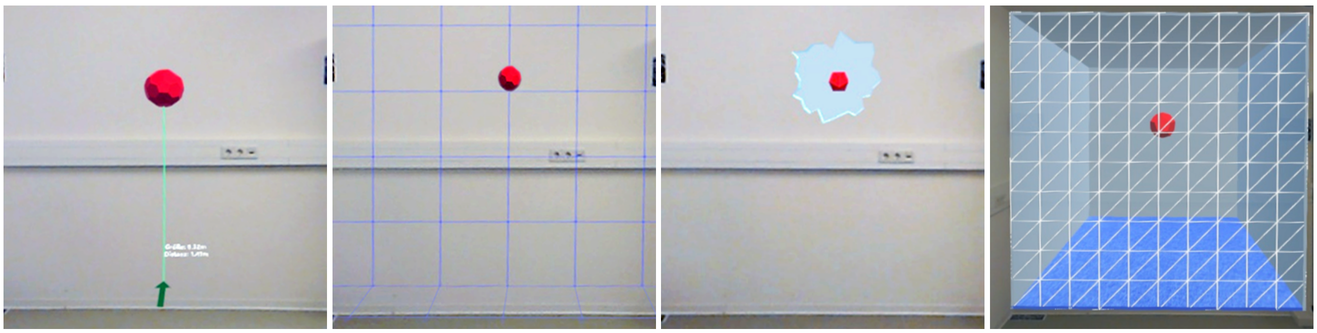
# Behind the Scenes: Comparing X-Ray Visualization Techniques in Head-mounted Optical See-through Augmented Reality

Uwe Gruenefeld
University of Duisburg-Essen
Essen, Germany
uwe.gruenefeld@uni-due.de

Yvonne Brück
University of Oldenburg
Oldenburg, Germany
yvonne.brueck@uol.de

Susanne Boll
University of Oldenburg
Oldenburg, Germany
susanne.boll@uol.de



**(a) *Baseline* technique.** **(b) *Grid* technique.** **(c) *Cutout* technique.** **(d) *Wireframe* technique.**

**Figure 1: The compared techniques for visualizing occluded objects in Augmented Reality. Each technique as seen from the user's perspective through the Hololens (captured with built-in RGB camera and merged with displayed AR content).**

## ABSTRACT

Locating objects in the environment can be a difficult task, especially when the objects are occluded. With Augmented Reality, we can alternate our perceived reality by augmenting it with visual cues or removing visual elements of reality, helping users to locate occluded objects. However, to our knowledge, it has not yet been evaluated which visualization technique works best for estimating the distance and size of occluded objects in optical see-through head-mounted Augmented Reality. To address this, we compare four different visualization techniques derived from previous work in a laboratory user study. Our results show that techniques utilizing additional aid (textual or with a grid) help users to estimate the distance to occluded objects more accurately. In contrast, a realistic rendering of the scene, such as a cutout in the wall, resulted in higher distance estimation errors.

## CCS CONCEPTS

• **Human-centered computing → Mixed / augmented reality**; *Visualization techniques*; *User studies*.

## KEYWORDS

optical see-through, head-mounted, augmented reality, occlusion, occluded objects, x-ray, visualization techniques

**ACM Reference Format:**

Uwe Gruenefeld, Yvonne Brück, and Susanne Boll. 2020. Behind the Scenes: Comparing X-Ray Visualization Techniques in Head-mounted Optical See-through Augmented Reality. In *19th International Conference on Mobile and Ubiquitous Multimedia (MUM 2020), November 22–25, 2020, Essen, Germany*. ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3428361.3428402

## 1 INTRODUCTION

Many of us dream of having supernatural abilities. In particular, people often imagine what it would feel like to have enhanced visual perception. A source of inspiration are fictional superheros such as Superman[1] who demonstrate the potential that x-ray vision offers. Having x-ray vision allows one to spot occluded objects, by removing objects from reality or making them see-through. Removing objects from reality is often referred to as Diminished Reality (DR). Combined with recent technological advances in tracking occluded objects [4], Augmented Reality can make the dream of x-ray vision a reality. Possessing this ability may be beneficial in various scenarios. For example, x-ray vision can help someone to spot a lost object (e.g., a keychain) in the next room, can enable one to see the locations of underground water pipes [35], or can assist a surgeon who needs to see inner organs during an operation [27].

With the latest improvements in Augmented Reality (AR) technology (e.g., rendering quality [33], refresh rate [20], or registration accuracy [26]), environments can be augmented to show additional information, to remove visual elements from the scene, or to do

[1] en.wikipedia.org/wiki/Superman, last retrieved October 13, 2020



*MUM 2020, November 22–25, 2020, Essen, Germany*



ACM ISBN 978-1-4503-8870-2/20/11…$15.00
https://doi.org/10.1145/3428361.3428402

**Table 1: Overview of visualization techniques for occluded objects in previous research.**

| Author | Reality | Form-factor | Display / Device | Year | Study | Task |
|---|---|---|---|---|---|---|
| Elmqvist et al., [7] | Virtual | desktop | widescreen LCD display | 2007 | yes | find and count object(s) |
| Martin-Gomes et al., [22] | Virtual | head-mounted | AMOLED display / HTC Vive | 2019 | yes | align object in 6 DoF |
| Furmanski et al., [11] | Augmented | desktop | 17” flat panel display | 2002 | yes | estimate distance |
| Kalkhofen et al., [18] | Augmented | desktop | not specified | 2009 | no | - |
| Zollmann et al., [35] | Augmented | desktop | not specified | 2014 | yes | rate depth percep.; trace object |
| Tsuda et al., [34] | Augmented | handheld | 5" SVGA display / Sony Vaio U | 2006 | yes | rate perceptibility |
| Schall et al., [32] | Augmented | handheld | 4.5" LCD display / Sony Vaio UX | 2008 | yes | free exploration of visualization |
| Zollmann et al., [36] | Augmented | handheld | 12.1" LCD display / Motion J3400 | 2012 | yes | explore specific object |
| Colley et al. [6] | Augmented | handheld | Projector / Samsung Galaxy Beam | 2014 | yes | change occlusion viewport |
| Mori et al. [25] | Augmented | handheld | not specified / Laptop | 2017 | no | - |
| Eren and Balcisoy [9] | Augmented | handheld | IPS Display / LG Nexus 5X | 2018 | yes | judge vertical depth distance |
| Reiners et al. [30] | Augmented | head-mounted | video see-through / Virtual i-O glasses | 1999 | no | - |
| Livingston et al., [21] | Augmented | head-mounted | video see-through / Sony Glasstron | 2003 | yes | determine relative distance |
| Robertson et al. [31] | Augmented | head-mounted | video see-through / Sony Glasstron | 2008 | yes | place partially occluded objects |
| Avery et al., [1] | Augmented | head-mounted | video see-through / Sony Glasstron | 2009 | no | - |
| Lilija et al., [19] | Augmented | head-mounted | optical see-through / Hololens | 2019 | yes | perform occluded interactions |
| Gruenefeld et al. [14] | Augmented | head-mounted | optical see-through / Hololens | 2019 | yes | locate occluded objects |

both at the same time. As it is experienced in a head-mounted device, people can use such technologies hands-free and while mobile. This has advantages in many spatial working environments where machines have to be operated by hand, or in situations in which the user is moving. Furthermore, the head-mounted device allows one to visualize the locations of occluded objects from an egocentric perspective, requiring no additional views [8]. Thereby, the cognitive load required to mentally integrate the displayed information is minimal [3, 5, 23]. Head-mounted AR devices can also be combined with safety helmets that workers may be required to wear.

Several techniques have already been proposed for visualizing occluded objects in Augmented Reality (see Table 1). However, most of these techniques have been designed with form-factors such as desktop monitors or handheld devices in mind. Compared to head-mounted AR, these devices cannot utilize our binocular depth perception because they show the same visualization to both eyes. In the past, little research has been conducted with regard to head-mounted devices. Therefore, Mori et al. refer to optical see-through Diminished Reality as a “virtually unexplored area” [24]. When looking into previous work that addresses occluded objects in head-mounted AR, they either investigate video see-through devices utilizing a single camera [1, 21, 30, 31], or they focus on the interaction with occluded objects that are physically accessible (e.g., objects that can be touched [14, 19]). Moreover, to our knowledge, no previous work investigates different visualizations regarding perception of size in optical see-through AR.

Therefore, in this paper, we focus on occluded objects that cannot be visually perceived and are physically not accessible. In such cases, users must fully rely on the visualization of the occluded objects. It is unclear how well users can understand the spatial attributes of such objects. Therefore, we analyze previous work to derive promising techniques for visualizing occluded objects. We then implement and evaluate four visualization techniques in a user study, in which we ask participants to estimate the distance and size of occluded objects visualized in optical see-through AR.

## 2 RELATED WORK

*Depth Perception.* Images projected onto the retina are two-dimensional. The brain then builds a third dimension from the 2D images. To enable the brain to create the impression of depth, various available cues must contribute [12]. The optical axes of the left and right eyes always meet at a fixed point. From these axes, the brain determines a convergence angle that can be used as a measure for the distance of the fixed point [17]. Binocular disparity and convergence only work at a close range. From a distance of 6m, monocular criteria for depth perception are used, because from this distance the binocular disparity decreases to a negligible small value [17]. To benefit from the binocular features of optical see-through AR, we focus on occluded objects within a 6m range.

*Occlusion Visualization.* Elmqvist and Tsigas defined a taxonomy for 3D occlusion management [8]. They classified a total of 50 different techniques using their taxonomy and stated that virtual x-rays can be used very well for object detection. By selectively removing distractors that occlude the targets, access to the relevant objects is simplified. Another advantage of x-ray visualization is that an AR application requires only one view to display occluded objects. However, the selective removal of distractors that occlude the target reduces depth perception, which makes it more difficult to understand spatial relationships [8, 21]. Besides the environment, it is also important how the object itself is rendered. For example, when occluded objects are rendered as solid 3D objects, they are often perceived as being in front of the real-world objects [11]. Therefore, a rendering technique that shows less detail may be beneficial for perceiving the object as being behind the occlusion [10, 14, 21, 22]. Nevertheless, by diminishing parts of reality, objects can be renders as solid 3D objects while users still can understand spatial relationships (e.g., if a wall is in front of an object).

In the last decades, researchers have started exploring occlusion visualizations for Virtual and Augmented Reality. Both Virtual and Augmented Reality share the ability to alter human perception of the world by either immersing the user in a virtual environment or adding virtual objects to the real environment. The latter has great potential for visualizing occluded information in various real world scenarios [2]. Several form-factors have been explored in the past, from desktop monitors, to handheld devices such as smartphones or tablets, to head-mounted devices (see Table 1). Head-mounted AR devices in particular have several advantages, such as hands-free operation and users being mobile. An early occlusion visualization for head-mounted devices was presented by Reiners et al. [30]. Later more such visualizations followed. However, these visualizations are either evaluated with objects that are physically accessible (e.g., objects that can be touched [14, 19]), or investigated in video see-through devices utilizing a single camera [1, 21, 30, 31]. For this reason, Mori et al. refers to Diminished Reality as a "virtually unexplored" area [24].

## 3 VISUALIZATION

To investigate how to best visualize occluded objects in head-mounted Augmented Reality, we analyzed different visualization techniques described in previous work (see Table 1). Overall, we derived four different visualization techniques. All visualization techniques render the occluded object as a solid 3D object. Thereby, all available information about the 3D object is rendered and users may be able to perceive the object very well (cf. [18, 30, 32]). Nevertheless, the high level of detail results in the occlusion of other virtual or real objects in the environment. However, in this work, we focus on visualizing on a single occluded object.

*Baseline.* For the *Baseline* condition, we use a simple arrow spanning between the occluding environment (i.e., a wall in our case) and the occluded object with text that specifies the distance to the object and the size of the object (see Figure 1a). Hence, our baseline condition represents the upper-performance limit in our study. Furthermore, it unveils the accuracy possible for some of the measures applied in our user study, such as blind walking.

*Grid.* A ground grid can be useful for estimating the locations of occluded objects (cp. [34]). Not only can it help the user to estimate the position of a single object, it can also aid in their understanding of spatial relationships among multiple occluded objects. When an additional grid is placed perpendicular to the ground grid, it may also aid size estimations of occluded object (see Figure 1b). For this purpose, we render the perpendicular grid at the position of the occluded object, intersecting the object in its center. The *Grid* visualization does not require diminishment of occluding objects in the environment, since it lays the grid on top of the environment. Thereby, the user's view is only minimally obscured and all objects in the real environment are still visible.

*Cutout.* The *Cutout* visualization is inspired by many existing rendering techniques for occluded objects that have not yet been tested for distance and size estimation accuracy on head-mounted optical see-through AR displays (cf. [1, 7, 11, 34]). Furthermore, a similar technique is used in the built-in Hololens game Robo Raid[2]. Instead of altering large parts of the virtual environment (e.g., changing a complete wall in the scene [11]), the *Cutout* visualization makes the occluding environment see-through only where necessary. This can be done, for example, by using dynamic transparency [7] or a cutout visualization [1]. Thereby, the user can see the occluded object without losing the environment as a reference point to judge spatial relationships (see Figure 1c).

*Wireframe.* The most common render technique for the occluding environment is a see-through visualization (cf. [34]). Basically, real world objects are diminished to allow the user to look through them. This can be done using a transparency effect or by completely removing the objects from the scene. However, using too little transparency can increase the visual complexity of the scene [7]. On the other hand, using too much transparency can make it harder to perceive the occluding environment, which is disadvantageous for estimating spatial relationships [34]. Therefore, we propose a wireframe-based visualization that fully replaces the occluding environment (see Figure 1d), similar to [21].

## 4 USER STUDY

To investigate different techniques for visualizing occluded objects in head-mounted optical see-through Augmented Reality, we performed a laboratory user study.

### 4.1 Procedure

At the start of the study, the participants were introduced to Augmented Reality and the Microsoft Hololens head-mounted display. Afterwards, we calibrated the Hololens for each participant (individually adjusting the interpupillary distance). We tested four blocks, one for each visualization technique (see Section 3). All blocks were counterbalanced using a Latin-square design. Each block contained one training trial and four measured trials.

In each trial, participants wearing the Hololens stood facing a wall behind which the object was placed (see Figure 2). The occluded object was then shown on the Hololens using the visualization technique of that block. The participants were asked to specify the distance to the occluded object. To measure this distance, we used a combination of different measures because distance estimation is most important for locating occluded objects [14]. We asked participants to physically demonstrate the perceived distance (by the well-known measure of "blind walking" [16]) and to verbally specify it. Participants were informed that they stood exactly one meter away from the wall (see Figure 2). Standing in front of the wall behind which the occluded object is shown, the participants rotated 180 degrees to walk the distance in the open space of the room not obstructed by walls. To confirm the walked distance, the participants pressed a button on a controller. The object would then appear at the confirmed position in the room. Participants could then adjust the walked distance by demonstrating it another time if they wanted to. After the distance had been selected, the participants were asked to specify the size of the occluded object. For this purpose, the participants used the joystick on the controller to change the size of the displayed object in the room to match that of the occluded object (in 2cm steps). Participants were asked to fill out a Likert-items questionnaire at the end of each block. At the end of the experiment, they filled out a questionnaire to compare the four visualization techniques as well as a demographic questionnaire. Each participant took approximately 45 minutes to finish the experiment.

[2]www.microsoft.com/en-us/p/roboraid/9nblggh5fv3j, last retrieved October 13, 2020

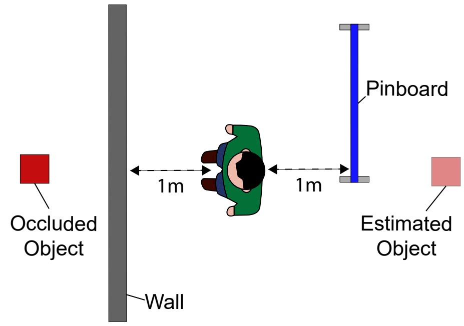



**Figure 2: Apparatus of the user study. *Best seen in color.***

## 4.2 Apparatus

In the experiment, the participant puts on the Microsoft Hololens and stands in front of a wall one meter away (see Figure 2). All generated occluded objects are located behind the wall. The participant can see the occluded objects by looking through the device. Behind the participant, the setup is mirrored, using a pinboard to represent the wall. Thereby, participants can walk to the location where they expect the object to be (in the mirrored setup). Using an xbox controller, participants could confirm actions and input the estimated object size. To input the size, participants can resize the presented object using the joystick on the controller. All visualization techniques were developed using Unity3D[3] and the Mixed Reality Toolkit[4]. We used image marker to measure the locations of the wall and floor (partially visible in Figure 1). After the initial calibration with the image marker, we used the spatial awareness system of the Hololens to continuously track the environment. During the study, we used dialog boxes on the Hololens to inform participants about the next step.

## 4.3 Study Design

To explore different visualization techniques for occluded objects, we conducted a within-subjects controlled laboratory study in Augmented Reality with the Microsoft Hololens. Our independent variable was technique with four levels (*Baseline* vs. *Grid* vs. *Cutout* vs. *Wireframe*; see Figure 1). Each technique was tested in a block consisting of four measured trials with each trial showing a different object (dodecahedron, truncated octahedron, icosahedron, and truncated icosahedron; see Figure 3). We varied the objects' shape for each technique to ensure our results are generalizable to different shapes. The order in which the different object shapes were presented was randomly selected for each technique. All blocks, each consisting of one technique, were counterbalanced using a Latin-square design. In each trial, we randomly selected a distance between 1 to 5 meters and a object size between 10 to 40cm (in 2cm steps). We used quantitative methods to evaluate user performance, taking distance and size estimation error, and subjective Likert-items as our dependent variables.

For this study, we asked: **(RQ) Which of the techniques works best for visualizing the distance and size of occluded objects in optical see-through head-mounted Augmented Reality?**

[3]unity.com, last retrieved October 13, 2020

[4]microsoft.github.io/MixedRealityToolkit-Unity, last retrieved October 13, 2020

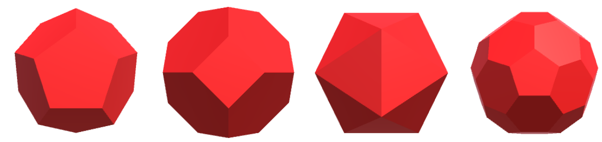


**Figure 3: Object shapes tested in study.**

$H_1$ For the distance estimation, we expect our *Baseline* condition to perform best, followed by the *Grid* visualizations, with *Cutout* and *Wireframe* performing worst because they offer no additional (visual) support to determine the distance.

$H_2$ We expect the *Grid* visualization to perform better for the object size estimation than *Cutout* and *Wireframe* because the fixed grid size may help users to estimate the size of the occluded object more accurately.

## 4.4 Participants

We recruited 16 volunteer participants (5 female), aged between 24 and 50 years (M=30.38, SD=9.06). None suffered from color vision impairment, 10 had normal vision, and 6 had corrected-to-normal vision. We asked the participants to rate their level of experience with Augmented Reality on a 5-point Likert-scale. The participants on average reported limited prior experience with AR (Md=1, IQR=1) and head-mounted displays (Md=1.5, IQR=2).

## 4.5 Results

*Distance estimation error.* We asked participants to estimate the distance to the occluded object twofold: by the measure of "blind walking" [16] and by verbally estimating it. The medians (inter-quartile range) distance estimation errors for each measure are: blind walking=22.5cm (IQR=74.3cm) and verbal estimating=0cm (IQR=14cm). A Shapiro-Wilk-Test showed that both are not normally distributed ($p<0.001$). A Wilcoxon Signed-rank test showed a significant difference between both measures (W=26584, Z=8.5469, $p<0.001$, r=0.38) (r: $>0.1$ small, $>0.3$ medium, and $>0.5$ large effect). In the following, we analyzed the absolute estimation errors for both distance measures individually.

The medians (inter-quartile range) of the absolute distance estimation errors (blind walking) for each technique are: *Baseline*=27cm (24.5cm), *Grid*=37cm (IQR=41.5cm), *Cutout*=49cm (IQR=62.5cm), and *Wireframe*=48cm (IQR=60cm) (see Figure 4).

A Shapiro-Wilk-Test showed that the absolute distance estimation errors (blind walking) are not normally distributed ($p<0.01$). Therefore, we ran a Friedman test that revealed a significant effect of technique on absolute distance estimation error ($\chi^2(3)=11.47$, $p=0.009$, N=16). A post-hoc test using Wilcoxon Signed-rank with Bonferroni correction showed significant differences between some of the conditions (see Table 2). For distance estimation error (blind walking) we can conclude: *Baseline* $<$ *Cutout*, *Wireframe* and *Grid* $<$ *Wireframe*.

The medians (inter-quartile range) of the absolute distance estimation errors (verbal estimation) for each technique are: *Baseline*=0cm (IQR=0cm), *Grid*=4cm (IQR=10cm), *Cutout*=33.5cm (IQR=44.3cm), and *Wireframe*=50cm (IQR=59.5cm) (see Figure 5).

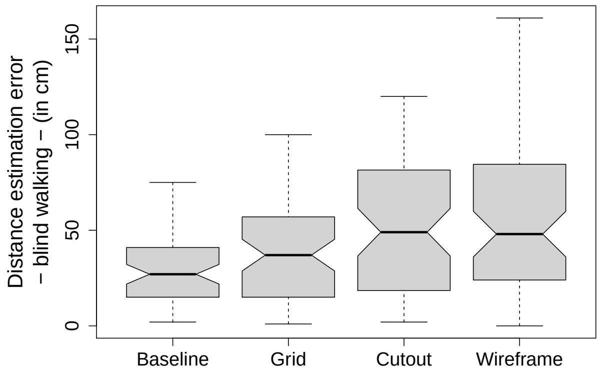



**Figure 4: Boxplots of absolute distance estimation error ("blind walking") for each technique.**

**Table 2: Pairwise comparisons of techniques with for the absolute distance estimation error (blind walking).**

| Comparison | | | W | Z | p | r |
|---|---|---|---|---|---|---|
| *Baseline* | vs. | *Grid* | 841 | -1.13 | 1 | 0.12 |
| *Baseline* | vs. | *Cutout* | 585 | -2.90 | 0.013 | 0.26 |
| *Baseline* | vs. | *Wireframe* | 610 | -2.54 | 0.042 | 0.22 |
| *Grid* | vs. | *Cutout* | 657 | -2.40 | 0.063 | 0.21 |
| *Grid* | vs. | *Wireframe* | 596.5 | -2.82 | 0.018 | 0.28 |
| *Cutout* | vs. | *Wireframe* | 945.5 | -0.23 | 1 | 0.02 |

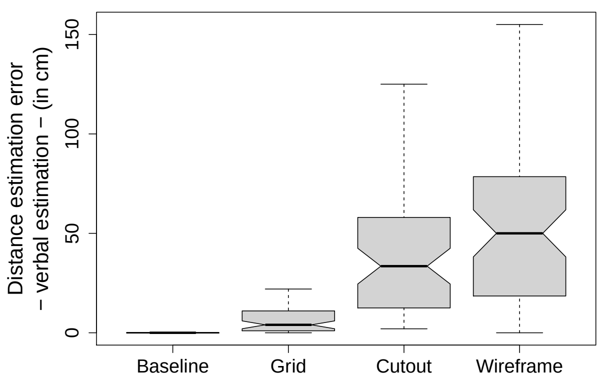



**Figure 5: Boxplots of absolute distance estimation error (verbal estimation) for each technique.**

A Shapiro-Wilk-Test showed that the absolute distance estimation errors (verbal estimation) are not normally distributed ($p<0.001$). Hence, we ran a Friedman test, which revealed a significant effect of technique on absolute distance estimation error ($\chi^2(3)=132.72$, $p<0.001$, N=16). A post-hoc test using Wilcoxon Signed-rank with Bonferroni correction showed significant differences between some of the conditions (see Table 3). For distance estimation error (verbal estimation) we can conclude: *Baseline* < *Grid* < *Cutout*, *Wireframe*.

*Size estimation error.* For the size estimation error, we determined the absolute difference between estimated size and size of the occluded object. The medians (inter-quartile range) of the absolute size estimation errors for each technique are: *Baseline*=2cm (IQR=4cm), *Grid*=2cm (IQR=4cm), *Cutout*=2cm (IQR=4.5cm), and *Wireframe*=4cm (IQR=4cm).

**Table 3: Pairwise comparisons of techniques for the absolute distance estimation error (verbal estimation).**

| Comparison | | | W | Z | p | r |
|---|---|---|---|---|---|---|
| *Baseline* | vs. | *Grid* | 0 | -6.92 | <0.001 | 0.61 |
| *Baseline* | vs. | *Cutout* | 0 | -6.96 | <0.001 | 0.61 |
| *Baseline* | vs. | *Wireframe* | 0 | -6.89 | <0.001 | 0.61 |
| *Grid* | vs. | *Cutout* | 207 | -5.57 | <0.001 | 0.49 |
| *Grid* | vs. | *Wireframe* | 168.5 | -5.83 | <0.001 | 0.52 |
| *Cutout* | vs. | *Wireframe* | 840 | -1.34 | 0.731 | 0.12 |

A Shapiro-Wilk-Test showed that the size errors are not normally distributed ($p<0.001$). Therefore, we ran a Friedman test, which revealed a significant effect of technique on absolute size estimation error ($\chi^2(3)=10.24$, p=0.017, N=16). A post-hoc test using Wilcoxon Signed-rank with Bonferroni correction showed significant differences between some of the conditions (see Table 4). For size estimation error we can conclude: *Grid* < *Baseline*, *Wireframe*.

**Table 4: Pairwise comparisons of techniques for the absolute size estimation error.**

| Comparison | | | W | Z | p | r |
|---|---|---|---|---|---|---|
| *Baseline* | vs. | *Grid* | 740 | 2.90 | 0.013 | 0.26 |
| *Baseline* | vs. | *Cutout* | 800.5 | 1.93 | 0.212 | 0.17 |
| *Baseline* | vs. | *Wireframe* | 522 | -0.88 | 1 | 0.08 |
| *Grid* | vs. | *Cutout* | 405 | -0.97 | 1 | 0.09 |
| *Grid* | vs. | *Wireframe* | 301.5 | -2.59 | 0.036 | 0.23 |
| *Cutout* | vs. | *Wireframe* | 501.5 | -2.26 | 0.093 | 0.20 |

*Correlation between distance and size estimation error.* We tested the Spearman's rank-order correlation coefficient between absolute distance estimation error (verbal estimation) and absolute size estimation error (see Figure 6). We found a moderate effect for the correlation of the two variables ($r_s$=0.46, $p<0.001$). With larger distance estimation errors, participants were more likely to also make larger size estimation errors. Specifically, participants made objects too large in size, if they perceived them as being father away.

*Subjective measures.* After each condition, we asked the participants to answer questions with 5-point Likert-scale items (1=strongly disagree, 5=strongly agree). Participants stated that they had a good understanding of distance with the techniques *Baseline* (Md=5, IQR=0) and *Grid* (Md=5, IQR=1), while they were neutral for *Cutout* (Md=3, IQR=1.5) and *Wireframe* (Md=3, IQR=1.25).

In addition, the participants indicated that each of the techniques successfully conveys the size of the occluded object: *Baseline* (Md=5, IQR=1), *Grid* (Md=4, IQR=1.5), *Cutout* (Md=4, IQR=2), and *Wireframe* (Md=4, IQR=2).

## 5 DISCUSSION

*Distance Estimation.* During the study, we asked participants to estimate the distance to the occluded object in two different ways:

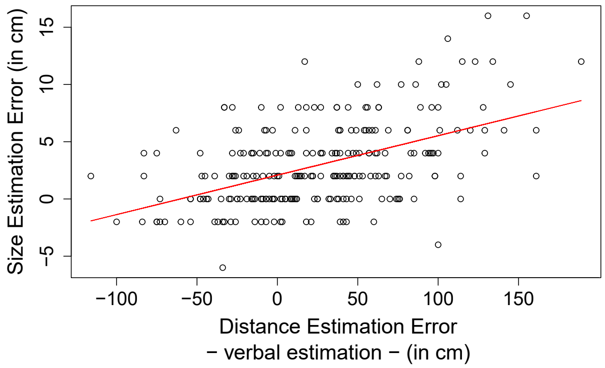



**Figure 6: Correlation between absolute distance estimation error (verbal estimation) and absolute size estimation error (size estimation in 2cm steps on xbox controller).**

by the measure of blind walking and by verbally estimating. Most participants tried to walk the distance via one-meter steps. However, the steps participants took were often shorter than one meter. This is in line with previous research that found that walked distances often do not match verbally-communicated estimations of that distance [28]. Participants tend to walk too far if the target is closer than 2m, but not far enough if it is farther than 2m away. For a distance of 6m, the error is about 84cm [29]. However, the environment in which one has to walk is also critical for determining the distance walked. For example, people make fewer errors when outdoors than when indoors because humans tend to reduce their step size when indoors [15]. Since our setup was indoors, the results reflect the fact that participants were better at verbally communicating the distance than walking it. Nevertheless, distance estimations for all techniques were affected equally by the measure, and we deployed the measure to ensure that participants could not simply repeat the displayed numbers (*Baseline* technique) or count the cells (*Grid* technique) while having problems to imagine how the values relate to the real world.

In our hypothesis $H_1$, we expected the *Baseline* visualization to perform better than the *Grid* visualization. We could proof this for the verbal estimation of the distance; however, it does not apply for the measure of blind walking. Therefore, we cannot accept our hypothesis $H_1$. Nevertheless, *Baseline* and *Grid* outperformed *Cutout* and *Wireframe* for both measures as hypothesized in $H_1$.

For our study, we can say that AR techniques like *Cutout* that look realistic in the sense that the visualization is part of the existing environment are outperformed by AR techniques that give additional (visual) aid such as *Baseline* and *Grid*. Therefore, we recommend using a *Grid*-based technique for visualizing occluded objects to best encode the spatial relationships between and distances to these objects. This is supported by the subjective measures, where the *Grid* technique received high ratings.

*Size Estimation.* Participants were able to estimate sizes of occluded objects similarly well for each technique. Nevertheless, we found that participants could estimate the object's size more precisely with the *Grid* technique than with *Baseline* and *Wireframe*.

Furthermore, our results show that the estimation of size becomes less accurate if the distance has already been determined inaccurately. This is because the two criteria for depth perception are interdependent. The perceived distance of an object impacts its perceived size [13]. Therefore, the perception of size is distorted when the perception of distance of the object is already incorrect.

In general, the size can be better estimated if there are multiple objects in the environment. In this study, only one object was displayed at a time. In real life, there may be many objects in one room. By evaluating several objects in a room, one can better perceive the size and distance of any single object. Here, previous knowledge about possible object constellations in the environment also plays a major role in solving ambiguity problems in the environment. Problems of ambiguity arise when multiple possible interpretations of the positions of objects in a room are generated by a two-dimensional retinal image [12].

*Limitations.* In the study, one occluded object was displayed to the participants. However, it is possible that multiple occluded objects in the field of view are relevant for the user. The position of the object was also only modified in terms of its distance to the user. This was done to eliminate influence of the object position on estimations of distance and size. The display of a single occluded object runs the risk of giving the user ambiguous impressions. On the other hand, if too many occluded objects are displayed, this may lead to an information overload so that the user no longer understands the spatial relationships of objects [21].

A further restriction is the maximum distance of the objects, which was five meters from the room wall (i.e. six meters from the participant). This maximum distance was adapted to the given room size and to the technical specifications of the Hololens as well as to the range within which humans have binocular depth perception [17]. Therefore, in the future, it should be evaluated whether the selected visualizations are suitable for longer distances.

Another aspect is that there was a limited number of object shapes (geometric bodies) and the shapes were similar to one other, so that we cannot assume our visualizations perform similarly well on any possible object shape. Nevertheless, using different object shapes in the study ensured that the results are not for one shape only and more generalizable towards different shapes.

## 6 CONCLUSION

In this paper, we first presented a comprehensive overview of researched techniques that allow one to visualize occluded objects. We identified that very little research has investigated occlusion visualizations in optical see-through head-mounted Augmented Reality so far. Hence, we evaluated four different techniques for visualizing occluded objects in head-mounted optical see-through Augmented Reality and compared them with regard to distance and size estimation error. Our results show that techniques with additional (visual) aid such as *Baseline* and *Grid* perform best for distance estimation. Furthermore, participants estimate objects as too large in size, if they perceive them as being father away.

## REFERENCES

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

[2] Ronald T. Azuma. 1997. A Survey of Augmented Reality. *Presence: Teleoperators and Virtual Environments* 6, 4 (1997), 355–385. https://doi.org/10.1162/pres.1997.6.4.355 arXiv:https://doi.org/10.1162/pres.1997.6.4.355
[3] Woodrow Barfield, Craig Rosenberg, and Thomass A Furness III. 1995. Situation awareness as a function of frame of reference, computer-graphics eyepoint elevation, and geometric field of view. *The International journal of aviation psychology* 5, 3 (1995), 233–256.
[4] Peter Barnum, Yaser Sheikh, Ankur Datta, and Takeo Kanade. 2009. Dynamic seethroughs: Synthesizing hidden views of moving objects. In *2009 8th IEEE International Symposium on Mixed and Augmented Reality*. 111–114. https://doi.org/10.1109/ISMAR.2009.5336483
[5] Andy Cockburn, Amy Karlson, and Benjamin B. Bederson. 2009. A Review of Overview+Detail, Zooming, and Focus+Context Interfaces. *ACM Comput. Surv.* 41, 1, Article 2 (Jan. 2009), 31 pages. https://doi.org/10.1145/1456650.1456652
[6] Ashley Colley, Olli Koskenranta, Jani Väyrynen, Leena Ventä-Olkkonen, and Jonna Häkkilä. 2014. Windows to Other Places: Exploring Solutions for Seeing through Walls Using Handheld Projection. In *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational* (Helsinki, Finland) *(NordiCHI '14)*. Association for Computing Machinery, New York, NY, USA, 127–136. https://doi.org/10.1145/2639189.2639226
[7] Niklas Elmqvist, Ulf Assarsson, and Philippas Tsigas. 2007. Employing Dynamic Transparency for 3D Occlusion Management: Design Issues and Evaluation. In *Human-Computer Interaction – INTERACT 2007*, Cécilia Baranauskas, Philippe Palanque, Julio Abascal, and Simone Diniz Junqueira Barbosa (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 532–545.
[8] Niklas Elmqvist and Philippas Tsigas. 2008. A Taxonomy of 3D Occlusion Management for Visualization. *IEEE Transactions on Visualization and Computer Graphics* 14, 5 (Sep. 2008), 1095–1109. https://doi.org/10.1109/TVCG.2008.59
[9] Mustafa Tolga Eren and Selim Balcisoy. 2018. Evaluation of X-ray visualization techniques for vertical depth judgments in underground exploration. *The Visual Computer* 34, 3 (01 Mar 2018), 405–416. https://doi.org/10.1007/s00371-016-13465
[10] Steven Feiner, Blair Macintyre, and Dorée Seligmann. 1993. Knowledge-based augmented reality. *Commun. ACM* 36, 7 (1993), 53–62.
[11] Chris Furmanski, Ronald Azuma, and Mike Daily. 2002. Augmented-Reality Visualizations Guided by Cognition: Perceptual Heuristics for Combining Visual and Obscured Information. In *Proceedings of the 1st International Symposium on Mixed and Augmented Reality (ISMAR '02)*. IEEE Computer Society, USA, 320.
[12] Simon Grondin. 2016. *Depth Perception.* Springer International Publishing, Cham, 103–122. https://doi.org/10.1007/978-3-319-31791-5_7
[13] Howard E. Gruber. 1954. The Relation of Perceived Size to Perceived Distance. *The American Journal of Psychology* 67, 3 (1954), 411–426. http://www.jstor.org/stable/1417933
[14] Uwe Gruenefeld, Lars Prädel, and Wilko Heuten. 2019. Locating Nearby Physical Objects in Augmented Reality. In *Proceedings of the 18th International Conference on Mobile and Ubiquitous Multimedia* (Pisa, Italy) *(MUM '19)*. Association for Computing Machinery, New York, NY, USA, Article 1, 10 pages. https://doi.org/10.1145/3365610.3365620
[15] Marco Iosa, Augusto Fusco, Giovanni Morone, and Stefano Paolucci. 2012. Walking there: environmental influence on walking-distance estimation. *Behavioural brain research* 226, 1 (2012), 124–132.
[16] F. E. Jamiy and R. Marsh. 2019. Distance Estimation In Virtual Reality And Augmented Reality: A Survey. In *2019 IEEE International Conference on Electro Information Technology (EIT)*. 063–068.
[17] James Kalat. 2015. *Biological psychology.* Nelson Education.
[18] Denis Kalkofen, Erick Mendez, and Dieter Schmalstieg. 2009. Comprehensible Visualization for Augmented Reality. *IEEE Transactions on Visualization and Computer Graphics* 15, 2 (March 2009), 193–204. https://doi.org/10.1109/TVCG.2008.96
[19] Klemen Lilija, Henning Pohl, Sebastian Boring, and Kasper Hornbundefinedk. 2019. Augmented Reality Views for Occluded Interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, Article 446, 12 pages. https://doi.org/10.1145/3290605.3300676
[20] Peter Lincoln, Alex Blate, Montek Singh, Turner Whitted, Andrei State, Anselmo Lastra, and Henry Fuchs. 2016. From Motion to Photons in 80 Microseconds: Towards Minimal Latency for Virtual and Augmented Reality. *IEEE Transactions on Visualization and Computer Graphics* 22, 4 (April 2016), 1367–1376. https://doi.org/10.1109/TVCG.2016.2518038
[21] Mark A Livingston, J Edward Swan, Joseph L Gabbard, Tobias H Hollerer, Deborah Hix, Simon J Julier, Yohan Baillot, and Dennis Brown. 2003. Resolving multiple occluded layers in augmented reality. In *The Second IEEE and ACM International Symposium on Mixed and Augmented Reality, 2003. Proceedings.* IEEE, 56–65.
[22] Alejandro Martin-Gomez, Ulrich Eck, and Nassir Navab. 2019. Visualization Techniques for Precise Alignment in VR: A Comparative Study. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 735–741. https://doi.org/10.1109/VR.2019.8798135
[23] Paul Milgram and Fumio Kishino. 1994. A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems* 77, 12 (1994), 1321–1329.
[24] Shohei Mori, Sei Ikeda, and Hideo Saito. 2017. A survey of diminished reality: Techniques for visually concealing, eliminating, and seeing through real objects. *IPSJ Transactions on Computer Vision and Applications* 9, 1 (June 2017), 17. https://doi.org/10.1186/s41074-017-0028-1
[25] Shohei Mori, Momoko Maezawa, and Hideo Saito. 2017. A work area visualization by multi-view camera-based diminished reality. *Multimodal Technologies and Interaction* 1, 3 (2017), 18.
[26] Alessandro Mulloni, Hartmut Seichter, and Dieter Schmalstieg. 2011. Handheld augmented reality indoor navigation with activity-based instructions. In *Proceedings of the 13th international conference on human computer interaction with mobile devices and services.* ACM, 211–220.
[27] Nassir Navab, Joerg Traub, Tobias Sielhorst, Marco Feuerstein, and Christoph Bichlmeier. 2007. Action- and Workflow-Driven Augmented Reality for Computer-Aided Medical Procedures. *IEEE Computer Graphics and Applications* 27, 5 (Sep. 2007), 10–14. https://doi.org/10.1109/MCG.2007.117
[28] John W Philbeck and Jack M Loomis. 1997. Comparison of two indicators of perceived egocentric distance under full-cue and reduced-cue conditions. *Journal of Experimental Psychology: Human Perception and Performance* 23, 1 (1997), 72.
[29] John W Philbeck, Jack M Loomis, and Andrew C Beall. 1997. Visually perceived location is an invariant in the control of action. *Perception & Psychophysics* 59, 4 (1997), 601–612.
[30] Dirk Reiners, Didier Stricker, Gudrun Klinker, and Stefan Müller. 1999. Augmented reality for construction tasks: Doorlock assembly. In *Proceedings of the international workshop on Augmented reality: placing artificial objects in real scenes: placing artificial objects in real scenes.* AK Peters, Ltd., 31–46.
[31] Cindy M Robertson, Blair MacIntyre, and Bruce N Walker. 2008. An evaluation of graphical context when the graphics are outside of the task area. In *2008 7th IEEE/ACM International Symposium on Mixed and Augmented Reality.* IEEE, 73–76.
[32] Gerhard Schall, Erick Mendez, Ernst Kruijff, Eduardo Veas, Sebastian Junghanns, Bernhard Reitinger, and Dieter Schmalstieg. 2009. Handheld augmented reality for underground infrastructure visualization. *Personal and ubiquitous computing* 13, 4 (2009), 281–291.
[33] William Steptoe, Simon Julier, and Anthony Steed. 2014. Presence and discernability in conventional and non-photorealistic immersive augmented reality. In *2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 213–218. https://doi.org/10.1109/ISMAR.2014.6948430
[34] Takahiro Tsuda, Haruyoshi Yamamoto, Yoshinari Kameda, and Yuichi Ohta. 2005. Visualization Methods for Outdoor See-through Vision. In *Proceedings of the 2005 International Conference on Augmented Tele-Existence* (Christchurch, New Zealand) *(ICAT '05)*. Association for Computing Machinery, New York, NY, USA, 62–69. https://doi.org/10.1145/1152399.1152412
[35] Stefanie Zollmann, Raphael Grasset, Gerhard Reitmayr, and Tobias Langlotz. 2014. Image-Based X-Ray Visualization Techniques for Spatial Understanding in Outdoor Augmented Reality. In *Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures: The Future of Design* (Sydney, New South Wales, Australia) *(OzCHI '14)*. Association for Computing Machinery, New York, NY, USA, 194–203. https://doi.org/10.1145/2686612.2686642
[36] Stefanie Zollmann, Gerhard Schall, Sebastian Junghanns, and Gerhard Reitmayr. 2012. Comprehensible and Interactive Visualizations of GIS Data in Augmented Reality. In *Advances in Visual Computing*, George Bebis, Richard Boyle, Bahram Parvin, Darko Koracin, Charless Fowlkes, Sen Wang, Min-Hyung Choi, Stephan Mantler, Jürgen Schulze, Daniel Acevedo, Klaus Mueller, and Michael Papka (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 675–685.