Evaluating Depth Perception of Photorealistic Mixed Reality Visualizations for Occluded Objects in Outdoor Environments

Arindam Dey* Andrew Cunningham† Christian Sandor‡

Magic Vision Laboratory
School of Computer and Information Science
University of South Australia

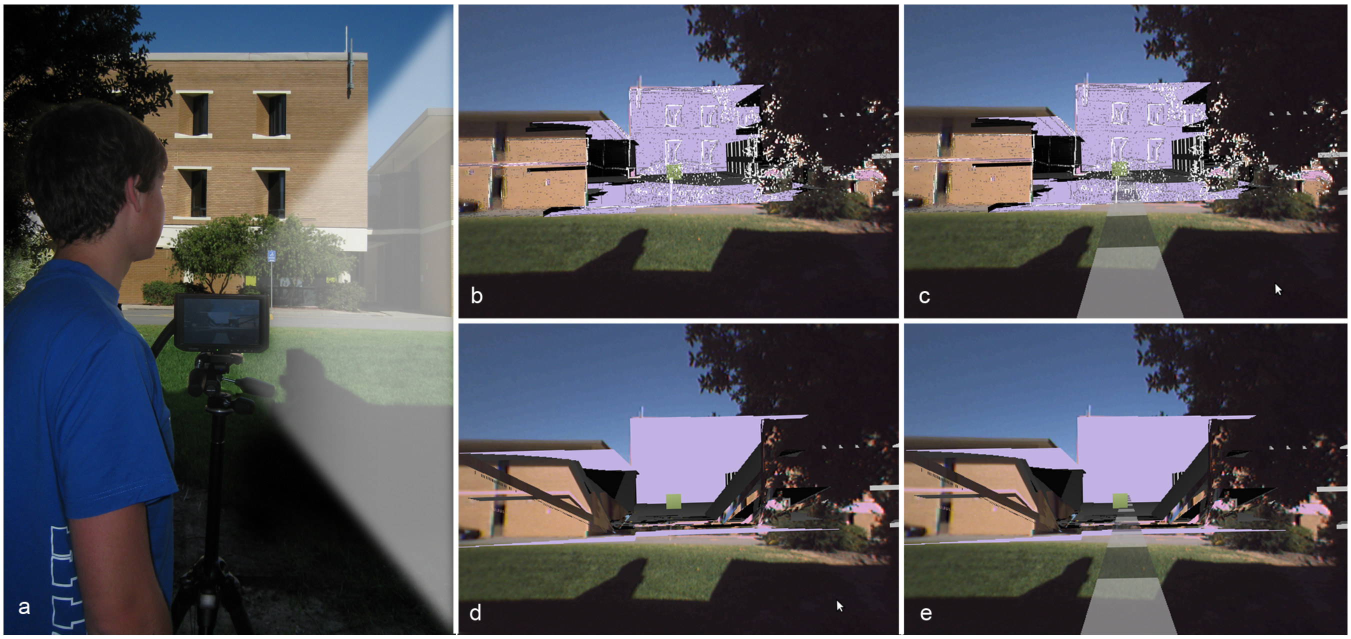

Figure 1: *Our experimental setup and conditions. (a) shows our experimental setup and a user performing the task. The top row shows our X-ray visualization without (b) and with (c) a depth cue. The bottom row shows our Melt visualization without (d) and with (e) a depth cue.*

Abstract

Enabling users to accurately perceive the correct depth of occluded objects is one of the major challenges in user interfaces for Mixed Reality (MR). Therefore, several visualization techniques and user evaluations for this area have been published. Our research is focused on photorealistic X-ray type visualizations in outdoor environments. In this paper, we present an evaluation of depth perception in far-field distances through two photorealistic visualizations of occluded objects (X-ray and Melt) in the presence and absence of a depth cue. Our results show that the distance to occluded objects was underestimated in all tested conditions. This finding is curious, as it contradicts previously published results of other researchers. The Melt visualization coupled with a depth cue was the most accurate among all the experimental conditions.

CR Categories: H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems—Artificial, Augmented, and Virtual Realities; H.5.2 [User Interfaces]: Ergonomics, Evaluation / Methodology—Screen Design, Style Guides; H.1.2 [Models and Principles]: User/Machine Systems—Human factors

Keywords: Evaluation, Mixed Reality, Augmented Reality, Depth Perception, Photorealistic Visualization, Depth Cues, Handheld Display

1 Introduction

Many interactive Mixed Reality (MR) visualizations have been developed to display occluded objects in outdoor environments, such as the previously described *Melt* visualization [Sandor et al. 2010] to display occluded objects by virtually melting the occluder, or the outdoor *X-ray* visualization [Avery et al. 2009], where the edges of occluding objects were highlighted in the video image to provide a realistic sense of occlusion. We categorize these techniques as *photorealistic visualizations*, as they attempt to realistically portray the occluded and occluder objects, creating a natural blend between the real and virtual space of the outdoor MR environment. Visualizing occluded objects is a challenging task, since the distance of virtual

*e-mail: Arindam.Dey@postgrads.unisa.edu.au
†e-mail: Andrew.Cunningham@postgrads.unisa.edu.au
‡e-mail: christian@sandor.com

objects is often misperceived. Numerous studies, for example [Livingston et al. 2003], have investigated this phenomenon through X-ray vision using head-mounted displays (HMDs). However, the visualization techniques in these studies have neither been photorealistic, nor have they explored the effects of handheld displays on depth perception. While HMDs have been extensively used in MR research and are suitable for many applications, they are currently costly and cumbersome, whereas handheld displays are not [Wagner and Schmalstieg 2006]. We assert that the use of handheld displays and improved visualizations can accelerate the acceptance of mixed reality by end users.

Keeping this assertion in mind, we have developed an MR prototype that purposefully uses a handheld display. With this prototype, we have implemented a Melt and an X-ray visualization to display occluded objects, along with a graphical depth cue, and compared depth perception under these conditions at far-field distances in an outdoor environment. Notably, we found that depth is underestimated (see Figure 2), contradictory to previous studies. Our photorealistic Melt visualization improved participants performance in distance estimation over X-ray, especially when the graphical cue was present. Through subjective responses we found that participants generally preferred Melt over X-ray for the task at hand.

We present our experiment in six sections. Section 2 discusses previous related research. Section 3 describes the detailed design of the experiment and our hypotheses. In the two subsequent sections, we present a thorough analysis of the collected data and detailed discussion of the results. Our concluding section points out future research directions.

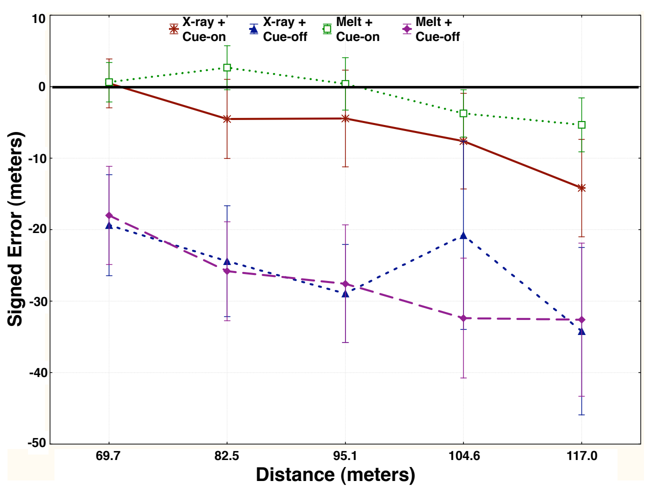

Figure 2: *Our experiment revealed a consistent underestimation of distances in an outdoor MR environment. The thick line at 0.0 denotes veridical perception (*± 0.95 *confidence interval is shown).*

2 Related Work

Depth perception is based on the interpretation of one or more available depth cues. Ten depth cues are especially important [Howard and Rogers 2002]: binocular disparity, binocular convergence, accommodative focus, atmospheric haze, motion parallax, linear perspective and foreshortening, occlusion, height in the visual field, shading, and texture gradient. In real-world environments, some or all of these cues are available to aid in determining distance. In virtual and MR environments however, not all of these cues are present due to the display technology used (being monocular, low field-of-view, or low resolution) and unnatural perceptual circumstances (such as X-ray vision [Swan et al. 2007]). In MR, various artificial depth cues have been proposed to aid distance estimation such as ground grid [Tsuda et al. 2005], tunnel cut-out [Avery et al. 2009], opacity [Livingston et al. 2003], shadow planes[Wither and Hollerer 2005], color encoded markers[Wither and Hollerer 2005], and edge map[Livingston et al. 2003]. Wither and Höllerer [2005] evaluated a set of pictorial depth cues for outdoor mobile augmented reality for absolute and relative depth perception.

There have been some studies performed where distance perception through X-ray vision was measured. Livingston et al. [2003] presented an evaluation of various methods of representing virtual occluded objects placed in three different layers at far-field distances. They studied the effect of drawing style, opacity, intensity, ground plane, and stereo on depth judgment. They found improved performance in decreasing opacity and intensity coupled with a "wire-fill" drawing style. However, they did not record the perceived distance but the layer in which the virtual object was placed. Avery et al. [2008] presented an evaluation of X-ray vision in an outdoor environment against an indoor video sequence of the same occluded locations and found X-ray to be a better option in terms of accuracy and response time. Tsuda et al. [2005] conducted a subjective evaluation of five visualization methods of occluded objects in an outdoor environment using a handheld display and found the combination of a ground grid with an overlaid model of the occluding object and top-down view to be optimal.

There have been studies where depth perception in MR was evaluated in indoor environments [Jones et al. 2008; Messing and Durgin 2005; Interrante et al. 2004; Plumert et al. 2005; Knapp and Loomis 2004; Livingston et al. 2003; Swan et al. 2007; Swan et al. 2006; Livingston et al. 2005]. These studies employed various tasks to measure perceived distances, such as a perceptual matching task [Swan et al. 2006], visually directed walking [Knapp and Loomis 2004; Messing and Durgin 2005], and verbal response [Knapp and Loomis 2004; Messing and Durgin 2005]. Most of the studies reported depth underestimation, where the distance was incorrectly perceived to be less than the actual distance. Swan et al. [2006] interestingly reported that egocentric distance is overestimated beyond around 23 meters in indoor environments. However, the majority of these studies were performed in indoor setups using HMDs. Livingston et al. [2009] compared distance estimation aided with two different depth cues—gridpoint and tramline—in both an indoor and an outdoor environment using HMDs. Contrary to results in indoor environments, they found that distance is overestimated in outdoor environments in medium to far-field distances. We took inspiration for our research from this work, however we focus on depth perception of occluded scenes in outdoor MR environments.

3 Experiment

We have investigated photorealistic visualizations for viewing occluded objects in outdoor MR. Previously, both an X-ray technique [Avery et al. 2009] and a virtual Melt technique [Sandor et al. 2010], which combine a real-world environment with a virtual replica of the occluded environment have been developed. The effects of these photorealistic visualizations on depth perception have not been investigated in detail.

Our design goal throughout the experiment is to study the effect that an MR visualization of an occluded scene has on depth perception. We intend to explore this in far-field distances, as these are most applicable to the intended use cases of the visualizations (i.e. standing across the street from a building).

3.1 Experimental Task

Participants stood in front of a display mounted on a tripod facing a building at 29 meters away (Figure 1a). Each participant completed 30 trials, where in each trial they were presented with one of the two photorealistic occluded object visualizations: X-ray or Melt. These visualizations reveal a virtual rendering of a real-world occluded target some distance behind the building front. We used a photo-realistic, correctly-scaled model of the occluded area and deliberately removed the textures from the virtual rendering of the model to reduce the chance that the prior knowledge of the environment would influence the participant's perception. Between participants, a graphical depth cue, as described in Section 3.3, was either presented or not.

Initially, participants were informed of (a) the exact distance from their current position to the building front and (b) the constraint of target objects being on the ground plane. The experimenter started a stopwatch in the experimental software at the beginning of each trial and stopped the watch when participants uttered "done" or "OK". Participants then reported the perceived distance.

In a training session, participants were informed about the usage of the system and the experimental task. None of the participants reported any difficulty in understanding or performing the task.

The visualizations were presented on a 7" screen with 640 × 480 resolution attached to a laptop with a 2.4 GHz Intel core 2 duo processor, 1 GB RAM, and nVidia GeForce 8600M GT graphics card. The laptop was placed on a table alongside another laptop that captured the participant's responses and dependent variables. The screen was mounted on a tripod with a fixed height of 1.5 meters and in a fixed orientation.

3.2 Participants

Twenty participants were recruited from the student population of our university and the general population through advertisement, having ages ranging from 18 to 31 years (M=25, SD=3.8). All of them had normal or corrected to normal vision. Two participants were female and 18 were male. Six of the participants had previous experience with MR and some of the participants played computer games regularly. We equally distributed participants into two matched groups—A and B. Participation was voluntary as no monetary or other benefits were provided.

3.3 Independent Variables

This experiment was based on four independent variables (Table 1). All variables were within subject except the Depth Cue and Participant group.

- **Visualization** $\in$ { X-ray, Melt} *within subjects*
  We implemented an X-ray visualization following Avery et al. [2009] in our mobile MR setup. This X-ray detects the prominent edges of the occluder and render them as thin white lines, and displays virtual representations of the real-world occluded objects through these detected edges (Figure 3). This edge overlay provides a depth cue as well as conveying a realistic sense of occlusion.

  The Melt visualization [Sandor et al. 2010] virtually melts the occluder objects in the viewer's field of view (Figure 4). The melt volume is defined by a circle sector shaped volume originating from the viewer in the direction of the point of interest (POI). Unlike X-ray, this Melt visualization can melt multiple occluding layers and reveal the occluded objects directly without having any occluders in between.

  Both visualizations were animated to help participants maintain context between the virtual and the real imagery. The X-ray visualization progressively fades the edge overlay, while cutting away at the virtual scene until the target was visible. The Melt visualization progressively melts the occluder until it reached the ground plane. Both animations took half a second to reveal the target object.

- **Graphical Cue** $\in$ {On, Off} *between subjects*
  It has been shown that distances in virtual environments are commonly underestimated and can be addressed by rendering graphical cues [Surdick et al. 1997]. We provide either no cue or a graphical cue as a between subjects variable since we aimed to precisely measure the effect of this cue on depth estimation, without any skill transfer. Our graphical cue is adapted from Livingston et al. [2009], as we recognize that, it may be applicable to outdoor MR as well. The cue is rendered as a semi-transparent segmented path originating from the user in the direction of the target. Each segment is 10 meters in length and alternates between black and white in color. We informed the properties of the graphical cue to participants and expected them to count the sections to more accurately judge the distance of the targets.

  During a pilot study the graphical depth cue was set as a randomized within subjects variable. From that study we found that subjects would base all consecutive depth estimations on their initial estimation of the first target presented. This means that a participant presented initially with the graphical cue would be consistently more accurate than one without, skewing the results. Therefore, we balanced the experiment by making the graphical cue a between subjects variable. By a coin toss group A was selected to work on the Cue-on condition and group B on the Cue-off condition.

- **Distance** $\in$ {69.7, 82.5, 95.1, 104.6, 117.0} *within subjects*
  Our MR prototype is intended to be used to visualize distant POIs in outdoor environments so it was required to choose longer distances than shorter. The target objects were placed at five different far-field distances from the participants' position.

- **Repetition** $\in$ {1 to 3} *within subjects*
  A same set of ten trials were repeated three times for each participant resulting them to perform thirty different trials with one target object in each of them.

Table 1: *Independent variables.*

Name	No. of Levels	Description
Visualization	2	X-ray, Melt
Depth Cue	2	On, Off
Distance	5	69.7m, 82.5m, 95.1m, 104.6m, 117.0m
Repetition	3	1, 2, 3

3.4 Dependent Variables

Four quantitative variables (see Table 2) were derived from the responses of the participants along with two subjective measurements. Accuracy was measured as a percentage of the actual distance, as determined by the following equation:

$$accuracy = \left(1 - \left|\frac{PD - AD}{AD}\right|\right) \times 100\% \quad (1)$$

Figure 3: *X-ray visualization: the occluder's (a) edges are overlaid on the occluded object (b) to provide X-ray vision (c).*

Figure 4: *Melt visualization: the occluded object is revealed by virtually melting the occluder (a-d).*

where *PD* was the participant's perceived distance and *AD* was the actual distance. Signed error (*SE*) was measured as the difference between the perceived distance and actual distance of virtual objects in meters. Hence, a positive *SE* indicates an overestimation of depth, while a negative *SE* indicates underestimation. We have also measured Absolute Error as $|SE|$. Time taken by the participants to respond on each of the trials was recorded in milliseconds. After the experiment participants were asked to report their experience in a NASA TLX form [Hart and Staveland 1988] to record subjective task load. They were also given a subjective questionnaire to provide qualitative feedback on the visualizations.

3.5 Controlled Variables

During the experiment the following variables were carefully controlled without affecting the experiment's generalizability.

Target object — Shape, Color and Number: All trials contained only one target object of identical shape and color. We selected a green cube with side lengths of 3 meters as the target object. Initially, the target was colored red for being distinguishable from other colors. After running a pilot study, we found this color to be misleading as it appeared to "pop" out of the environment, making users wrongly perceive the distance. We settled upon the neutral green color after testing several variations through an expert study.

Effect of sunlight and brightness: To control sunlight reflection into the participant's eye from the handheld display, and to make the display more legible, we performed the experiment in a shaded area. Before each session we adjusted the brightness of the screen depending on each individual participant's needs.

Movement of the display: Participants were prevented from moving the screen. This controlled setup helped us to explicitly identify the effects of the experimental variables without any other confounding factors, such as registration errors.

3.6 Experimental Design

In this mixed design experiment each participant experienced both visualizations across all five distances, achieving ten unique treatments per participant. Each treatment was repeated in a randomized order three times, resulting in 30 trials per participant. The graphical depth cue was treated as a between-subjects condition, with ten participants experiencing the graphical cue (Cue-on) and ten participants without a graphical cue (Cue-off). The total experiment resulted in 2(*visualizations*) × 5 (*distances*) × 2 (*graphical cues*) × 10 (*participants*) × 3(*repetitions*) = 600 data points.

3.7 Hypotheses

Before conducting the experiment we had the following hypotheses:

[H1] Distance will be overestimated in all of the conditions. This hypothesis is supported by the findings of Livingston et al. [2009].

[H2] Our X-ray creates a certain amount of occlusion from its edge detection rendering. This visual noise will cause negative effects on performance.

[H3] Melt eliminates all of the occlusions in the scene, therefore it will perform better than X-ray; both in terms of accuracy and response time.

[H4] The presence of the graphical depth cue will improve accuracy but, it will take participants longer to respond as they will have to count the segments of the depth cue. However, with increasing trials response time will decrease and accuracy will increase.

4 Results

The collected raw data was prepared and analyzed using the Statistica and SPSS analysis package. We describe our analysis in five

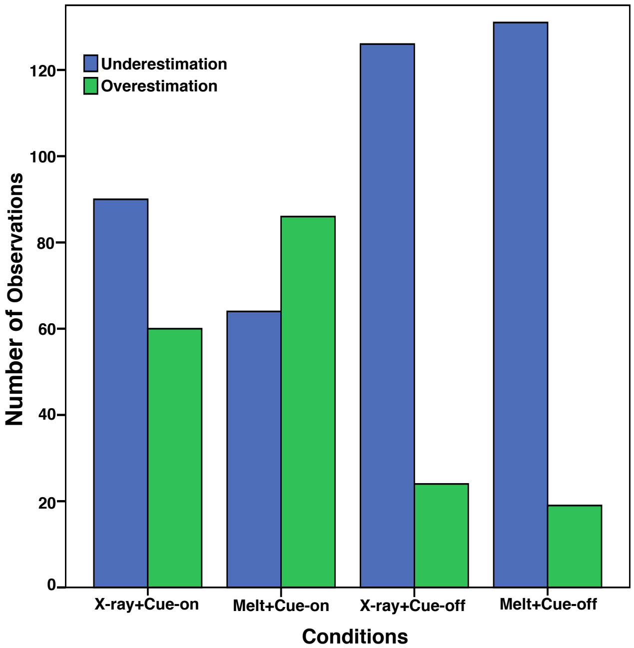

Figure 5: *Number of observations of distance estimation shows that, unlike others conditions, the Melt+Cue-on caused more over-estimations than underestimations.*

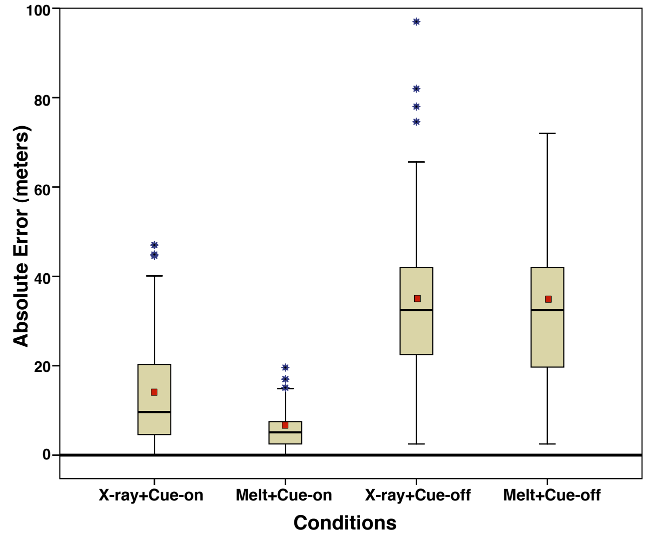

Figure 6: *Absolute error showed that Melt+Cue-on condition was the best among all condition tested (±1 standard error is shown). Thick black line at 0.0 shows the veridical perception, red squares indicate mean, and blue stars indicate outliers.*

different sections. The next section discusses results of signed error, followed by absolute error, accuracy, and response time in the three subsequent sections. The last section discusses subjective analysis.

4.1 Signed Error

Distance was mostly underestimated in all conditions, which contradicts previous studies which reported that distance is commonly overestimated in an outdoor environment [Livingston et al. 2009]. We discuss this observation in Section 5. The graphical cue significantly reduced error in both Melt and X-ray ($p<0.001$). Two separate ANOVAs revealed a main effect of distance on the signed error for both Melt ($F(4,295)=2.4389$, $p<0.05$) and X-ray ($F(4,295)=3.1249$, $p=0.016$). Tukey's HSD post-hoc test revealed significant ($p<0.03$) differences between 117.0 meters and (a) 69.7 meters, and (b) 82.5 meters for both visualizations. We found that the number of observations where an underestimation occurred was more than overestimations in each experimental condition except Melt+Cue-on, where most of the observations were overestimations (Figure 5). Overall, there were 411 underestimations, compared to 189 overestimations. A chi-square test indicated a significant relationship between conditions and depth estimation ($\chi^2(3,N=600)$=92.774, $p<0.001$). The effect size was .393.

4.2 Absolute Error

In the case of absolute error, we crossed visualization and depth cue to create four unique conditions and ran an one-way ANOVA with these conditions being independent factors—$F(3,596)=147.935$, $p<.001$ (see Figure 6). A Tukey's HSD test showed that the Melt+Cue-on condition was significantly ($p<.001$) better than all other conditions, including X-ray+Cue-on. However, there were no significant differences between Melt and X-ray in the Cue-off conditions. Through another factorial ANOVA, with distance, depth cue, and visualization being independent factors, we also found significant ($p<.05$) interaction effects between (a) visualization and depth cue, and (b) distance and depth cue.

4.3 Accuracy

An analysis of accuracy showed that our graphical depth cue aided participants to perceive distance more accurately than without the cue. We ran a one-tailed t-test to analyze differences between the means for the X-ray and Melt conditions. We found that X-ray+Cue-on was significantly ($p<0.001$) more accurate than X-ray+Cue-off, similarly Melt+Cue-on was significantly ($p<0.001$) more accurate than Melt+Cue-off. Interestingly, we found that in the presence of the graphical cue, Melt was significantly ($p<0.001$) more accurate than X-ray and, with increasing distance and the graphical cue, the accuracy of Melt stayed constant, while X-ray lost accuracy (Figure 7). We attribute this result to the visual noise created by our X-ray (hypothesis *H3*).

4.4 Response Time

In this experiment we found that the graphical cue caused a delayed response from participants. This makes sense, since participants had to count the segments of the cue to accurately interpret the distance. The mean response time of the Cue-on was significantly ($p<0.001$) higher than Cue-off condition for both X-ray and Melt visualizations (see Figure 8). In the case of the Cue-on condition Melt was significantly ($p<0.001$) faster then X-ray. This is also consistent with our hypothesis (*H3*). We predict that the low visibility due to the edge overlay of our X-ray visualization made it harder to count the segments of the graphical cue.

We also noted that response time reduced significantly with later trials in the case of X-ray+Cue-off condition — $F(14,135)=5.57$, $p<0.001$ (Figure 9). It was clear that in this condition, participants gave up in later trials and replied unexpectedly faster than earlier trials. In the case of Melt there was a similar trend of reduced response time with later trials, however, this was not statistically significant ($p<0.08$).

We observed that in the presence of the graphical cue, participants

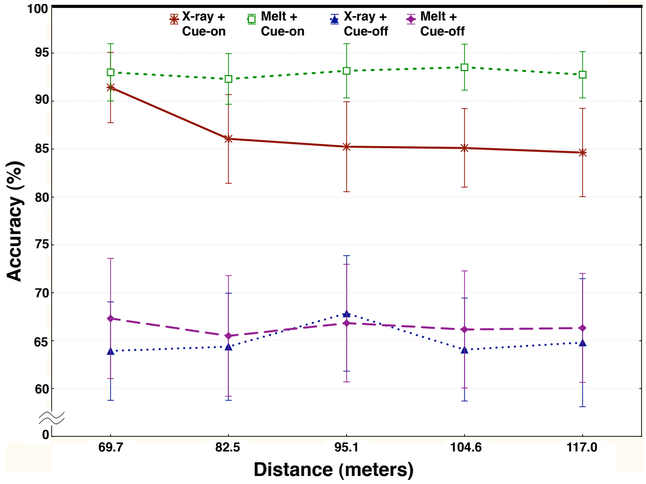

Figure 7: *Accuracy shows an expected and significant improvement with the graphical cue conditions across both photorealistic visualizations (±0.95 confidence interval is shown).*

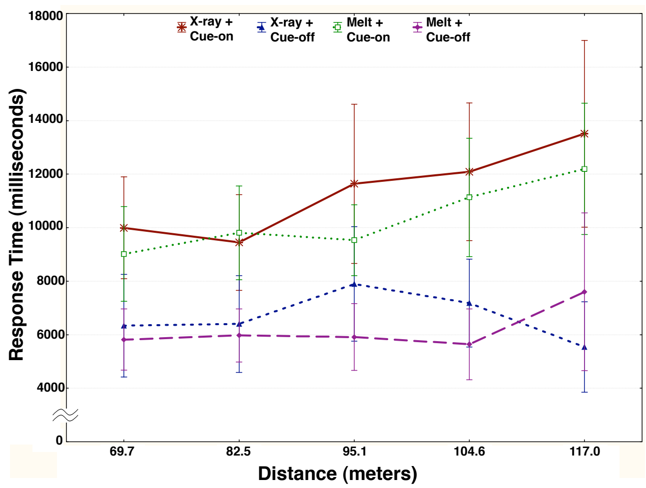

Figure 8: *Response time at different distances in different experimental conditions with ±0.95 confidence interval.*

would explicitly count the segments of the cue in earlier trials, whereas in later trials they did not count all the segments and replied faster while still maintaining accuracy (Figure 10). This indicates that participants made a mental model of the occluded area and primarily relied on the size of the target object to determine its distance. It also suggests a learning effect. However, it was not significant with the number of observations we studied.

4.5 NASA TLX and Subjective Analysis

The evaluation of NASA Task Load Index showed that participants rated the Cue-on (M=42.5, SD=15.3) and the Cue-off (M=41.9, SD=13.2) similarly. It is hard to conclude anything from this small difference but we noted that most of the participants indicated "mental demand" as the biggest contributor to the task load in both conditions.

We asked participants to rate the visualizations depending on how confident they were in depth judgment, visibility of the target object, and usability of the visualizations, on a scale of 100. We

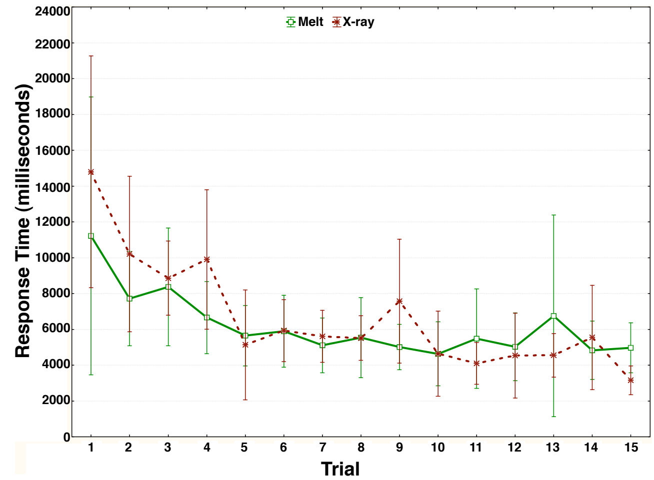

Figure 9: *Average response time at different trials in Cue-off conditions with ±0.95 confidence interval.*

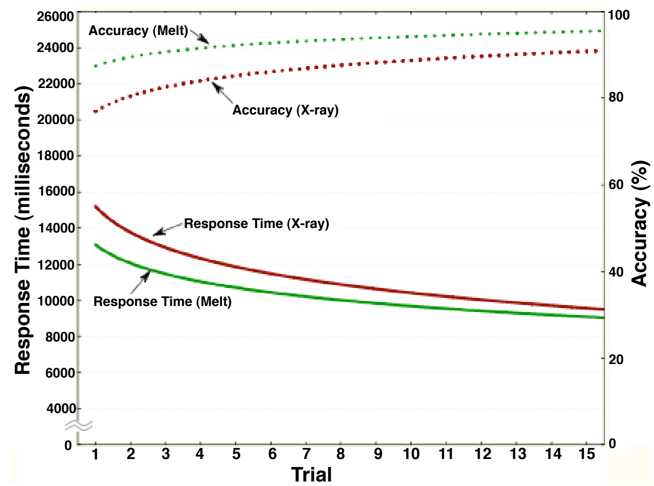

Figure 10: *The trend lines show that, in the Cue-on condition response time decreased with increasing trial but a certain level of accuracy was maintained across the trials for both of the visualizations.*

conducted a paired two tailed T-test to analyze the responses for these three aspects. Overall, Melt rated significantly higher than X-ray for all aspects (Figure 11). Participants were significantly ($p<0.001$) more confident in distance estimation with Melt (M=69.5, SD=20.6) than X-ray (M=46, SD=15.4). Visibility rated significantly ($p<0.001$) higher for Melt (M=80.5, SD=11.0) than X-ray (M=42, SD=24.6), and usability was rated significantly ($p<0.001$) higher for Melt (M=69.5, SD=19.0) than X-ray (M=55.5, SD=16.1).

As a reason, some of the participants indicated that the edge overlay made the target object less visible, especially in the case of distant objects (supporting hypothesis *H3*). S6 suggested to implement an edge overlay with semi-transparent lines. We also noted that participants responded faster in later trials than earlier trials. As a reason for this, S12 indicated that "…after about 10 trials I felt a training effect".

Table 2: *Measurement, Mean (M), and Standard Deviation (SD) of quantitative dependent variables.*

Conditions	Dependent Variables							
	Signed Error (±meters)		Absolute Error (meters)		Accuracy (%)		Response Time (msec)	
	M	SD	M	SD	M	SD	M	SD
X-ray + Cue-on	-6.05	16.54	13.01	11.84	86.50	11.78	11337.60	7088.70
Melt + Cue-on	-1.07	9.31	6.59	6.64	92.96	7.05	10338.42	5262.54
X-ray + Cue-off	-25.53	26.13	32.79	16.02	64.99	15.36	6672.57	4945.52
Melt + Cue-off	-27.27	22.70	31.52	16.23	66.43	16.14	6191.57	4525.69

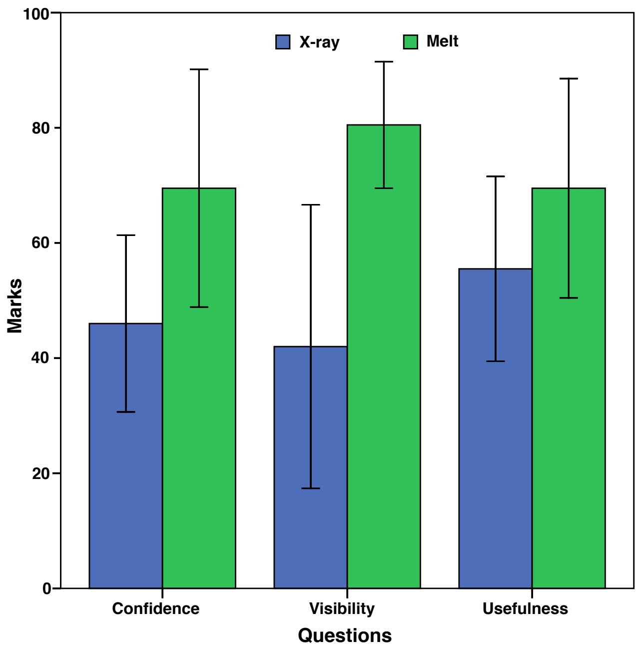

Figure 11: *Subjectively, participants preferred our Melt visualization over our X-ray implementation for visibility, usability and confidence (whiskers represent ±1 standard deviation).*

5 Discussion

The most important result of this experiment is the observation of an underestimation of the distance of occluded objects in outdoor environments, which refutes our hypothesis *H1*, and seemingly contradicts the findings of Livingston et al. [2009]. We can imagine five possible reasons for this contradiction: use of a non-near eye display, use of a video see-through setup, use of our MR system, and the dense edge overlay of our X-ray visualization.

- **Use of a non-near-eye display:** Unlike HMDs, the 7" display used in our study provided only a small, low-resolution window to the augmented world, reducing immersion. This might be the reason for the distance underestimation.
- **Use of a video see-through setup:** We assume another reason for the underestimation is that our experiment was conducted in a video see-through setup, whereas Livingston et al. [2009] conducted their experiment in an optical see-through setup.
- **Use of our MR system:** The effect of depth-compression is well known in virtual reality systems. We believe that our MR setup is more akin to a virtual reality system than Livingston et al's [2009], considering that participants were required to judge depth primarily in a rendering of a virtual representations of the occluded objects. This would result in the underestimation.
- **Longer experimental distances:** The distances used in our experiment ranged between 69.7 meters and 117.0 meters. However, the range of distances experimented in Livingston et al's [2009] experiment was between 4.83 meters and 38.64 meters. We assume use of longer egocentric distances caused underestimation. This finding also leads to an interesting future work—to find out the distance where outdoor depth perception switches from overestimation to underestimation. We assume that distance to be somewhere between 38 meters to 70 meters in MR environments.
- **The dense edge overlay of our X-ray visualization:** In the case of X-ray, the dense edge overlay adversely affected visibility of the target object, which we assume to be a factor causing the underestimation (though this assumption requires further experimentation). In the case of Melt, aided with the graphical depth cue, there was less underestimation than overestimation. On average, participants began underestimation after a distance of 100 meters, and judged distance almost accurately before that (Figure 2). This is because the scene was clear of any occlusion in the case of Melt and segments on the graphical cue were clearly legible up until the 10th segment (i.e. 100 meters).

However, we are unable to explicitly identify the effect of the above factors had on distance estimation. We assume these reasons might have both a combined effect as well as individual effects on the result.

Though the graphical cue was more accurate, it adversely effected response time. While this supports our Hypothesis *H2*, we acknowledge that a more efficient graphical cue may be designed. Participant feedback supported this notion. One of the participants suggested to add a distinguishable color to improve the cue.

As we predicted the remaining visual noise in the case of X-ray visualization (especially in the Cue-on condition) adversely effected the response time and accuracy in distance perception. This effect was very clear, since participants complained about the white edges over the target objects. This supports our Hypothesis *H3*.

Melt provided a clear scene and made the target object completely visible, which resulted in significantly positive results in favor of Melt over X-ray, both in terms of accuracy and response time. This finding was consistent with our Hypothesis *H4*. However, in the case of the Cue-off condition the results were not significant. We assume that the reason of this insignificance is that subjects did not make a real effort to guess the distance correctly, especially in later trials, as they had no cue to guide them other than the size of the object.

6 Future Work and Conclusion

This experiment was one of the first efforts to evaluate photorealistic visualizations for distance perception in MR. Contradicting expectations based on previous research, we observed distance underestimation in an outdoor environment. We plan to verify the validity of our results with a more extensive user study and aim to contrast this result with a HMD. In this experiment, we deliberately mounted the display on a tripod to ameliorate the confounding effect of tracking errors. We consider to conduct another study using a freely-movable handheld display in the future. We also aim to investigate, when depth overestimation switches to underestimation in an outdoor MR environment. If such an effect is not found, it can be inferred that use of longer distances was not a cause of underestimation in our current experiment. From the results we found that the edge overlay of X-ray caused negative effects; it will be interesting to determine the optimum level of edge overlay versus performance in a future study. We have also found that the graphical depth cue aided performance. In future studies we plan to compare various synthetic depth cues available for MR environments.

We have learned that there are fundamental differences between depth perception in non-near eye displays and HMDs. We plan to isolate the main reason for the different result to Livingston et al. We are confident that we are on the right track to investigate handheld MR, and we hope our findings will encourage other researchers to investigate handheld displays in mixed and augmented reality environments.

Acknowledgments

The authors wish to thank anonymous participants for their voluntary participation. Rhys Moyne, Graeme Jarvis, and Thahn Nguyen for their enormous support during the experiment.

References

AVERY, B., THOMAS, B. H., AND PIEKARSKI, W. 2008. User evaluation of see-through vision for mobile outdoor augmented reality. In *ISMAR '08: Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*, IEEE Computer Society, Washington, DC, USA, 69–72.

AVERY, B., SANDOR, C., AND THOMAS, B. H. 2009. Improving spatial perception for augmented reality x-ray vision. In *VR '09: Proceedings of the IEEE Conference on Virtual Reality*, IEEE, 79–82.

HART, S., AND STAVELAND, L. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Human mental workload 1*, 139–183.

HOWARD, I., AND ROGERS, B. 2002. *Seeing in Depth*, vol. 2. I Porteous, University of Torento Press.

INTERRANTE, V., ANDERSON, L., AND RIES, B. 2004. An experimental investigation of distance perception in real vs. immersive virtual environments via direct blind walking in a high-fidelity model of the same room. In *APGV '04: Proceedings of the 1st Symposium on Applied Perception in Graphics and Visualization*, ACM, New York, NY, USA, 162–162.

JONES, J. A., SWAN, J., SINGH, G., KOLSTAD, E., AND ELLIS, S. R. 2008. The effects of virtual reality, augmented reality, and motion parallax on egocentric depth perception. In *APGV '08: Proceedings of the 5th Symposium on Applied Perception in Graphics and Visualization*, ACM, New York, NY, USA, 9–14.

KNAPP, J., AND LOOMIS, J. 2004. Limited field of view of head-mounted displays is not the cause of distance underestimation in virtual environments. *Presence: Teleoperators & Virtual Environments 13*, 5, 572–577.

LIVINGSTON, M. A., SWAN, J., GABBARD, J. L., HÖLLERER, T. H., HIX, D., JULIER, S. J., BAILLOT, Y., AND BROWN, D. 2003. Resolving multiple occluded layers in augmented reality. In *ISMAR '03: Proceedings of the 2nd IEEE/ACM International Symposium on Mixed and Augmented Reality*, IEEE Computer Society, Washington, DC, USA, 56.

LIVINGSTON, M., ZANBAKA, C., SWAN, J., AND SMALLMAN, H. 2005. Objective measures for the effectiveness of augmented reality. *VR '05: Proceedings of IEEE Conference on Virtual Reality (Poster Session)*, 287–288.

LIVINGSTON, M. A., AI, Z., SWAN, J., AND SMALLMAN, H. S. 2009. Indoor vs. outdoor depth perception for mobile augmented reality. In *VR '09: Proceedings of the IEEE Conference on Virtual Reality*, IEEE Computer Society, Washington, DC, USA, 55–62.

MESSING, R., AND DURGIN, F. H. 2005. Distance perception and the visual horizon in head-mounted displays. *ACM Transactions on Applied Perception 2*, 3, 234–250.

PLUMERT, J. M., KEARNEY, J. K., CREMER, J. F., AND RECKER, K. 2005. Distance perception in real and virtual environments. *ACM Transactions on Applied Perception 2*, 3, 216–233.

SANDOR, C., CUNNINGHAM, A., ECK, U., URQUHART, D., JARVIS, G., DEY, A., BARBIER, S., MARNER, M., AND RHEE, S. 2010. Egocentric space-distorting visualizations for rapid environment exploration in mobile mixed reality. In *VR 2010: Proceedings of IEEE Conference on Virtual Reality*, 47–50.

SURDICK, R., DAVIS, E., KING, R., AND HODGES, L. 1997. The perception of distance in simulated visual displays: A comparison of the effectiveness and accuracy of multiple depth cues across viewing distances. *Presence: Teleoperators and Virtual Environments 6*, 5, 513–531.

SWAN, J. I., LIVINGSTON, M. A., SMALLMAN, H. S., BROWN, D., BAILLOT, Y., GABBARD, J. L., AND HIX, D. 2006. A perceptual matching technique for depth judgments in optical, see-through augmented reality. In *VR '06: Proceedings of the IEEE Conference on Virtual Reality*, IEEE Computer Society, Washington, DC, USA, 19–26.

SWAN, J., JONES, A., KOLSTAD, E., LIVINGSTON, M., AND SMALLMAN, H. 2007. Egocentric depth judgments in optical, see-through augmented reality. *IEEE Transactions on Visualization and Computer Graphics 13*, 3, 429–442.

TSUDA, T., YAMAMOTO, H., KAMEDA, Y., AND OHTA, Y. 2005. Visualization methods for outdoor see-through vision. In *ICAT '05: Proceedings of the International Conference on Augmented Tele-Existence*, ACM, New York, NY, USA, 62–69.

WAGNER, D., AND SCHMALSTIEG, D. 2006. Handheld augmented reality displays. In *VR '06: Proceedings of the IEEE Conference on Virtual Reality*, IEEE Computer Society, Washington, DC, USA, 321.

WITHER, J., AND HOLLERER, T. 2005. Pictorial depth cues for outdoor augmented reality. In *ISWC '05: Proceedings of the Ninth IEEE International Symposium on Wearable Computers*, IEEE Computer Society, Washington, DC, USA, 92–99.