Adapting VST AR X-Ray Vision Techniques to OST AR

Conference Paper · October 2022

DOI: 10.1109/ISMAR-Adjunct57072.2022.00104

5 authors, including:

Thomas Clarke
University of South Australia

Joanne Zucco
University of South Australia

Brandon Matthews
University of South Australia

Ross T. Smith
University of South Australia

Adapting VST AR X-Ray Vision Techniques to OST AR

Thomas J. Clarke*
IVE, Wearable Computer Lab, University of South Australia

Wolfgang Mayer†
Industrial AI Research Centre, University of South Australia

Joanne E. Zucco‡
IVE, Wearable Computer Lab, University of South Australia

Brandon J. Matthews§
IVE, Wearable Computer Lab, University of South Australia

Ross T. Smith¶
IVE, Wearable Computer Lab, University of South Australia

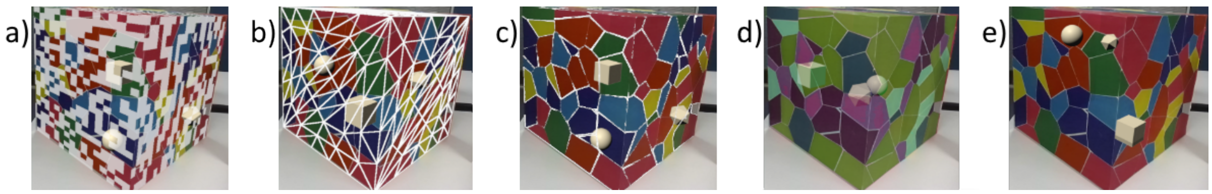

Figure 1: The x-ray visualizations shown augmented on a physical cube with a printed Voronoi pattern evaluated in this paper a) Random Dot [23], b) Tessellation, c) Edge Based [16], d) Saliency [28] and e) No Visualization (None)

ABSTRACT

When merging physical and virtual objects with optical see-through augmented reality there is little research that has focused on x-ray vision visualizations considering depth perception. We investigate partial occlusion visualizations when merging visual cues and real-world objects to explore the effect of the visualizations with a procedural placement task. We adapted existing x-ray visualization techniques designed for Video See-Through (VST) Augmented Reality to operate on Optical See-Through (OST) devices and investigated how these techniques affect accuracy in a placement task within arms reach. We evaluate the visualizations' impact on accuracy when user movement is unrestricted and report the perceived usability and mental load of each visualization. Our findings indicate that although the type of x-ray visualization is important, the presence of other virtual objects in the scene appears to have a stronger impact on the users' accuracy in placing objects and user experience.

Index Terms: Augmented Reality—Visualization techniques—X-ray Vision—X-ray Visualizations; Human Computer Interaction—Visualization—Visualization Design and Evaluation Methods—User Studies

1 INTRODUCTION

This is the first paper to adapt x-ray visualizations, specifically Saliency- and Edge-Based visualizations, to an Optical See Through (OST) display to better understand the impact of different visualizations on the user's ability to place virtual objects in an augmented space. Augmented Reality (AR) devices such as the Microsoft HoloLens provide us with the ability to overlay virtual information over the physical world. When placing virtual objects behind or inside an object in the physical world there are depth and spatial perception challenges when understanding their spatial relationship. This investigation is focused on x-ray visualizations where there is a need for a tight coupling between the physical and virtual information, such that users perceive the virtual and physical information as one. Our goal is to further build knowledge about x-ray visualizations during a placement task with OST AR display technologies.

Overlaying three-dimensional virtual content on the physical world can be problematic in terms of aligning the perceived depth of the virtual and physical information in AR. This is made worse when virtual objects are placed beyond physical objects [3] as our sense of depth is influenced by our ability to recognize visual relationships. When visual cues in AR are not carefully designed, the user may perceive virtual objects to be smaller instead of further away [4]. X-ray vision is one approach that has been explored to address some of the challenges of correctly perceiving where virtual objects are positioned in the physical world.

X-ray vision in AR is defined as the use of partial occlusion to provide an understanding of the depth of virtual objects that are rendered behind physical world objects [2]. There are three main approaches to x-ray vision that can be used individually or together. One can render a texture over the surface that creates a degree of occlusion that the user can relate to [16, 23, 28], create a visualization that conveys to the user the distance through objects such as a tunnel [1], or show objects that are hidden in the physical world by rendering them as virtual objects [3, 19, 30].

OST AR is preferred over VST AR in applications that require the user's vision to be unobstructed like driving [11], Surgery [27], and Security operations [26]. VST technologies suffer latency issues and are vulnerable to hardware failure which could be catastrophic for mission-critical tasks like surgery. OST does not suffer from these limitations and may be better suited to the applications mentioned above motivating further investigation into the use of x-ray vision techniques on OST technologies.

A strength of x-ray vision on VST AR devices is its ability to improve the visual connection between the real world and virtual objects, It allows us to judge a partially occluded object to better explain where a virtual object is in relative space in the virtual world [16, 23, 28, 36]. X-ray visualizations designed for VST AR enable us to integrate real-world visual information into how virtual objects are rendered. We believe that these visualizations will create a better sense of depth when used with more objects.

This research aims to adapt existing x-ray vision techniques to OST AR and to determine which techniques perform better in a placement task. We focus on within arms reach (action space) interactions with an x-ray vision method for viewing information

*e-mail: thomas.clarke@mymail.unisa.edu.au
†e-mail: wolfgang.mayer@unisa.edu.au
‡e-mail: joanne.zucco@unisa.edu.au
§e-mail: brandon.matthews@mymail.unisa.edu.au
¶e-mail: ross.smith@unisa.edu.au

inside a physical object, such as viewing human internal anatomy. We also investigate the impact of nearby objects and how grouped virtual reference objects impact placement accuracy.

The contributions of this paper are:

- A method to apply computer vision techniques to an OST AR display allowing for comparison between four x-ray vision effects.
- A study of the effects of the four x-ray visualizations (Saliency, Edge, Random Dot, Tessellation) and the introduced reference objects on the users' ability to precisely place an object.

This paper explores how x-ray vision blending effects can create depth cues that help understand a physical space that is hidden in the real world, such as the interior of an opaque physical object. In Section 2 we discuss related work to describe the progression of previous x-ray vision methods. In Section 3 we describe how we adapted the VST techniques to the OST AR device. Section 4 highlights our experiment setup including our hypotheses, research questions, recruitment practices, study design, and procedure. The results of our study are presented in Section 5, their implications are discussed in Section 6, and the limitations of the study are discussed in Section 7.

2 Background

Kalkofen et al. [16] created an x-ray vision effect by highlighting the edges of real-world objects. These edges served to highlight details in the foreground and leverage these details to partially obscure any object found behind it. This is known as Edge-Based x-ray visualization. Avery et al. [1] used a type of virtual box with no back plate called tunnelling. To give the user an accurate sense of depth through an object, the relative distance through that object was represented as a tunnel that helped the user understand the depth/distance between one point and another. The Edge-Based effect was used to prevent the tunnel from appearing on top of the visualization.

Saliency is an x-ray visualization technique that removes elements of the scene of interest and allows users to see through an object [28]. This creates areas of interest that are affected by x-ray vision and areas that are of less interest to the user (presented in the background based on Zollmann et al.'s [37] ghosting x-ray visualization). Sandor et al. [28] found that this visualization was preferred over the previous Edge-Based techniques.

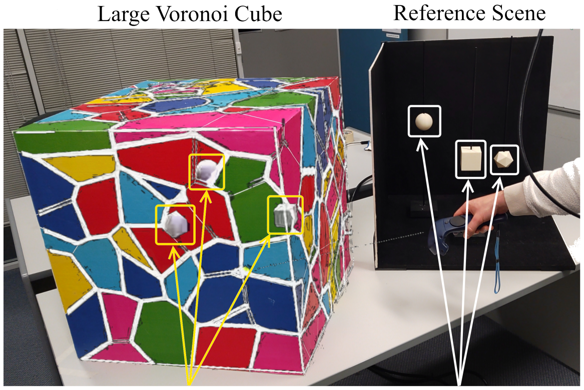

Figure 2: The study environment, taken from a 3rd person's Microsoft HoloLens. Left is a physical 60cm^3 cube with a Voronoi pattern where x-ray vision techniques are performed (Edge based visualization in this photo) and the virtual objects are displayed inside. Right is a reference scene with movable 3D printed objects that need to be replicated during the task.

Otsuki et al. [23] created an effect that was fast to render and provided an occlusive cue that was not dependent on any optical cues and could take advantage of a depth map, allowing the user to rely on binocular disparities. Their visualization used square dots that aligned with the user's view of the real world. Otsuki et al. [22] found that larger dot sizes resulted in a higher level of depth perception and enabled some transparency.

A wireframe visualization was first used by Tsuda et al. [32] who found that a wireframe cue could help the user determine an object's relative position to a reference frame. The planes on this visualization would also darken everything found behind them, lowering the opacity of a virtual object behind it. It was later found out that too much transparency could have a negative effect on depth perception when using a wireframe visualization [10], and Gruenefeld et al. [14] therefore attempted to use a denser mesh of triangles or a Tessellation effect to represent the object being viewed into.

Gruenefeld et al. [14] also produced one of the first depth perception studies done using an OST headset utilizing x-ray vision, including Grid and Cut-out effects. The grid effect, placed on the ground and a perpendicular wall, could be used to approximate a relationship between the wall and the object. In contrast, the cut-out visualization provided a hole in the wall, the user could look through to get a sense of where the object was on the other side. Gruenefeld et al.'s [14] found that the weakest of the depth cues were the wireframe and the cut-out. This was probably since neither of these cues conveyed depth indications to the user.

Martin-Gomez et al. [20] studied the difference between several x-ray vision effects (None, virtual hole, ghosting, and Random Dot) on both VST AR and OST AR devices in the near field. Their work is interesting as it is one of the few approaches that place objects close to the user. They found that users were better at using x-ray vision on a VST AR headset than on an OST AR device. Martin-Gomez et al. [20] investigated different rendering techniques for x-ray vision (shading, hashing, ghosting) and levels of brightness, finding that bright clear objects work best in OST AR.

Interestingly, only a few studies considered within arms reach x-ray vision, like Martin-Gomez et al. [20] and Santos et al. [29]. Ozgur et al. [24] showed that highlighting an object as the user gets closer to it will improve the position of the object in a surgical situation. This lack of work may be due to many of the depth perception tasks in the near field of AR using reaching and matching techniques that do not work as well with physical obstacles [21,31].

3 Video See-Though Overlay

To take advantage of x-ray visualizations that were traditionally designed for VST AR systems, a system was developed that could transmit calibrated camera data to an OST AR display. This system was designed for displaying the output of a video feed over a user's vision for x-ray vision, allowing us to highlight areas of interest in the user's vision but its potential uses are not limited to just this task.

The HoloLens2 was used because it afforded us the ability to add additional sensors positioned relative to the display (to allow for more precise controls and a faster image input), the option to increase the power of the device, and a single static depth plane. We also used a Zed Mini as our input camera as it was able to take more images faster than the HoloLens web camera and didn't reduce the HoloLen's frame rate below 60fps.

Calibrating video see-through techniques, such as Saliency or Edge-Based visualizations, for video see-through headsets and mobile devices is a challenging calibration problem [1, 28]. To date, we found only one method for aligning computer vision techniques with someone's actual bifocal vision [15]. This method developed by Hamadouche [15] created a method to display edge-based x-ray vision using virtual projectors in collaboration with the HoloLens which used a homography to align a video feed with a 3D model. Instead of a homography approach we use a lens distortion calibration method between the camera and the HoloLens2. Although no performance numbers were reported, we expect that our technique can achieve a higher frame rate.

Each time a new image from the video feed was received, the image was filtered to represent either a Saliency or an Edge-Based map and then distorted to align with the users' vision. This image was then used as input into a virtual projector from the perspective of the user and interacted with the x-ray-able object. For Saliency, shades of black are projected onto the surface. These appear as transparent or semi-transparent in OST AR devices, partially occluding the objects rendered within the cube. For the Edge-Based visualization, white was used to highlight the edge features.

Radial distortion is required to match the virtual representation of the real world so that the processed video feed from the camera is accurately aligned with the user's vision [7, 9, 35]. By applying this to the visualization frame, we matched the edges of the display to the user's vision of the real world. Since the camera was not located at the same height as the user's eyes we applied a distance-based distortion from the headset to the visualization. A set of tests were performed at different distances to determine the offset required to account for the offset. These values were used to compute a mapping between the distance from the visualization and the offset required to align the camera to the user's eyes. Linear interpolation was used to fill in between the known samples. This gave the user the impression the projection was displayed correctly on the box.

To overcome the latency created by the connection between the camera, the PC, and the Hololens; we used high-speed cables (the shortest being 4m) and found the latency of our system was 48ms using a system from Gruen et al.'s [13]. The results informed the parameters of an asynchronous re-projection algorithm which recorded the position the user was looking at each frame and used the position recorded at the time the image was taken [18, 33].

Asynchronous re-projection made two more issues apparent: a) the difference between the estimated and actual positions of where the picture was taken and b) the user's involuntary head movements. To solve this issue, the position of the output image was first sent into a one euro filter [8]. The filter provided the flexibility to allow large sweeping head movements to appear as if they had no latency but filtered values when the user stopped moving. This kept the image where the participant expected it without restricting their movement.

4 User Study

We designed a placement task aimed at assessing how well participants were able to place virtual objects within the 60cm^3 Voronoi cube shown in Figure 2, while users were able to walk freely within a prescribed area in the study environment. The x-ray visualizations were chosen because they share similar properties in that they all relied on a back pane showing the user the other side of the box. We tested the user's accuracy as the distance between the actual placement position and the target position. We tested this against two independent variables: a) The presence of the reference objects and b) the x-ray visualizations (Random Dot [23], Tessellation(A more dynamic variation of wire-frame), Edge-Based [16], Saliency [28] and a baseline utilizing none of the visualizations). The x-ray visualizations in our study are shown in **??**. Additional measures included the participant's cognitive load and usability indicators.

Our research questions were as follows:

R.1 Is there a difference in accuracy when placing a virtual object when different x-ray visualization effects are used?

R.2 Are there differences in perceived usability and cognitive load when using different x-ray visualizations on an OST headset?

4.1 Pilot Study

A short pilot study was run on a cohort of 5 participants to test the viability of the study design to test depth perception using both a VST AR device and an OST AR device. Similar to the design of studies done by Outski et al. [12] and Martin-Gomez et al. [20] also using the Voronoi cube (shown in Figure 2) where the participants were asked to identify what geometric shapes were inside or outside of the box. This pilot had similar results to Martin-Gomez et al.'s [20] test regarding both devices, where both devices produced a significantly different effect in regards to depth-based accuracy with VST AR being more accurate but having more uncertain answers. The visualization effect was not significant; however, Saliency proved the most effective x-ray visualization for depth perception, and no visualization performed the worst. A System Usability Scale (SUS) [6] showed that participants preferred to use the camera-based effects, but that result was not statistically significant. We selected our hypotheses based on these results and modified the conditions and system parameters based on our participant feedback.

4.2 Hypotheses

We hypothesized users will be able to place the object more accurately using Saliency because it performed the best in our preliminary tests that focused on depth perception, and we expected similar results for Saliency when placing objects (**H.1**). We expected an improvement in depth perception due to the impact the reference objects would have on the relative size and density because of findings by cutting [34] (**H.2**) The benefits to depth perception are both highly effective at any distance and we believe that they will have an even greater impact when used with any x-ray visualizations (except for None).

Although Saliency may be a powerful depth cue assisting with more accurate placement of objects, we suspected it required much more attention from users which may result in a greater cognitive effort because it relies on the correct interpretation of the Saliency visualization by the user (**H.3**).

Our hypotheses of the study are as follows:

H.1 Participants will place the virtual icosahedron closer to the correct target position when they are using the Saliency visualization.

H.2 Participants will place the virtual icosahedron closer to the correct target when reference objects are present with any x-ray visualization (other than None).

H.3 Participants will find Saliency subjectively difficult to use and require a higher cognitive load than other x-ray vision effects.

4.3 Participants

The study recruited 22 participants between the ages of 22-44 (mean:29.35 sd:6.43). Two of the participants were female and 20 were male. All participants were required to have a normal or corrected to normal vision regarding depth perception. This was determined by user self-assessment prior to the study. Two male participants who performed were removed from the results one was found to have vision issues regarding depth after the study had concluded and the other did not understand the task.

4.4 Software Implementation and Technical Design

Participants were provided with a HoloLens2 headset fitted with a Zed mini mounted on the front and a Vive puck mounted on the back using 3D printed mounts. To interact with the virtual objects, the user was given a Vive controller for which the position relative to the hololens was calculated using the relative position of the Vive puck. To lower the sensitivity of the Vive sensors and the HoloLens2 connection, a one-euro filter [8] was used to remove any noise caused by the Vive's hardware that may be transmitted to the HoloLens2.

4.4.1 Task and Reference Scene

The physical study environment consisted of two physical cubes placed in an area that the participant could freely walk within. The area in which the participant could walk was constrained so that

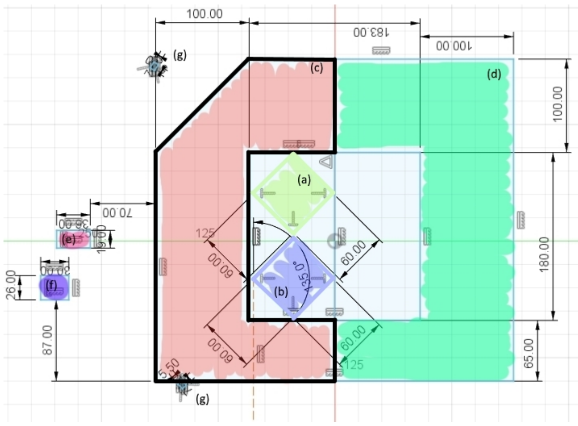

Figure 3: Study setup: a) Voronoi Cube; b) Reference Scene; c) Area the participant could traverse (outlined in black); d) Examiner area; e) Main Processing Server, f) Projector, g) Vive sensors

they were at most 1m from the physical scene. The study area arrangement is depicted in Figure 3.

The cubes, shown in Figure 2, were very distinct in appearance and each measured 60cm along each axis. The reference scene on the right-hand side of Figure 2 held smaller objects held up by stilts placed in the cube. The objects would be used as a reference by the participant. The Voronoi cube pattern was designed to support the Edge and Saliency effects. This cube held the virtual objects displayed by the HoloLens2 that the user could interact with.

The reference scene was designed to be a 1:1 physical representation of the virtual world displayed in the Voronoi cube. Both of these cubes were designed to hold three geometric objects: a cube, a sphere, and an icosahedron. The 3D-printed objects within the reference scene were ivory color to ensure that the shadows on the objects would be similar to the virtual ones displayed by the Hololens2 and were presented in the same orientation as the virtual objects. Each stilt featured a square base measuring 10cm x 10cm, ensuring that the center of each object was positioned at least 5cm away from any other object. This was done so that placing the virtual object required participants to use their spatial estimation skills.

The geometric objects were identical in the physical reference scene and the virtual space. Each object had similar dimensions. The sphere (113.3cm^3) and the icosahedron (109.9cm^3) were made to be the largest size possible that would fit into the reference cube (216cm^3). The placement of each reference object was decided randomly along all axes while accounting for collisions and ensuring that each object was wholly within the reference scene.

4.5 Procedure

At the start of each session, participants completed a demographics survey, and then they were asked to don a Hololens2 and given an HTC Vive controller to perform interactions. Participants could interact with the interior of the box by using a length-adjusting ray that protruded out of their Vive controller. Participants were given a tutorial that explained how to use the controller to change the size of the ray and that the color of the virtual object would change when it was selected.

Participants were instructed to place the virtual icosahedron in the same position in the Voronoi cube as the physical icosahedron was positioned in the physical reference scene. They were free to move around area (c) outlined in Figure 3 and they could replace the virtual icosahedron as many times as required until they believed the object was correctly placed. The other reference objects and their physical counterparts in the reference scene were also pointed out to them. In each iteration of the study, one of four different x-ray

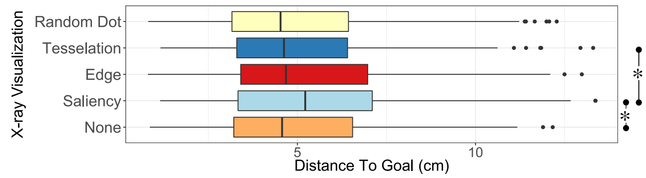

Figure 4: Accuracy data graphed showing the distance to the correct target position for the various blending effects ($*$ is $p < 0.5$, $**$ is $p < 0.01$, $***$ is $p < 0.001$).

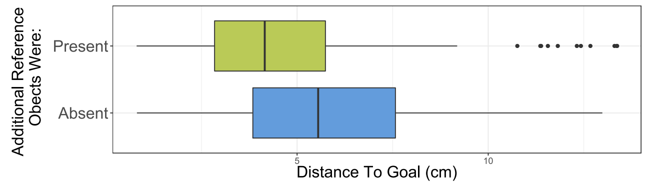

Figure 5: Accuracy data graphed showing the distance to the correct target position for the different clustering effects ($p < 0.0001$).

visualizations or no visualization would be randomly chosen and shown on the outside of the cube. The positions of the geometrical objects were generated before the study using a pseudo-random selection. The initial position of the virtual icosahedron was at 5cm along all axes measured from the bottom front corner of the Voronoi cube. If virtual reference objects were present for the iteration, they would be located at the correct relative position indicated by their physical counterparts.

A spatial augmented reality (SAR) calibration guidance tool was used to position the physical objects in 3D space by the examiner within the reference scene every iteration [5]. The desired height position of each object was indicated using a wireframe representation of these models to ensure that they were in the correct position and facing the correct way.

Participants were given three practice iterations of each task, where they were given instructions and guidance before data collection began. Following the practice iterations, participants were presented with 10 iterations for each x-ray visualization effect. Five of these would include two extra reference objects, and five would only include the icosahedron but no reference objects, presented in random order. After completing the ten iterations, the participants answered a questionnaire about the visualization technique which included a System Usability Survey [6], the PAAS subjective rating scale [25], and several other custom questions. This procedure was repeated in a random order for each x-ray visualization and the baseline condition.

5 Results

The study followed a within-group design, with a single group of participants completing all conditions (presence of reference objects and x-ray visualizations).

5.1 Placement Accuracy Results (H.1 & H.2)

A linear mixed effect model was used because it accounts for the random effect from our participants [17]. The model was specified with the factors of the presence of virtual reference objects and x-ray visualization effect techniques to analyze the accuracy of the user's placement of the virtual icosahedron object as compared to the target object's position, with a random effect of participant on the intercept. Significance values were extracted using Type II Wald chi-square tests. This model showed a significant fixed effect of the x-ray visualization effects ($\chi2(4, N{=}\ 20) = 13.897$, p

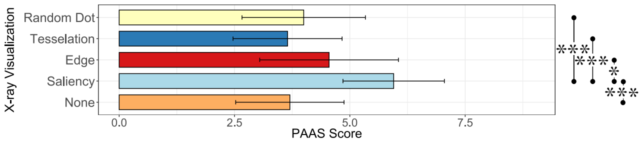

Figure 6: PAAS cognitive load results. Higher scores indicate a higher cognitive load required. Error bars indicate ± standard deviation ($*$ is $p < 0.5$, $**$ is $p < 0.01$, $***$ is $p < 0.001$).

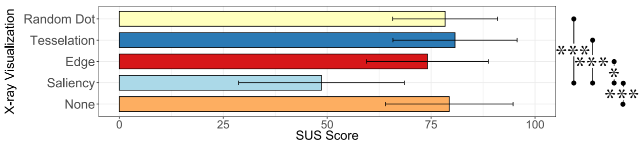

Figure 7: SUS results, higher scores indicate better usability. Error bars indicate ± standard deviation ($*$ is $p < 0.5$, $**$ is $p < 0.01$, $***$ is $p < 0.001$).

= 0.007) and the presence of the additional reference objects ($\chi 2(1, N = 20) = 77.641$, $p < 0.0001$), with no significant interaction effect. To further evaluate these findings posthoc pairwise comparisons using Tukey's HSD for multiple comparisons showed significantly improved accuracy when additional reference objects were present ($p < 0.0001$, df = 989, t = 8.771) and the Saliency visualization showed significantly lower accuracy than Tessellation (p = 0.0414, df = 989, t = 2.8). Saliency was also significantly less accurate than None (p = 0.0358, df = 989, t = -2.853).

5.2 Subjective Results (H.3)

The PAAS results showed a significant difference between the x-ray visualizations using a Friedman rank sum test ($\chi 2(4, N = 20) = 41.185$, $p < 0.0001$). Post-hoc analysis with pairwise Wilcoxon signed-rank tests was conducted with a Bonferroni correction applied, comparisons showed significantly increased cognitive load between Saliency and Random Dot (p = 0.00028), Tessellation (p = 0.00001), Edge (p = 0.02077), and None (p = 0.00001).

The SUS results showed a significant difference between the x-ray visualizations using a Friedman rank sum test $\chi 2(4, N = 20) = 45.234$, $p < 0.0001$. Post-hoc analysis with pairwise Wilcoxon signed-rank tests was conducted with a Bonferroni correction applied, resulting in a significantly lower usability score for Saliency than Random Dot ($p = 0.00020$), Tessellation ($p = 0.00018$), Edge ($p = 0.00119$), and None ($p = 0.00018$).

6 Discussion

The results in Section 5.1 did not support H.1 since Tessellation and None achieved better accuracy than Saliency. All other results were not statistically significant, but the results in Figure 4 show that all other visualizations outperformed Saliency slightly in terms of accuracy of placement. This is likely due to the amount of partial occlusion that this effect caused making it difficult to place objects in the near field [30].

The presence of reference objects had a large impact on the placement of the virtual objects since a significant difference was found in placement accuracy, as reported in Section 5.2. In Figure 5 there is an improvement in accuracy that is universal for all visualizations. This result does not support H.2, as the improvement applies to all visualizations. This result led us to conjecture that the presence of reference objects in OST AR x-ray vision is as efficient as with no visualization. This may indicate that the relationship between the background and the virtual objects has a much greater impact than the x-ray visualization in regards to depth perception.

We hypothesized that participants would find Saliency hard to use (H.3), and the results from our qualitative data (reported in Section 5.2) seem to support this hypothesis. Most visualizations had a relatively similar PAAS score across all activities as seen in Figure 6, and this was also reflected in the SUS shown in Figure 7. We believe that most participants disliked Saliency because it was too occlusive for how close they were to the object making the small virtual shapes hard to see and may have caused some frustration. However, we should note that this observation was not universal, as one participant stated "It was easier to position the objects accurately as opposed to the others" about Saliency. Some participants reported that Random Dot and Saliency blocked their vision too much, which may have been the reason for its lack of popularity and may hint that certain x-ray visualizations are preferable for different tasks and at different distances.

7 Limitations and Future Considerations

This study has several limitations that may have impacted this work. The number of participants recruited, considering the high number of conditions, reduces the strength of the study findings. In addition, the gender disparity could have influenced some of our findings. Our Saliency implementation took in the recommendations from Santos et al.'s [30] prior work on illumination and opacity, which was focused on Mobile VST AR displays, and it is possible that our system would produce different results than theirs.

This system utilizes a software-based workaround for this applying the visualizations over someone's sight. Ideally, the video overlay system would be implemented into the hardware itself as manufacturers could precisely calibrate integrated cameras to the lenses of the display. Similarly, integrated support for video overlay could reduce the latency between the frame being captured and rendered. The HoloLens 2 has a relatively long and varied delay when accessing the camera feed directly and displaying them to the user. Such a system could be possible on a custom-built device but is not currently possible on commercially available systems. Another method that will be considered in future studies is mounting the camera in a stationary position and applying the same or similar techniques for testing Saliency-based and Edge-Based visualizations or static texture renderings. This would reduce the impact of head motion and latency, however, the camera would need to be set up such that the user doesn't occlude the camera's view.

8 Conclusion

This research has shown that the x-ray visualization effects that work on VST AR devices can be transferred to OST AR devices, and it highlighted that they may not be as effective as existing OST AR visualizations. Visualizations like Saliency were much less accurate and they were harder to use in the near field than other techniques when used like tessellation which can accommodate the limitations of the system more easily. Our findings revealed that x-ray vision in OST AR devices may be made significantly stronger by using a back plate and some extra 3D objects for reference than when using a sophisticated x-ray vision effect.

Acknowledgments

Research supported by an Australian Government Research Training Program Scholarship. The authors would like to thank the participants for their time and feedback.

References

[1] B. Avery, C. Sandor, and B. H. Thomas. Improving Spatial Perception for Augmented Reality X-Ray Vision. In *2009 IEEE Virtual Reality Conference*, pp. 79–82, 2009. doi: 10.1109/VR.2009.4811002

[2] B. Avery, B. H. Thomas, and W. Piekarski. User evaluation of see-through vision for mobile outdoor augmented reality. *Proceedings - 7th*

IEEE International Symposium on Mixed and Augmented Reality 2008, ISMAR 2008, pp. 69–72, 2008. doi: 10.1109/ISMAR.2008.4637327

[3] M. Bajura, H. Fuchs, and R. Ohbuchi. Merging virtual objects with the real world: seeing ultrasound imagery within the patient. *Computer Graphics (ACM)*, 26(2):203–210, 1992. doi: 10.1145/142920.134061

[4] M. Berning, D. Kleinert, T. Riedel, and M. Beigl. A study of depth perception in hand-held augmented reality using autostereoscopic displays. *ISMAR 2014 - IEEE International Symposium on Mixed and Augmented Reality - Science and Technology 2014, Proceedings*, pp. 93–98, 2014. doi: 10.1109/ISMAR.2014.6948413

[5] O. Bimber and R. Raskar. *Spatial Augmented Reality Merging Real and Virtual Worlds*. A K Peters Ltd., Wellesley, 1 ed., 2004.

[6] J. B. Brooke. Sus: A 'quick and dirty' usability scale. 1996.

[7] D. Brown. Decentering Distortion of Lenses - The Prism Effect Encountered in Metric Cameras can be Overcome Through Analytic Calibration. *Photometric Engineering*, 32(3):444–462, 1966.

[8] G. Casiez, N. Roussel, and D. Vogel. 1€ filter: A simple speed-based low-pass filter for noisy input in interactive systems. *Conference on Human Factors in Computing Systems - Proceedings*, pp. 2527–2530, 2012. doi: 10.1145/2207676.2208639

[9] J. P. de Villiers, F. W. Leuschner, and R. Geldenhuys. Centi-pixel accurate real-time inverse distortion correction. *Optomechatronic Technologies 2008*, 7266(1):726611, 2008. doi: 10.1117/12.804771

[10] N. Elmqvist, U. Assarsson, and P. Tsigas. Employing dynamic transparency for 3D occlusion management: Design issues and evaluation. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 4662 LNCS(PART 1):532–545, 2007. doi: 10.1007/978-3-540-74796-3_54

[11] J. L. Gabbard, G. M. Fitch, and H. Kim. Behind the glass: Driver challenges and opportunities for AR automotive applications. *Proceedings of the IEEE*, 102(2):124–136, 2014. doi: 10.1109/JPROC.2013.2294642

[12] S. Ghasemi, M. Otsuki, P. Milgram, and R. Chellali. Use of Random Dot Patterns in Achieving X-Ray Vision for Near-Field Applications of Stereoscopic Video-Based Augmented Reality Displays. *Presence: Teleoperators & Virtual Environments*, 26(1):42–65, 2017. doi: 10.1162/PRES

[13] R. Gruen, E. Ofek, A. Steed, R. Gal, M. Sinclair, and M. Gonzalez-Franco. Measuring System Visual Latency through Cognitive Latency on Video See-Through AR devices. *Proceedings - 2020 IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2020*, pp. 791–799, 2020. doi: 10.1109/VR46266.2020.1580498468656

[14] U. Gruenefeld, Y. Brück, and S. Boll. Behind the Scenes: Comparing X-Ray Visualization Techniques in Head-mounted Optical See-through Augmented Reality. *ACM International Conference Proceeding Series*, pp. 179–185, 2020. doi: 10.1145/3428361.3428402

[15] I. Hamadouche. *Ilyas Hamadouche Augmented Reality X-ray vision on Optical See-Through Head Mounted*. PhD thesis, 2018.

[16] D. Kalkofen, E. Mendez, and D. Schmalstieg. Focus and Context Visualization for Medical Augmented Reality Focus and Context Visualization for Medical Augmented Reality. *ISMAR '07: Proceedings of the 2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality*, 6:1–10, 2007.

[17] M. Kaptein. Using Generalized Linear (Mixed) Models in HCI. In J. Robertson and M. Kaptein, eds., *Modern Statistical Methods for HCI*, pp. 251–274. Springer International Publishing, Cham, 2016. doi: 10.1007/978-3-319-26633-6_11

[18] M. Landvoigt and J. Cardoso. Analysis about publications on Facebook pages: finding of important characteristics. (February):8, 2017. doi: 10.1145/123

[19] M. Lerotic, A. J. Chung, G. Mylonas, and G. Z. Yang. Pq-space based non-photorealistic rendering for augmented reality. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 4792 LNCS, pp. 102–109, 2007. doi: 10.1007/978-3-540-75759-7_13

[20] A. Martin-Gomez, J. Weiss, A. Keller, U. Eck, D. Roth, and N. Navab. The Impact of Focus and Context Visualization Techniques on Depth Perception in Optical See-Through Head-Mounted Displays. *IEEE Transactions on Visualization and Computer Graphics*, XX(X):1–16, 2021. doi: 10.1109/TVCG.2021.3079849

[21] D. Medeiros, M. Sousa, D. Mendes, A. Raposo, and J. Jorge. Perceiving depth: Optical versus video see-through. *Proceedings of the ACM Symposium on Virtual Reality Software and Technology, VRST*, 02-04-Nove:237–240, 2016. doi: 10.1145/2993369.2993388

[22] M. Otsuki and K. Hideaki. Effect of translucent random dot mask on depth perception in stereo AR environment. *Research report Human-computer interaction (HCI) (Research Report Human-Computer Interaction (HCI))*, 2016(2):1–8, 2016.

[23] M. Otsuki, Y. Kamioka, Y. Kitai, M. Kanzaki, H. Kuzuoka, and H. Uchiyama. Please show me inside: Improving the depth perception using virtual mask in stereoscopic AR. *SIGGRAPH Asia 2015 Emerging Technologies, SA 2015*, (February 2018):2–5, 2015. doi: 10.1145/2818466.2818469

[24] E. Ozgur, A. Lafont, and A. Bartoli. Visualizing In-Organ Tumors in Augmented Monocular Laparoscopy. *Adjunct Proceedings of the 2017 IEEE International Symposium on Mixed and Augmented Reality, ISMAR-Adjunct 2017*, pp. 46–51, 2017. doi: 10.1109/ISMAR-Adjunct.2017.30

[25] F. G. Paas. Training Strategies for Attaining Transfer of Problem-Solving Skill in Statistics: A Cognitive-Load Approach. *Journal of Educational Psychology*, 84(4):429–434, 1992. doi: 10.1037/0022-0663.84.4.429

[26] N. Phillips, B. Kruse, F. A. Khan, and J. E. S. Ii. *Window for Law Enforcement Operations*. Springer International Publishing, 2020. doi: 10.1007/978-3-030-49695-1

[27] J. P. Rolland and H. Fuchs. Optical Versus Video See-Through Head-Mounted Displays in Medical Visualization. *Presence: Teleoperators and Virtual Environments*, 9(3):287–309, 2000. doi: 10.1162/105474600566808

[28] C. Sandor, A. Cunningham, A. Dey, and V. V. Mattila. An augmented reality X-ray system based on visual saliency. *9th IEEE International Symposium on Mixed and Augmented Reality 2010: Science and Technology, ISMAR 2010 - Proceedings*, pp. 27–36, 2010. doi: 10.1109/ISMAR.2010.5643547

[29] M. E. C. Santos, I. de Souza Almeida, G. Yamamoto, T. Taketomi, C. Sandor, and H. Kato. Exploring legibility of augmented reality X-ray. *Multimedia Tools and Applications*, 75(16):9563–9585, 2016. doi: 10.1007/s11042-015-2954-1

[30] M. E. C. Santos, M. Terawaki, T. Taketomi, G. Yamamoto, and H. Kato. Development of handheld augmented reality X-Ray for K-12 settings. *Lecture Notes in Educational Technology*, (9783662444467):199–219, 2015. doi: 10.1007/978-3-662-44447-4_11

[31] J. E. Swan, G. Singh, and S. R. Ellis. Matching and Reaching Depth Judgments with Real and Augmented Reality Targets. *IEEE Transactions on Visualization and Computer Graphics*, 21(11):1289–1298, 2015. doi: 10.1109/TVCG.2015.2459895

[32] T. Tsuda, H. Yamamoto, Y. Kameda, and Y. Ohta. Visualization methods for outdoor see-through vision. *ACM International Conference Proceeding Series*, 157:62–69, 2005. doi: 10.1145/1152399.1152412

[33] J. M. Van Waveren. The asynchronous time warp for virtual reality on consumer hardware. *Proceedings of the ACM Symposium on Virtual Reality Software and Technology, VRST*, 02-04-Nove:37–46, 2016. doi: 10.1145/2993369.2993375

[34] J. C. Vishton and P.M. Perceiving Layout and Knowing Distances : The Integration, Relative Potency, and Contextual Use of Different Information about Depth. In *Perception of Space and Motion*, vol. 22, pp. 69–117. 1995. doi: 10.1016/B978-012240530-3/50005-5

[35] Zhengyou Zhang. A Flexible New Technique for Camera Calibration Zhengyou. *Technical Report MSR TR-98-71 Microsoft*, (last updated on Aug. 13, 2008), 1998.

[36] S. Zollmann, R. Grasset, T. Langlotz, W. H. Lo, S. Mori, and H. Regenbrecht. Visualization Techniques in Augmented Reality: A Taxonomy, Methods and Patterns. *IEEE Transactions on Visualization and Computer Graphics*, 2626(c):1–20, 2020. doi: 10.1109/TVCG.2020.2986247

[37] S. Zollmann, D. Kalkofen, E. Mendez, and G. Reitmayr. Image-based ghostings for single layer occlusions in augmented reality. *9th IEEE International Symposium on Mixed and Augmented Reality 2010: Science and Technology, ISMAR 2010 - Proceedings*, pp. 19–26, 2010. doi: 10.1109/ISMAR.2010.5643546