buttons. (b) A participant grabbing and observing a 3D virtual object in VR.

B. Experimental Environments and Observation Tasks

As shown in Fig. 2a, we implemented an experimental environment for observation tasks. Users start the experiment by pressing the `Get Object` button on the right. The timer starts and an object shows up on the bench at a distance in front of the user. Users then need to grab and manipulate the virtual object to find the mark on the object (see Fig. 2b for examples), and classify it to the corresponding category by pressing the buttons on the left. The object is dismissed when the classification is made, and the timer for the trial stops. We set up objects of various sizes when they are instantiated to simulate various object sizes in the real world (see Fig. 2c. Users are allowed to scale them to an ideal manipulation size after grabbing them.

IV. Study Design

A. Research Questions

The study aims to answer four research questions:

RQ1. Does virtual object complexity in occlusion affect observation efficiency in VR?

RQ2. Does virtual object complexity in structure affect observation efficiency in VR?

RQ3. Does virtual object complexity in texture affect observation efficiency in VR?

RQ4. What is the optimal size and distance for virtual object observation in VR?

B. Study Conditions

1) Occlusion.: Object complexity in occlusion indicates the amount of hidden areas of an object. We used spheres with dents to simulate complexity in object occlusion, as shown in Fig. 3. The simulation of occlusion is achieved by indenting part of the surface of the sphere. The excess surface area

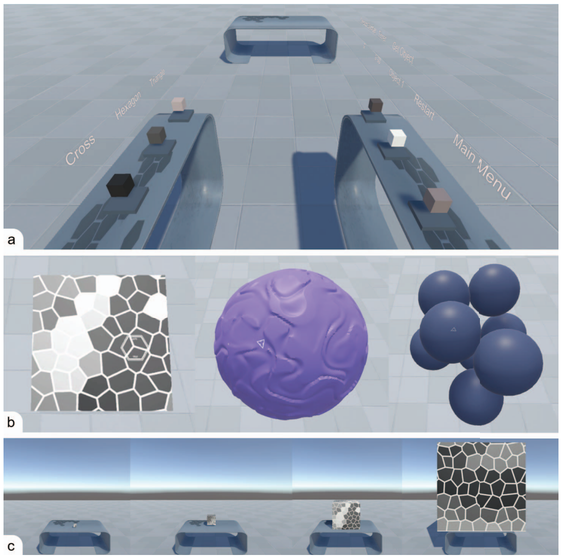

Fig. 2. Study setup. (a) The experimental scene with a fixed user position in all trials, with menu buttons on the right and classification buttons on the left. (b) Example objects with marks on them. (c) Example objects in four different sizes: 0.1 m, 0.3 m, 0.9 m, and 2.7 m.

compared to the surface area of a sphere of equal volume is the hidden area. The ratio of the hidden area to the surface area is the degree of occlusion complexity. An object is of low, medium, and high complexity in occlusion when the proportion of the hidden area reaches 10%, 40%, and 80%.

Fig. 3. Objects of low, medium, and high complexity in *Occlusion*.

2) Structure: Object complexity in structure is defined by the physical shape of an object. We create sample objects with 10 uniform spheres as shown in Fig. 4. The spheres are arranged in different ways, and the structure appears converged or scattered. The length of the approximate cylinder decides the degree of structural complexity. An object is of low, medium, and high complexity in structure when the length of the approximate cylinder is 1, 1.5, and 2 meters.

Fig. 4. Objects of low, medium, and high complexity in *Structure*.

3) Texture: Texture is the visual complexity of an object surface. We used a cube with texture mappings that were divided into multiple pieces of polygons, as shown in Fig. 5. The amount and the density of the polygons on the object's surface determine the degree of texture complexity. An object is of low, medium, and high complexity in texture when it has approximately 25 (5 × 5), 100 (10 × 10), and 400 (20 × 20) polygons on each face of the cube.

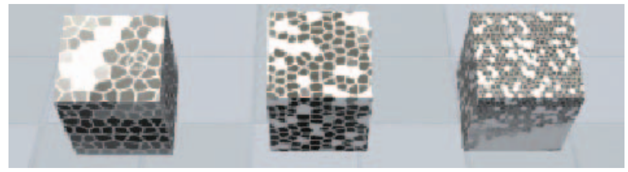

Fig. 5. Objects of low, medium, and high complexity in *Texture*.

C. Study Procedure and Tasks

Prior to the experimental sessions, participants' informed consent was collected and they were asked to complete a pre-experiment questionnaire to collect demographic information. Then, participants were instructed to enter a tutorial scene to familiarize themselves with the direct manipulation technique, as well as the layout and system controls in the experimental scene. After they finished the tutorial session and got prepared, they entered the experimental scene to get ready to start the experimental sessions. There are 9 sessions (3 Complexity Dimensions × 3 Complexity Levels) in total. We applied a randomized design for the experimental sessions to avoid the influence of the experimental order on the results. Each session had 4 trials of different object size (see Fig. 2c). Participants were asked to scale the object, confirm the best viewing size and distance, and search the mark on the object. Once participants completed the nine sessions, they were invited to attend a short interview to provide their comments. This study was conducted in a 2m × 2m space, where participants were seated on a chair. The whole study lasted for ∼25 minutes on average including the tutorial session (∼5 minutes).

D. Measures

We measured the categorization accuracy of each session (9 in total) and the task completion time for each trial (36 in total). The accuracy of each session (A) is calculated by $A = S/T$, where S is the number of successful trials and T is the number of total trials. The completion time is recorded as the duration between the time when the object is instantiated and the time it is classified.

E. Participants

Twenty participants (12 males, 8 females) aged between 18 and 25 ($M = 20.06, SD = 1.61$) voluntarily took part in this study. Five of the participants have never used VR; nine participants' total VR use time is under 10 hours; six of them use the VR system frequently, with a total VR use time above 40 hours. All participants have prior experience in 3D graphics, e.g., 3D games or 3D modeling software. They reported an average of 3.38 ($SD = 1.39$) for VR familiarity and an average of 3.75 ($SD = 1.01$) for 3D graphics familiarity (5 = extremely familiar).

V. Results

We performed the data analysis using IBM SPSS Statistics 26. For the objective data analysis, all of the recorded categorization accuracy was 100%, so we only analyzed the measurement of the task completion time. Extreme outliers were recognized by checking whether the studenized residual exceeded ±3 times the standard deviation and were removed from the analysis. We checked the distribution of the data by examining the Shapiro-Wilk test results and determined whether the interaction term of each group satisfies the condition of Mauchly's spherical hypothesis test. We took the Greenhouse-Geisser method to rectify the result when the spherical test condition was not satisfied. One-way ANOVA were conducted to test the effects of object complexity (low, medium, and high) on the task completion time. The results are presented in Fig. 6.

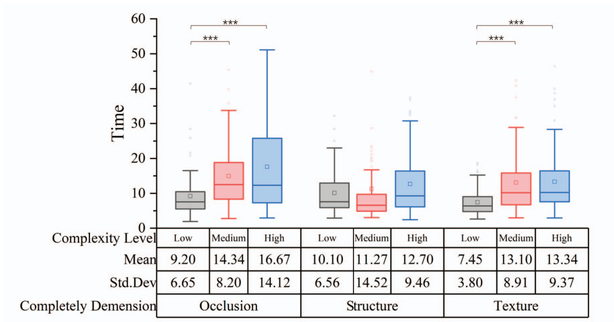

Fig. 6. The results of task completion time for objects with three levels of three complexity dimensions: occlusion, structure, and texture.

A. Occlusion

A one-way ANOVA showed a significant difference among complexity levels of object occlusion on task completion time, $F(2, 132) = 12.653, p < 0.001$. A Tukey post hoc test revealed that significantly more time was required for medium ($14.34 \pm 8.20s, p < 0.001$) and high ($16.67 \pm 14.12s, p < 0.001$) object occlusion than low object occlusion ($9.20 \pm 6.65s$). The difference between medium and high object occlusion conditions was insignificant ($p = 0.544$).

B. Structure

A one-way ANOVA showed no significant difference among complexity levels of object structure on task completion time, $F(2, 134) = 1.128, p = 0.319$.

C. Texture

A one-way ANOVA showed a significant difference among complexity levels of object texture on task completion time, $F(2, 134) = 14.687, p < 0.001$. A Tukey post hoc test revealed that significantly more time was required for medium ($13.10 \pm 8.91s, p < 0.001$) and high ($13.34 \pm 9.37s, p < 0.001$) texture complexity than low texture complexity ($7.45 \pm 3.80s$). The difference between medium and high texture complexity conditions was insignificant ($p = 1.00$).

D. Size and Distance

A regression analysis was conducted to model the relationship between the optimal object size and distance identified by the participants for object observation. A significant linear relationship was found between distance and size, $F(1, 79) = 63.492, p < 0.001$. The linear regression equation is $Size = 0.818 * Distance - 0.198$. The regression results and the detailed ranges are shown in Fig. 7. The interquartile range is reported as the optimal range: 0.6 to 0.7 meters for distance and 0.2 to 0.4 meters for size.

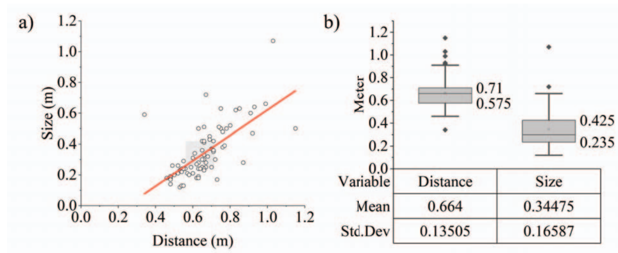

Fig. 7. (a) The regression results of the optimal distance and size. The shaded boxes indicate the range of optimal observation distance and size. (b) Detailed results of the optimal distance and size.

VI. Discussion

In this paper, we evaluated the effects of three dimensions of object complexity: occlusion, structure, and texture on 3D virtual object observation, and examined the optimal size and distance for object observation in VR. The 100% accuracy rate for all participants indicates that the direct manipulation technique is effective for users to observe objects and figure out the right answers. It also shows the satisfying usability of the technique, which excludes the influence of other variables such as users' familiarity with the interaction technique. The main differences were shown in task efficiency.

Overall, it was found that object complexity in occlusion (**RQ1**) and texture (**RQ3**) has a significant impact on users' observation task completion time, but structure (**RQ2**) showed no significant impact. Combined with participants' comments in the interview, we found that participants encountered difficulty in observing objects with uneven surfaces and irregular structures. Participants reported that observation tasks with these objects require close looks and examinations, especially when the object's occlusion complexity becomes high. As P2 said, *"when observing objects with high occlusion complexity, I have to closely examine the object's surface to find the mark."*. The direct manipulation technique we selected enabled users to observe virtual objects with the same mental model of object interaction in the real world, which is particularly helpful for novice users who are not at all familiar with 3D interfaces or VR. For example, one of our participants (P6) told us that she had never used gamepad or joysticks. We found that she could quickly acquire the technique. However, observing objects with high occlusion complexity was reported to be the most difficult task for most participants, including skilled

VR users. This implies the need to optimize the interaction design for tasks involving the manipulation of complex objects with occluding parts, such as car engines and human anatomy. Potential approaches include enabling the segmentation of parts and using cutting planes [18]. In the meantime, the texture complexity of objects also had a significant effect on the efficiency of the observation tasks. A growing body of VR systems are designed for sectors that contain objects of various textures, such as interior design, museums and exhibitions, and the fashion industry. These VR systems should consider improving the design for observing objects with complex textures, such as providing additional visual cues to direct users' attention towards a point of interest. Our analysis of user-identified optimal size and distance (**RQ4**) implies that within an arm's reach (0.6 to 0.7 meters), the optimal size is around 0.2 to 0.4 meters for direct manipulations in VR.

There are some limitations in this study. Our study only examined the direct manipulation technique based on the grasping metaphor for object observation in VR. It is worth testing other interaction techniques in VR, such as raycasting [12], HOMER [13], as well as indirect manipulation techniques [19], which may lead completely different results. In addition, our work roughly categorized the object complexity into three levels. Future work could attempt to quantify the complexity level and identify the benchmarks in order to inform more explicit design decision. Besides, using generic 3D shapes helps control the level of complexity, but also sacrifices realism. Studying user observation of real-world objects may yield new findings. In our future work, we also plan to repeat the study in real-world scenarios such as virtual museums to test if the results of the current study can be generalized to 3D virtual objects simulating real-world objects.

VII. Conclusion

In this paper, we present a study that investigated the effects of three dimensions of object complexity on 3D virtual object observation in VR: occlusion, structure, and texture. We found a significant impact of object complexity in occlusion and texture on observation task efficiency, but the complexity in structure showed no significant impact. We also found a significant linear correlation between users' perceived optimal size and distance for direct manipulation, showing an optimal size of 0.2 to 0.4 meters and an optimal distance of 0.6 to 0.7 meters. We discuss the design implications for future VR system that involve complexly occluded and textured objects in potential areas such as manufacturing, medical and healthcare, cultural heritage, and fashion. These findings are useful for the future design of object interaction techniques and VR systems.

Acknowledgment

The authors would like to thank all participants for their time and support. This work is supported by the National Natural Science Foundation of China (62207022), the Natural Science Foundation of the Jiangsu Higher Education Institutions of China (22KJB520038), and Xi'an Jiaotong-Liverpool University (RDF-20-02-47).

References

[1] N. Magnenat-Thalmann and G. Papagiannakis, "Virtual worlds and augmented reality in cultural heritage applications," 2005.

[2] S. Xu, X. Hong, and H. Yang, "The Application of VR Technology based on Computer Vision in Physical Education," in *Proceedings of the 7th International Conference on Information and Education Innovations*, 2022, pp. 176–181.

[3] S. Jeong, H. H. Yun, Y. Lee, and Y. Han, "Glow the Buzz: a VR Puzzle Adventure Game Mainly Played Through Haptic Feedback," in *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, 2023, pp. 1–6.

[4] M. Rettinger, N. Müller, C. Holzmann-Littig, M. Wijnen-Meijer, G. Rigoll, and C. Schmaderer, "VR-based equipment training for health professionals," in *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021, pp. 1–6.

[5] Y. Shao and X. Min, "VR and AR Technology used in Education in the Education of Packaging Design for Students," in *2021 2nd Asia-Pacific Conference on Image Processing, Electronics and Computers*, 2021, pp. 1073–1076.

[6] D. Mirk and H. Hlavacs, "Virtual tourism with drones: experiments and lag compensation," in *Proceedings of the First Workshop on Micro Aerial Vehicle Networks, Systems, and Applications for Civilian Use*, 2015, pp. 45–50.

[7] G. Schofield, G. Beale, N. Beale, M. Fell, D. Hadley, J. Hook, D. Murphy, J. Richards, and L. Thresh, "Viking VR: designing a virtual reality experience for a museum," in *Proceedings of the 2018 Designing Interactive Systems Conference*, 2018, pp. 805–815.

[8] D. A. Bowman and L. F. Hodges, "An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments," in *Proceedings of the 1997 Symposium on Interactive 3D Graphics*, ser. I3D '97. ACM, 1997, p. 35–ff. [Online]. Available: https://doi.org/10.1145/253284.253301

[9] R. Kopper, F. Bacim, and D. A. Bowman, "Rapid and accurate 3D selection by progressive refinement," in *2011 IEEE symposium on 3D user interfaces (3DUI)*. IEEE, 2011, pp. 67–74.

[10] J. J. LaViola Jr, E. Kruijff, R. P. McMahan, D. Bowman, and I. P. Poupyrev, *3D user interfaces: theory and practice*. Addison-Wesley Professional, 2017.

[11] Y. Shi, L. Zhao, X. Lu, T. Hoang, and M. Wang, "Grasping 3D Objects With Virtual Hand in VR Environment," in *The 18th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and its Applications in Industry*, 2022, pp. 1–8.

[12] T. Nukarinen, J. Kangas, J. Rantala, O. Koskinen, and R. Raisamo, "Evaluating ray casting and two gaze-based pointing techniques for object selection in virtual reality," in *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*, 2018, pp. 1–2.

[13] J. S. Pierce and R. Pausch, "Comparing voodoo dolls and HOMER: exploring the importance of feedback in virtual environments," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2002, pp. 105–112.

[14] M. Yamagami, S. Junuzovic, M. Gonzalez-Franco, E. Ofek, E. Cutrell, J. R. Porter, A. D. Wilson, and M. E. Mott, "Two-in-one: A design space for mapping unimanual input into bimanual interactions in vr for users with limited movement," *ACM Transactions on Accessible Computing (TACCESS)*, vol. 15, no. 3, pp. 1–25, 2022.

[15] M. Weise, R. Zender, and U. Lucke, "A comprehensive classification of 3D selection and manipulation techniques," in *Proceedings of Mensch und Computer 2019*, 2019, pp. 321–332.

[16] D. A. Bowman, D. B. Johnson, and L. F. Hodges, "Testbed evaluation of virtual environment interaction techniques," in *Proceedings of the ACM symposium on Virtual reality software and technology*, 1999, pp. 26–33.

[17] J. Bergström, T.-S. Dalsgaard, J. Alexander, and K. Hornbæk, "How to Evaluate Object Selection and Manipulation in VR? Guidelines from 20 Years of Studies," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, 2021.

[18] N. Elmqvist and P. Tsigas, "A taxonomy of 3D occlusion management for visualization," *IEEE transactions on visualization and computer graphics*, vol. 14, no. 5, pp. 1095–1109, 2008.

[19] Y. Wang, Y. Li, and H.-N. Liang, "Comparative Analysis of Artefact Interaction and Manipulation Techniques in VR Museums : A Study of Performance and User Experience," in *2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 2023, pp. 761–770.