Int. J. Human-Computer Studies 72 (2014) 704–716

Contents lists available at ScienceDirect

Int. J. Human-Computer Studies

journal homepage: www.elsevier.com/locate/ijhcs

# Lessons learned: Evaluating visualizations for occluded objects in handheld augmented reality[☆]

Arindam Dey[a,b,*], Christian Sandor[a,c]

[a] School of Information Technology & Mathematical Sciences, University of South Australia, Mawson Lakes, 5095, Australia
[b] School of Public Health, Tropical Medicine & Rehabilitation Sciences, James Cook University, Smithfield, 4878, Australia
[c] Interactive Media Design Lab, Nara Institute of Science and Technology, Japan

ARTICLE INFO

*Article history:*
Received 11 September 2013
Received in revised form
20 March 2014
Accepted 9 April 2014
Communicated by D.A. Bowman
Available online 28 April 2014

*Keywords:*
Augmented reality
Handheld devices
Experiments
Visualizations for occluded objects
Outdoor environments

ABSTRACT

Handheld devices like smartphones and tablets have emerged as one of the most promising platforms for Augmented Reality (AR). The increased usage of these portable handheld devices has enabled handheld AR applications to reach the end-users; hence, it is timely and important to seriously consider the user experience of such applications. AR visualizations for occluded objects enable an observer to look through objects. AR visualizations have been predominantly evaluated using Head-Worn Displays (HWDs), handheld devices have rarely been used. However, unless we gain a better understanding of the perceptual and cognitive effects of handheld AR systems, effective interfaces for handheld devices cannot be designed. Similarly, human perception of AR systems in outdoor environments, which provide a higher degree of variation than indoor environments, has only been insufficiently explored.

In this paper, we present insights acquired from five experiments we performed using handheld devices in outdoor locations. We provide design recommendations for handheld AR systems equipped with visualizations for occluded objects. Our key conclusions are the following: (1) Use of visualizations for occluded objects improves the depth perception of occluded objects akin to non-occluded objects. (2) To support different scenarios, handheld AR systems should provide multiple visualizations for occluded objects to complement each other. (3) Visual clutter in AR visualizations reduces the visibility of occluded objects and deteriorates depth judgment; depth judgment can be improved by providing clear visibility of the occluded objects. (4) Similar to virtual reality interfaces, both egocentric and exocentric distances are underestimated in handheld AR. (5) Depth perception will improve if handheld AR systems can dynamically adapt their geometric field of view (GFOV) to match the display field of view (DFOV). (6) Large handheld displays are hard to carry and use; however, they enable users to better grasp the depth of multiple graphical objects that are presented simultaneously.

## 1. Introduction

AR superimposes computer generated virtual objects on top of the real world with 3D registration and provides the ability to interact in real time (Azuma, 1997). AR, being a user interface technology, must be experimented with human users to evaluate their experience with the system and to identify usability issues. Historically, researchers in the AR domain have not conducted an adequate number of user studies as reported by Swan and Gabbard (2005) and Dünser et al. (2008). This scarcity is even more prominent in handheld AR. A typical AR system consists of multiple components including Display, Tracking, Visualizations, and Interaction (van Krevelen and Poelman, 2010). It is important to understand individual effects of these components on the overall perception of AR interfaces.

Displays present the AR interface to the observer. Current advances in computational capacity of handheld devices have significantly increased the interest of handheld AR among researchers and developers worldwide, promising a future adoption on the consumer level (Kruijff et al., 2010). The release of multiple AR browsers and development APIs by numerous companies is evidence of this widespread interest. AR applications can now be used for mobile information browsing in outdoor locations. Empirical studies in the AR domain have been predominantly conducted using HWDs and in indoor locations. There are differences in usability between handheld and head-worn displays, and hence, insights acquired from experiments using HWDs cannot be directly applied to handheld displays. First, HWDs

---

[☆] This paper has been recommended for acceptance by D.A. Bowman.

[*] Corresponding author at: School of Public Health, Tropical Medicine & Rehabilitation Sciences, James Cook University, Smithfield, 4878, Australia. Tel.: + 61 425808230.

*E-mail addresses:* aridey@gmail.com (A. Dey), christian@sandor.com (C. Sandor).

http://dx.doi.org/10.1016/j.ijhcs.2014.04.001
1071-5819/© 2014 Elsevier Ltd. All rights reserved.

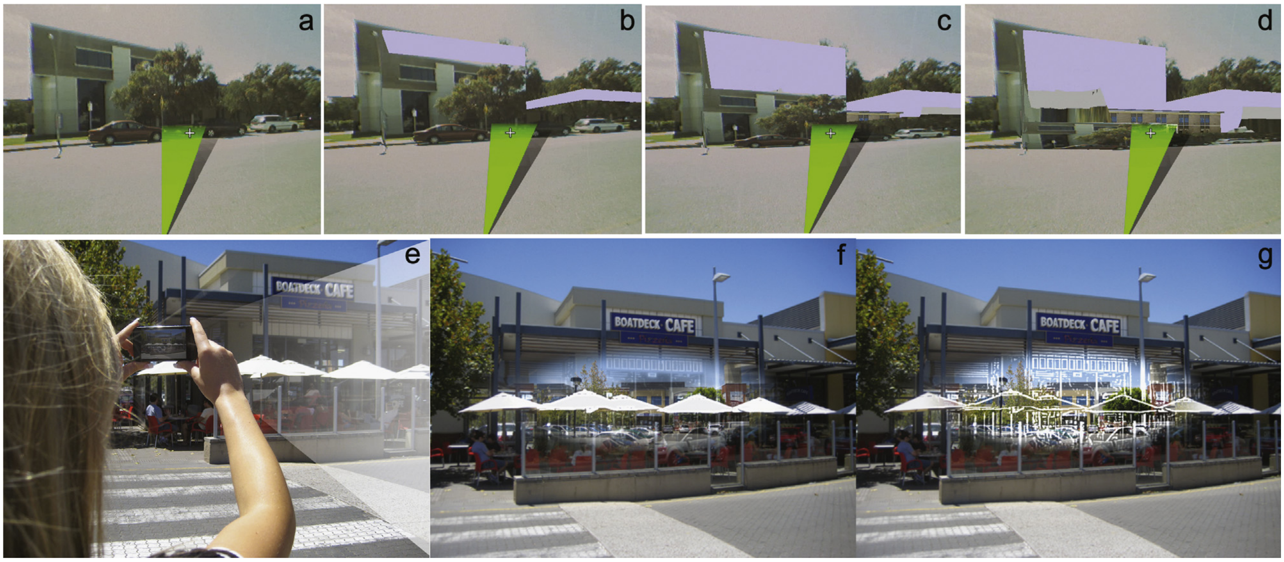

**Fig. 1.** Visualizations for occluded objects: Melt Vision virtually melts the occluding object and reveals the occluded object through an animation (a–d). When the user searches for occluded objects (e), these are revealed while foreground features are preserved: saliency-based X-ray preserves several salient features (f), while edge-overlay X-ray only preserves edges (g).

provide a higher level of immersion into the AR world than handheld displays. Second, when handheld displays are used outside, environmental effects, such as glare on the screen and moving objects in the visual field, make it even harder to see and perceive the objects present in the interface. Without proper understanding of the perceptual issues of handheld AR systems, effective user interfaces cannot be designed.

Another important component of AR systems is visualization. Visualizations portray the spatial relationship between virtual and real contents, and are another important component of any AR system. During information browsing, there are situations when points of interests (POIs) are hidden or occluded by opaque real-world objects. AR visualizations for occluded objects, commonly known as X-ray visualizations or see-through visualizations, aid in these types of situations by enabling observers to look through opaque objects which otherwise would be impossible using bare eyes (Swan et al., 2007). The metaphor of X-ray visualization was first applied to the work by Stix (1992). See-through visualizations can be categorized in two streams based on the rendering of the occluded scene: (a) Symbolic and (b) Photorealistic. On the one hand, symbolic X-ray visualizations show icons or symbols to represent the occluded objects, for example see Livingston et al. (2011). On the other hand, photorealistic AR X-ray visualizations show a realistic representation of the occluded scene through the occluding object as presented by Avery et al. (2009), Sandor et al. (2009, 2010), and Bane and Hollerer (2004). Both categories of X-ray visualizations have their own areas of application. We have investigated three different photorealistic visualizations for occluded objects: Edge-overlay X-ray, Saliency-based X-ray, and Melt (see Fig. 1).

Numerous experiments have investigated the perceptual properties of X-ray visualizations using HWDs and in indoor locations. Most of the studies used egocentric depth perception tasks to measure the general perception of the environment.

Until recently, the use of AR application was mostly restricted to indoor locations. Although, there were a few AR systems built for outdoor usage, such as Piekarski and Thomas (2001) and MARS (Feiner et al., 1997), their bulky form-factors compromised users' mobility to some extent. However, with recent advances in handheld display and tracking technologies, it is now possible to use AR in outdoor locations without compromising the mobility of the users. Notably, there are some fundamental differences between indoor and outdoor locations. In contrast to indoor locations, outdoor locations are more dynamic, noisy, and have variable lighting and environmental conditions. Cumulatively, these factors make outdoor locations uncontrollable and deviate users' attention. However, there is a clear lack of understanding about human perception in outdoor AR environments, even though AR systems are now usable in outdoor locations. Better understanding of the human perception in this environment will help researchers and developers to create effective AR interfaces.

### 1.1. Overview of experiments

To address these gaps in the empirically validated knowledge in the AR domain, over the last four years we have performed five experiments in outdoor locations using handheld displays. On a higher level we investigated the following research question: *How do users' perception and cognition get affected by the different AR visualizations for occluded objects, and their interaction with the other components of handheld AR systems in outdoor locations?*

To investigate this high level question, we have decomposed the question into three specific research questions. Our contributions are the answers to the questions mentioned in Table 1.

Below, we provide an overview of the experiments. However, we will not discuss the experiments in detail as those details are available in earlier publications. Table 1 shows the key insights acquired from these studies.[1]

Experiment 1 investigated Edge-overlay X-ray versus Melt visualizations (Dey et al., 2010). Melt was advantageous for egocentric depth judgment (distance to a target object) when a synthetic depth cue was present and egocentric depth was underestimated for both visualizations. This finding contradicts (Livingston et al., 2009) where a depth overestimation was reported.

Experiment 2 investigated the effect of egocentric distance to the target object and rendering techniques of X-ray visualizations on depth judgment. Depth was underestimated in medium- to far-field distances and different rendering techniques did not influence depth judgment.

---

[1] Please see Appendix A for detailed statistical analysis and data values.

**Table 1**
Summary of key insights acquired from respective experiments. Egocentric distance refers to the distance to an object from the observer and exocentric distance refers to the distance between two objects in the view. DFOV refers to the physical subtended angle from the eyes to the edges of the screen. GFOV is a property of any graphical application, and is defined as the subtended angle between the left, right, top, and bottom ends of the view frustum originating from a virtual camera used in that application.

| Research questions | Key insights | Experiments | | | | |
|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 |
| How do different visualizations for occluded objects affect human perception, and when one should be used over the others? | Visualization for occluded objects improves the depth perception of occluded objects to be the same as if they are not occluded | | X | X | X | |
| | Less occlusion causes better depth judgment | X | | | | |
| | Visualizations complement each other in different usage scenarios | | | | | X |
| How accurately do users perceive the spatial relationships among virtual objects in handheld AR? | Distances underestimated in outdoor locations | | | | | |
| | • Egocentric | X | X | X | | |
| | • Exocentric | | | | X | |
| How do the physical properties of handheld displays affect human perception? | Interplay between DFOV and GFOV affects depth judgment | | | X | X | |
| | Larger displays provide better ordinal depth judgment | | | | X | |

Experiment 3 and Experiment 4 investigated the effects of size and resolution of handheld displays on egocentric, exocentric (distance between two target objects), and ordinal (relative ordering of two target objects) depth judgment (Dey et al., 2012). Smaller displays caused less depth underestimation than larger displays, and larger displays were advantageous for ordinal perception.

Experiment 5 investigated Edge-overlay versus Saliency-based X-ray visualizations, and found that both of these visualizations complement each other in different use cases (Sandor et al., 2010).

### 1.2. *Contributions*

The main contribution of this paper is the summarization of key insights acquired from five different empirical studies on handheld AR, and the identification of design constraints in light of these insights that will improve the perception of handheld AR interfaces in outdoor locations. This paper addresses the gap in usability research in the handheld AR domain at a time when handheld devices like smartphones have become the primary mode of information browsing. We believe that these insights will enable developers and researchers to create effective handheld AR interfaces with improved user experience.

The rest of the paper is organized as follows. Section 2 reviews previous perceptual studies performed in the AR domain. Section 3 briefly describes the design principles of our photorealistic visualizations for occluded objects. Sections 4–8 describe the design and results of our experiments. We coalesce the experimental results in Section 9. Finally, we conclude by pointing towards future research in Section 10.

## 2. Related work

The success of AR applications, being a user interface technology, depends heavily on human perception of the environment. Often, to investigate the perception in a given environment researchers adopt depth judgment tasks. In this section we review user studies in the AR domain that investigated depth perception.

Perception of distance helps us to create a three-dimensional impression of the world by combining two two-dimensional, flat retinal images captured from slightly different viewpoints (Swan et al., 2006). Understanding depth perception in AR is necessary to portray the correct relationship between the real-world objects and the virtual objects to the observer as intended by the developer. A wide range of research has been carried out on this domain.

Perception is an invisible cognitive state and we measure the perception of depth through quantifiable depth judgments and perception is inferred from these judgments (Swan et al., 2006). To measure depth judgment, perceptual scientists have employed different types of tasks including verbal estimation, and closed- and open-loop action based tasks. A review of such tasks is presented by Loomis and Knapp (2003).

Two different types of depth judgment occur in our environment: *egocentric* and *exocentric* (Milgram et al., 1994). Egocentric depth perception refers to the distance to an object perceived from the observers viewpoint; exocentric depth perception refers to the distance between two objects in the view (Swan et al., 2006). Both egocentric and exocentric depths are underestimated in virtual environments (Pollock et al., 2012); however, in the real world, depth estimation is somewhat accurate (Loomis and Knapp, 2003). To solve this problem, various synthetic depth cues have been proposed (Livingston et al., 2003; Tsuda et al., 2005; Wither and Hollerer, 2005).

### 2.1. *Egocentric depth perception*

AR has been investigated widely, specifically using HWDs. Most depth judgment studies in AR, like Virtual Reality, have consistently reported depth underestimation of objects presented on a ground plane. However, the reason of this underestimation is not clearly understood.

Numerous studies have evaluated egocentric depth judgment in near-field distances. The effect of near-field distances along with an occluded surface, convergence, accommodation, age, and stereo displays was studied through a perceptual matching task in a series of studies (Ellis and Menges, 1998). Later on, the effect of motion parallax and system latency was explored by (McCandless et al., 2000). Recently, a study investigated reaching and matching tasks in near-field distances (Singh et al., 2012).

Egocentric depth judgment in medium- and far-field AR was evaluated using a perceptual matching protocol in Swan et al. (2006). This experiment interestingly reported a shift in bias from underestimation to overestimation at 23 m in an indoor environment, whereas depth underestimation is a common phenomenon in virtual environments. Later on, another experiment without using any X-ray vision reported depth overestimation of medium-field distances in an outdoor environment (Livingston et al., 2009).

Multiple visualizations for occluded objects were evaluated by Livingston et al. (2003, 2011). Recently, the effect of peripheral vision was evaluated by Jones et al. (2011). All the above studies used an optical see-through HWD.

In the last few years, however, handheld devices like mobile phones have become a promising platform for AR applications as their computational power has increased (Kruijff et al., 2010). While perceptual issues in AR are investigated predominantly using HWDs, handheld displays remained under-explored.

Some recent studies investigated perceptual issues with X-ray visualization using handheld displays. A set of depth cues for X-ray visualization was evaluated by Tsuda et al. (2005). We have investigated X-ray visualization with a target selection task in Sandor et al. (2010), with a depth judgment task in Dey et al., 2010, and with a real-world navigation task in Dey et al. (2011). Results in Dey et al. (2010), contradicting (Livingston et al., 2009), reported a consistent depth underestimation in an outdoor environment, however, of far-field distances. In this paper, we are further investigating if different ranges of distances affect depth judgment.

### 2.2. Exocentric depth perception

Exocentric depth refers to the distance between two objects in the visual field. Numerous experiments have investigated the exocentric depth judgment in real-world scenarios. It was reported that the change in the viewing angle can change the exocentric depth judgment of exactly the same stimuli (Levin and Haber, 1993). A mathematical model of the visual space was created by Foley et al. (2004) based on exocentric depth judgment. Loomis et al. (2002) proved a dissociation between perceived target location and perceived exocentric distance and shape. The effect of familiar and unfamiliar sizes of objects on depth judgment was investigated by Predebon (1991). However, in the AR domain exocentric depth judgment is not investigated at all.

In summary, it is evident that depth perception is a complex process that depends on multiple depth cues available in the environment. These depth cues are not always preserved in AR interfaces due to the displays and visualization methods used. Particularly, different visualizations for occluded objects render the scene differently and preserve different levels of information about the environment. These visualization methods influence depth judgment. On the one hand, the effects of symbolic visualizations for occluded objects on depth judgment are widely studied, however, the effects of photorealistic visualizations on depth judgment are not properly understood. On the other hand, most of the user studies were executed in indoor locations and using HWDs. AR depth judgment in outdoor locations using handheld displays has not been experimented. Similarly, all the earlier depth perception studies in the AR domain focused on egocentric depth judgment, either metric or ordinal. However, exocentric depth judgment was not investigated in AR.

## 3. Visualizations for occluded objects

In our experiments we have used three different visualizations for occluded objects: Edge-overlay X-ray, Saliency-based X-ray, and Melt Vision. These three visualizations were designed to solve a common purpose, but in different usage scenarios.

### 3.1. Edge-overlay X-ray

The purpose of Edge-overlay X-ray was to visually portray the relationship between the occluding and the occluded objects by preserving the edges of the foreground object (see Fig. 1g). By preserving the foreground edges and overlaying them on top of the occluded objects enable the impression of occlusion which would have been missing in a naïve overlay of occluded objects on top of the foreground.

### 3.2. Saliency-based X-ray

In addition to the edge information, preserving the other salient features—color hue, brightness, and moving objects—of the foreground can potentially attract human attention (see Fig. 1f). The primary aim of this visualization was to preserve the important information of the foreground, such as street signs, that would not be available in Edge-overlay X-ray. Similarly, preserving moving objects can increase the safety of the user while using this visualization.[2]

### 3.3. Melt Vision

Melt Vision completely removes the core area of the occluding object through an animation and provides a clear view of the occluded object (see Fig. 1a–d). While the core area is removed, the peripheral area of the foreground remains intact to provide a sense of occlusion.

## 4. Experiment 1: Edge-overlay X-ray versus Melt Vision

The purpose of this experiment was to study the effects of Edge-overlay X-ray and Melt Vision on outdoor depth judgment. We purposefully used 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).

### 4.1. Design and procedure

Twenty voluntary participants with ages ranging from 18 to 31 years (M=25, SD=3.8) participated in this experiment. Participants stood in front of a 7 in display (640 × 480 pixels) facing a building 29 m away. This building was used as the occluding surface. We rendered a 3 m × 3 m × 3 m green cube as the occluded object at five different distances: 69.7 m, 82.5 m, 95.1 m, 104.6 m, and 117 m. To visualize the target cubes, we rendered either Edge-overlay X-ray or Melt Vision as a within-subject variable (see Fig. 2). We also presented a synthetic depth cue as a between-subject variable. Participants had to verbally report the egocentric distance of the cubes. The display was mounted on a tripod and participants were not allowed to move the display during the depth judgment task.

### 4.2. Analysis

We found that Melt Vision was significantly more accurate in depth estimation than Edge-overlay X-ray vision. The graphical depth cue significantly improved the egocentric perception, but, it increased the response time as well (see Fig. 3). However, the most important insight of this experiment was a consistent underestimation of the egocentric distance in an outdoor AR environment.

In a previous work by Livingston et al. (2009), differing from our results, it was reported that egocentric distance was overestimated in an outdoor AR environment. However, their experimental setup differs from ours in the following aspects: head-worn optical

---

[2] In our prototype of the Edge-overlay X-ray, we first did the extremely compact edge overlay, which only highlights edges. Next, we explored the larger concept of saliency highlights (edges are only one example). Contrary to Edge-overlay X-ray, in Saliency-based X-ray, a moving object is considered to be salient and when it comes to a stop it is not considered to be salient unless it has other salient features such as brightness.

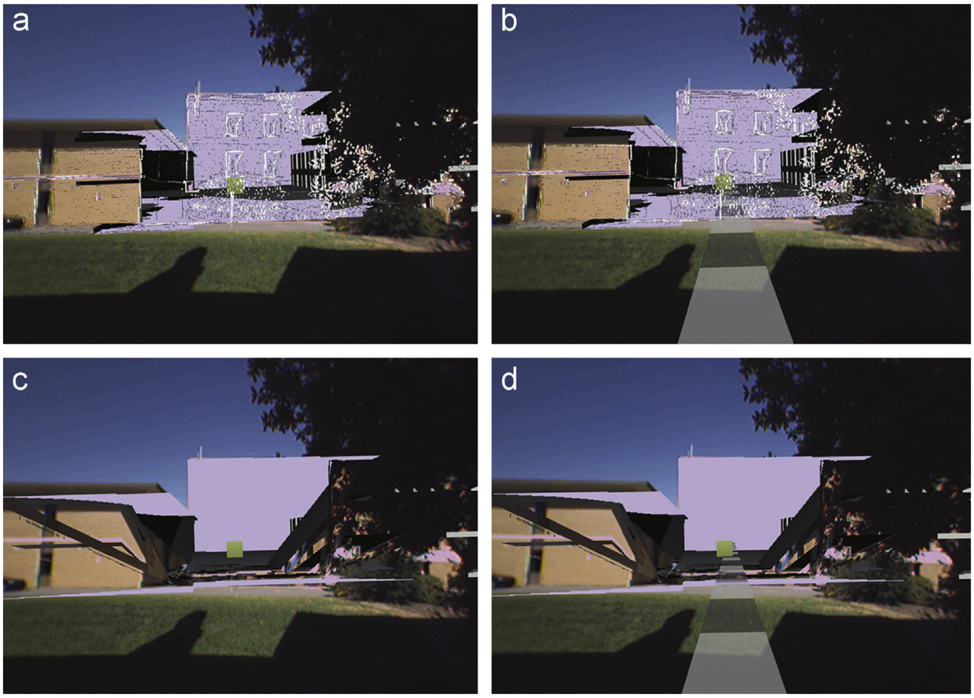

**Fig. 2.** Experiment 1 conditions: the top row shows our Edge-overlay X-ray visualization without (a) and with (b) a depth cue. The bottom row shows our Melt Vision without (c) and with (d) a depth cue. The depth cue originated from the participant's location and each black and white section was 10 m long. Participants had to estimate the distance to the green cube. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

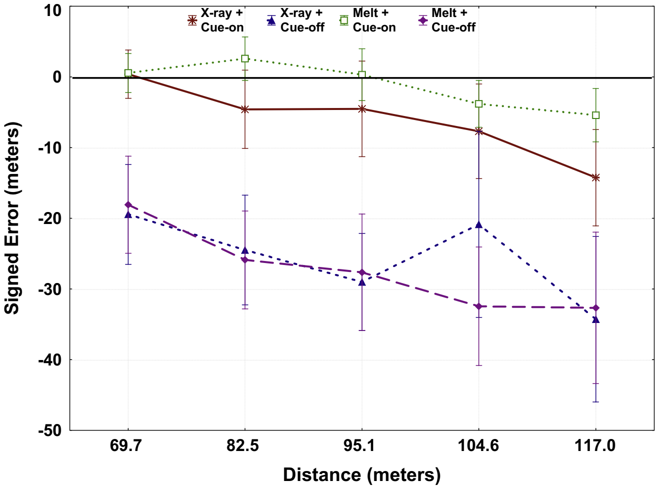

**Fig. 3.** Experiment 1 results: we have noticed a consistent underestimation of distance. Melt Vision was significantly more accurate than X-ray visualization. The synthetic depth cue aided distance judgment but increased response time. The thick line at 0 shows veridical perception. Whiskers show ±95% confidence interval.

see-through display, shorter distance field, no occluded objects, and a different experimental protocol. Due to these differences, our results cannot be directly compared. To further investigate and validate the reasons of this contradiction, we investigated the effect of distance-fields on egocentric perception in Experiment 2.

## 5. Experiment 2: effect of distance-field and X-ray visualizations

This user study was designed to investigate if the use of different distance-fields and the use of X-ray visualization had any effect on depth judgment that caused the contradiction noticed in Experiment 1. In this experiment, we investigated the effect of X-ray visualization and distance-fields on egocentric depth judgment using an iPhone. Additionally, we investigated the effect of tracking methods on depth judgment.

### 5.1. *Design and procedure*

Twenty-seven students and staff, ages ranging from 21 to 63 years, from the University of South Australia were recruited for the *mixed-factorial* experiment and were equally distributed into three matched groups. We used three spatial arrangements as a between subjects variable (see Fig. 4). On–On condition had visualized an occluded 3 m × 3 m × 3 m pyramid using Edge-overlay X-ray. On–Off condition visualized the occluded pyramid without any X-ray visualization and the pyramids appeared to be floating on the occluding object. Off–Off condition was similar to conditions used by Livingston et al. (2009), where virtual objects were not occluded. The experimental distances ranged between 19.3 m to 117 m, covering the distances used in Livingston et al. (2009) and Dey et al. (2010). We also varied Vision-based and Sensor-based tracking methods as a within-subjects variable. As we only used an orientation tracker, motion parallax was not available as a depth cue. Participants held an iPhone 4S (3.5 in diagonal screen), and reported egocentric distance to the tip of the pyramid. During the experiment we noticed that participants did not move the display a lot and held it within a comfortable distance of 35 cm to 45 cm from their eyes. This behavior was also visible in Experiments 3 and 4 presented in the following sections of this paper. We attribute this behavior to the unavailability of the motion parallax depth cue.

### 5.2. *Analysis*

Results of this experiment validated that depth is indeed underestimated in all distances in outdoor AR environments and using different ranges of experimental distances was not a reason for the contradiction with Livingston et al. (2009) found in Experiment 1 (see Fig. 5). Distance had a significant effect on depth

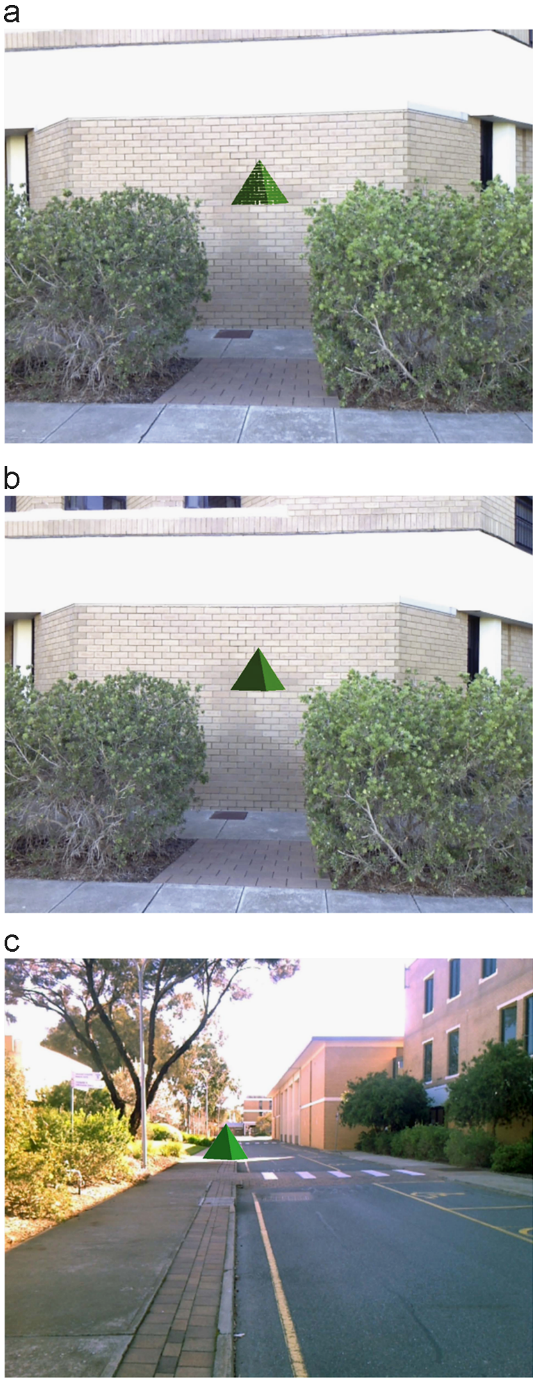

**Fig. 4.** Three spatial arrangements used in Experiment 2: in two of the conditions, the target pyramid was occluded; in one of them we used an Edge-overlay X-ray visualization to display the pyramid (a); in the other one, we did not use any X-ray visualization (b). The other control condition was similar to Livingston et al. (2009), where the target pyramid was presented without any occlusion (c).

judgment, with the increase of distance accuracy of perception decreased. We did not notice any significant effect of X-ray rendering and tracking quality on depth judgment. However, participants subjectively preferred vision-based tracking.

We have also noticed that the depth compression in far-field distances is noticeably more than our previous experiment. The experimental environments were similar in both experiments, however, in the earlier experiment we used a larger handheld display with a lower resolution. To investigate the reason of this difference more deeply, we varied size and resolution of handheld displays systematically in separate conditions, and found how the physical attributes of handheld displays affect depth judgment in outdoor AR environments.

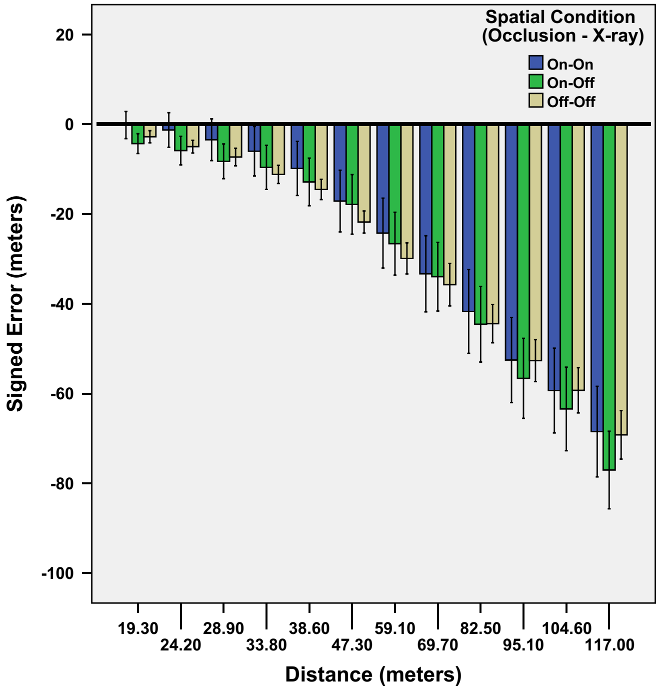

**Fig. 5.** Experiment 2 results: signed error shows a consistent underestimation of distances in all spatial conditions. The thick black line at 0 represents veridical perception and whiskers represent ±95% confidence interval.

## 6. Experiment 3: effect of display size and resolution on egocentric depth perception

This is an early experiment in the AR domain where the effects of size and resolution of handheld displays on depth perception are investigated. The experimental environment was the same as for the previous experiment. We used two different display devices. First, an iPhone 4S with 3.5 in screen size and 960 × 640 (326 ppi) resolution. Second, an iPad 3 with 9.7 in screen and 2048 × 1536 (264 ppi) resolution.

### 6.1. *Design and procedure*

This experiment was intentionally designed to be a within-subject experiment as we intended to evaluate all the conditions using the same participants, eliminating any errors induced by separate participant groups. Twelve participants (aged between 22 and 41 years) were recruited from the university staff and students; among them six participants participated in the Experiment 2 at least two weeks prior. We investigated three different display configurations. Small-Low condition used iPhone 4S with its native resolution, Big-Low condition used an iPad 3 with a resolution of iPhone 4S, and Big-High condition used iPad 3's native resolution.

The experimental task was the same as for Experiment 2, however, in this experiment we did not use On–Off condition as we did not notice any significant difference between spatial arrangements. We only used vision-based tracking in this experiment.

### 6.2. *Analysis*

Similar to the two earlier experiments, in this experiment, we noticed a consistent underestimation of depth in all conditions.

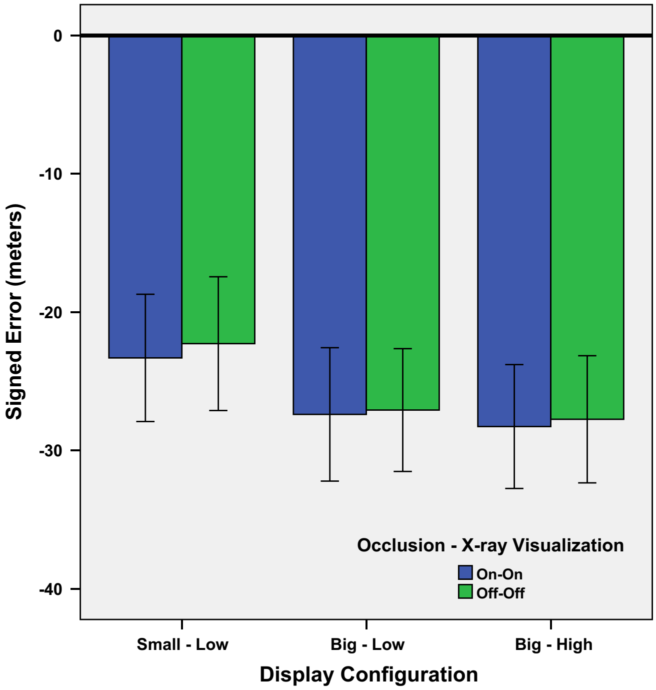

**Fig. 6.** Experiment 3 results: underestimation was less in the Off–Off spatial arrangement; this effect was significantly more in Small-Low display configuration. Whiskers represent ±95% confidence interval.

However, the underestimation was significantly less using a small screen. We did not notice any effect of resolution on depth judgment (see Fig. 6). Expectedly, errors increased with distance. Participants, however, subjectively preferred the larger screen size significantly, compared to the smaller one; we did not notice any significant effect objectively, though.

This conflict in objective and subjective measures prompted us to conduct a further experiment to investigate if the subjective preference is substantial. Accordingly, in the next experiment, we investigated the perception of exocentric and ordinal depths using the same displays.

## 7. Experiment 4: effect of display size on exocentric and ordinal depth perception

The primary goal of this experiment is to quantitatively investigate the validity of the subjective responses where participants reported significant difficulty in depth judgment using a smaller display. Accordingly, in this experiment we used exocentric and ordinal depth judgment tasks and showed two pyramids together on the screen. In ordinal perception, the identification of the difference in the heights of two pyramids was critical to guess which one of the two pyramids is closer. Users require exocentric and ordinal depth judgments in real-world AR applications where they see multiple POIs together, and they have to perceive the spatial relationship between them. However, this is the first experiment in the AR domain to experiment exocentric depth judgment.

### 7.1. Design and procedure

The same 12 participants from the last experiment were recruited for this *within-subject* experiment. However, there was at least one week gap between the two experiments for every participant. As we objectively investigated the validity of the subjective preference noticed in the last experiment; the use of the same participant group in both experiments was crucial.

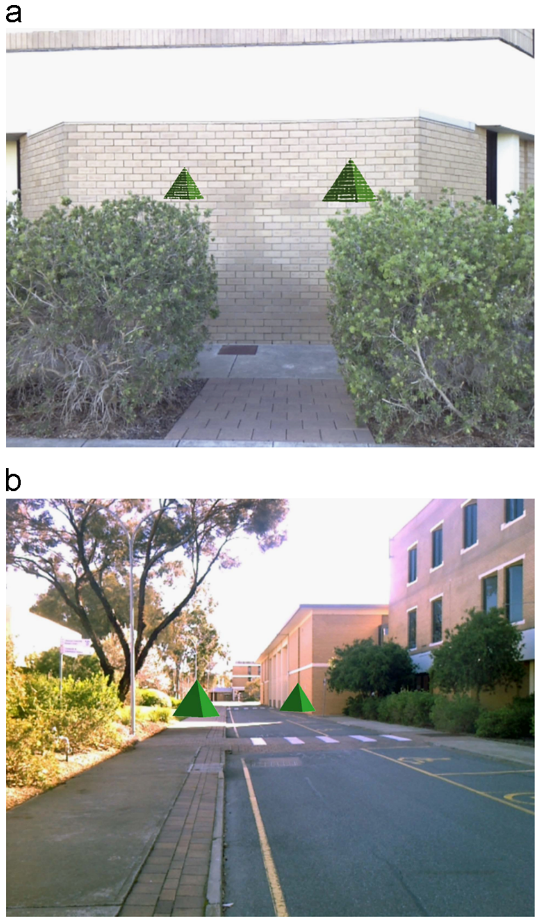

**Fig. 7.** Experiment 4 conditions: the On–On arrangement showed two occluded pyramids by rendering an Edge-overlay X-ray (a). The Off–Off arrangement presented the pyramids without any occlusion, and hence, no X-ray rendering was required (b).

We used On–On and Off–Off spatial conditions, however, in this experiment we placed two virtual pyramids of identical size on the scene (see Fig. 7). Participants first had to judge which of the two pyramids is closer by answering left, right, or equal (ordinal perception). Then they had to guess the distance between the tips of the pyramids (exocentric perception). In this experiment we excluded Big-Low display configuration as we did not find any significant effect of resolution on depth judgment.

### 7.2. Analysis

We confirmed that, like egocentric depth, exocentric depth is also underestimated in outdoor AR environments. Most importantly, we have found that larger displays provide better spatial perception than smaller displays where multiple objects are present on the scene (see Fig. 8). These findings alleviate the contradiction noticed in Experiment 3 between qualitative and quantitative responses as indeed larger displays are advantageous for overall perception of the AR interface.

We have noticed that AR Edge-overlay X-ray visualization does not affect depth judgment and provides similar performance like situations where augmented target objects are visible without any occlusion and no X-ray visualization was used. This finding encouraged us to investigate whether preserving more information while rendering X-ray visualizations will affect perception of the AR interfaces. Accordingly, in Experiment 5, we have compared Edge-overlay X-ray with Saliency-based X-ray through a subjective online survey.

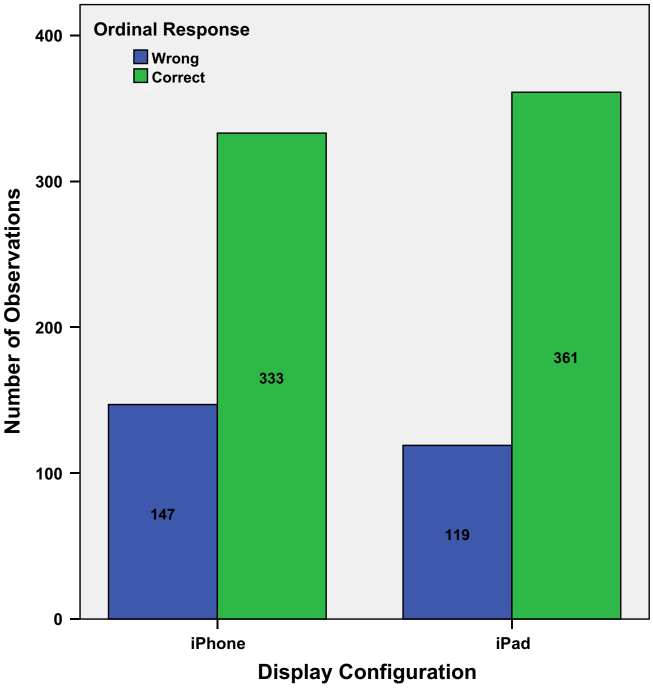

**Fig. 8.** Experiment 4 results: larger display performed significantly better than smaller display in terms of ordinal perception.

## 8. Experiment 5: edge-overlay versus saliency-based X-ray visualizations

Besides edge information of the occluding object, our Saliency-based X-ray visualization (Sandor et al., 2010) preserved hue, luminance, and motion features of the foreground. However, too much visual information from the occluding object can cause distracting visual noise and hide essential regions of the occluded area. Hence, it was important to understand if the use of Saliency-based X-ray visualization significantly affects users' performance in comparison with the previously investigated Edge-overlay X-ray visualization. This experiment was motivated to investigate the comparative performance of these two visualizations through an online survey.

### 8.1. Design and procedure

Twenty-five voluntary respondents responded to our online survey with ages ranging from 18 to 62 years. This within-subject survey was based on three independent variables: Brightness, Edge, and X-ray. We selected three levels for Brightness and Edge: High, Medium, and Low. An expert panel, consisting of eight members of our research group, carefully selected nine combinations of occluding and occluded regions from a set of 25 random combinations to correctly represent the different levels of brightnesses and edges. We then executed our X-ray visualizations to create 18 different images.

We instructed respondents to do a *see-perceive-score* task. The images were presented to respondents one at a time. After observing the image carefully, respondents had to score on a scale of 1 (worst) to 10 (best) to report how well the image conveyed information for occluding (foreground) and occluded (background) regions.

### 8.2. Analysis

Results showed that high levels of edges cause problems in the Edge-overlay X-ray visualization, whereas high levels of brightness cause problems in the Saliency-based X-ray visualization (see Fig. 9).

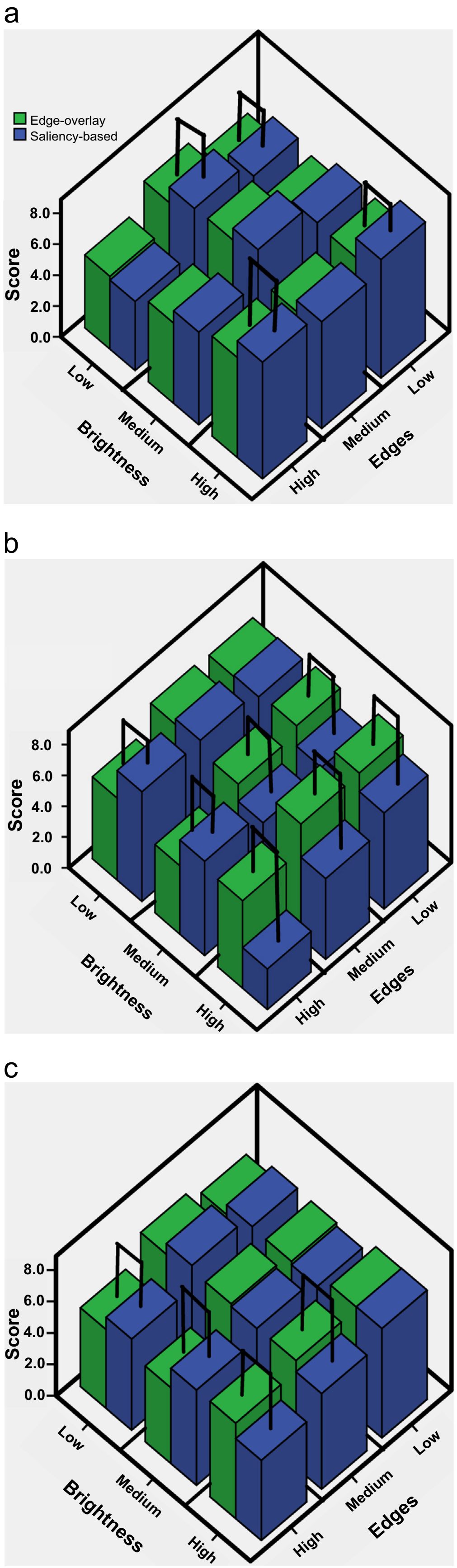

**Fig. 9.** Experiment 5 results: the bars represent mean scores provided by respondents. Connected bars represent significant differences between means. (a) Foreground. (b) Background. (c) Total.

This result is consistent with the way we have designed the Saliency-based and Edge-overlay X-ray algorithms.

The problem of highly salient foregrounds for Saliency-based X-ray is very similar to the problem of Edge-overlay X-ray when too many edges are present. On an abstract level, they both point towards the need for an adaptive classification of salient features.

The primary intention of using visualizations for occluded objects on handheld displays is to aid users, in real-world use cases perceive the occluded POIs in a proper spatial egocentric and exocentric relationship. However, there are other methods of displaying hidden POIs on handheld displays such as mobile maps. It is important to understand how AR X-ray visualizations compare to other commercial methods available in the market. While AR can present information in an egocentric perspective, mobile maps present information in an exocentric perspective.

## 9. Discussion

As mentioned in Section 1, the contribution of this paper is the answers to the three research questions we aimed at investigating. In this section, we discuss the findings and lessons learned in the process.

### 9.1. Effects of visualizations for occluded objects

Our results suggest that different photorealistic visualizations for occluded objects affect human perception differently.

*Use of visualizations for occluded objects improves depth perception akin to non-occluded objects:* In Experiments 2–4 our results showed that visualizations for occluded objects, despite preserving the information about the occluding object, do not negatively affect depth judgment. Occluded objects are perceived with the same accuracy as if they were not occluded. However, when occluded objects are naïvely rendered without any cues, as commonly done in current AR browsers, depth judgment deteriorates.

Depth judgment is subject to the subtle properties of the surrounding environment and not always accurate even in the real world (Lappin et al., 2006). A meta-analysis by Waller and Richardson (2008) showed that real-world depth perception is more accurate than for virtual worlds; however, a recent study by Jones et al. (2011) showed that when users can move while seeing both virtual and real objects (a similar situation as in handheld AR applications) their AR and VR depth judgments rapidly improve and become similar to that of the real world. Hence, for handheld AR interfaces intended to be used for environment exploration and navigation, such as AR browsers, using visualizations for occluded objects, the distance of the occluded objects can be estimated with similar accuracy of unoccluded objects. This estimation, even if initially less accurate than for the real world, can be made more accurately and rapidly and help in locomotion towards the target object as users make an internal representation of the visually perceived space and can accurately update it with visually directed action (Loomis et al., 1996). In fact, in a previous navigational study using X-ray visualization we noticed that X-ray visualization provides an equivalent navigational performance to that of the mobile maps (Dey et al., 2011).

Accordingly, for future AR interfaces we strongly advise others to use visualizations for occluded objects when visualizing objects that are hidden by some occluding objects. It will provide effective environmental context and aid in depth judgment.

*To aid depth judgments, visualizations of occluded objects should preserve only as much information of the occluding objects as required to portray the correct environmental context:* Visualizations presenting the occluded objects with the east amount of occlusion, such as the Melt Vision, enable better depth judgments than visualizations with more occlusion such as X-ray visualizations. Particularly, when a synthetic depth cue is used, the difference in perceptual accuracy with other visualizations increases significantly.

Research in neurophysiology has shown that humans can extract information about the target object from varying amounts of visual noise (Pratte et al., 2013) with attention; however, it is also reported that object recognition is fast and efficient in good viewing conditions (Vanrullen and Thorpe, 2001; Sugase-Miyamoto et al., 2011; DiCarlo et al., 2012). Additionally, in poor and noisy visual conditions, our visual system needs to utilize additional processing resources to effectively perceive the visual stimuli and as a result cognitive load and response time increase (Naya et al., 2001; Sheinberg and Logothetis, 2001; Ullman, 2009).

From our findings and earlier research it is clear that the level of visual noise (i.e. preserved occlusion in the visualizations) should be minimized for better perception. Either the AR system should adaptively minimize the level of noise based on the occluding object's surface, or at least there should be some mechanism provided to users to minimize the noise interactively. While doing so, developers should be careful to not eliminate environmental context or important information. For example, Melt Vision removes the occluding object completely, but this is not ideal for situations where occluding objects contain required information to perceive the environment.

*Handheld AR systems should provide a suite of complementary visualizations to be used in different conditions:* In conjunction with the above suggestion, AR designers should carefully consider various use cases of the AR system and provide adequate visualization tools to the users. We noticed that Edge-overlay and Saliency-based X-ray visualizations provide more information about occluding objects than Melt Vision while showing the occluded objects. Interestingly, these two X-ray visualizations are suitable for different use cases. Due to the underlying rendering algorithms, Saliency-based X-ray is more suitable for occluding surfaces with higher density of edges and where more information (beyond only edges) about the occluding object is needed; and Edge-overlay X-ray visualizations are suitable for occluding surfaces full of salient features, such as bright colors, that may impede the visibility of the occluded object. On the contrary, when a clear view and more accurate perception of the occluded objects is required Melt Vision should be used. Livingston et al. (2013) reviewed currently available visualizations for occluded objects and highlighted the need for further research on this topic to develop purposeful X-ray visualizations for different use cases. Ganapathy (2013) also argued for a visualization system based on user-selected values about how far she wants to see in the occluded region and also to equip her with multiple interaction methods to toggle for different needs. Overall, based on our research and other literature, we assert that for outdoor AR interfaces designers should provide a suite of visualizations for occluded objects to complement each other in different situations.

### 9.2. Effect of spatial relationships among virtual objects

In the AR domain, depth judgment studies have investigated egocentric depth judgment only. We believe that we are first to have also investigated exocentric depth judgments in Experiment 4 where participants judged the distance between two target objects.

*Exocentric distances are underestimated:* Our results indicated that, like egocentric distances, exocentric distances are also underestimated and the underestimation increases with distance. Similar to egocentric distance, smaller displays cause less underestimation of exocentric distance than larger displays.

### 9.3. Effect of handheld displays

In Experiments 3 and 4, we presented two different user studies where the effects of display size and resolution on depth judgment were investigated in outdoor environments. Our findings suggest that the size of the displays affects depth judgment. However, the resolution of the displays does not affect this judgment.

*Reasonably large displays should be used for better perception:* Our results also indicate that larger displays cause significantly faster response times and less errors than smaller displays for ordinal depth judgment where the relationship between two virtual objects needs to be perceived. An earlier study on the relationship between perceived image quality and display size also suggested that larger displays are preferred over smaller displays for images having more than 100 pixels (Barten, 1988). Earlier studies have found that the effect of auditory noise, which is available in abundance in outdoor locations, along with visual noise can interfere with the processing of visual stimuli (Acosta and Richard, 1976; Simon et al., 1981; Stoffels et al., 1985). Hence, in outdoor environments, where various visual and auditory noises distract our attention, a better image quality may provide help in perception. Accordingly, for outdoor handheld AR interfaces such as AR browsers, where multiple virtual objects are presented, we advise others that reasonably large displays should be used.

*Purposefully manipulating the ratio of Geometric Field of View (GFOV) and Display Field of View (DFOV) may result in better perception:* DFOV refers to the physical subtended angle from the eyes to the edges of the screen (Tan et al., 2006). Hence, DFOV changes with the size of the display and its distance from the eyes. GFOV is a property of any graphical application, and is defined as the subtended angle between the left, right, top, and bottom ends of the view frustum originating from a virtual camera used in that application. In AR, GFOV is typically defined to match the FOV of the physical camera used to capture the real environment (see Fig. 10). Anything within that defined frustum is rendered on a 2D plane, which is displayed to the user. Hence, with a constant GFOV, a smaller DFOV causes minification, and larger DFOV causes magnification of the environment (Steinicke et al., 2011).

In Experiments 3 and 4, all the experimental factors were the same for both displays except the DFOV. Yet, smaller displays consistently caused less underestimation than larger displays. Smaller displays with smaller DFOV caused minification of the AR environment; hence, participants perceived the target objects to be further away than they perceived them on the larger displays. This finding indicates that the interplay between GFOV of the AR application and DFOV of the display interferes with our depth perception using handheld AR systems. The fact that the interplay between GFOV and DFOV affects perception was reported earlier by researchers such as Rolland et al. (1995), Tan et al. (2006), Kuhl et al. (2009), and Steinicke et al. (2011). However, neither did they identify the effect of this interaction on handheld AR systems, nor they identified the effect of this interaction on depth perception. Receiving hints from these findings, further experimentations are required to identify the efficient ratio between DFOV and GFOV in handheld AR applications. Once identified, AR system designers and developers can purposefully manipulate the GFOV to improve visual perception of the AR content on the display.

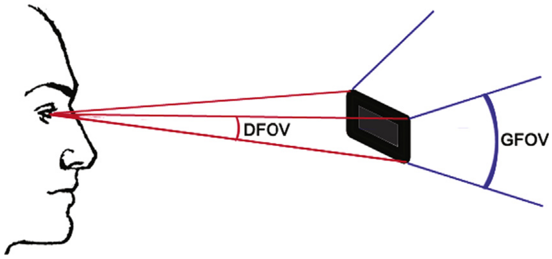

**Fig. 10.** The interplay between DFOV and GFOV affects depth judgment in handheld AR systems. By carefully manipulating the ratio of these two field of views may improve visual perception using these systems.

## 10. Conclusions and future work

The primary aim of our research was to investigate the perceptual and cognitive effects of different components of handheld AR systems with visualizations for occluded objects. This investigation was particularly important in the current state-of-the-art where handheld devices are increasingly used for AR applications. However, prior to our work researchers in this domain have predominantly chosen HWDs for usability evaluations. Hence, empirically validated knowledge of perceptual and cognitive effects of handheld AR systems was unavailable.

Overall, in this paper we have presented some key insights and design recommendations for handheld AR systems collected from a series of five different experiments. We believe that these insights will enable researchers and developers of handheld AR applications to create effective interfaces and conduct further usability research using handheld devices.

During the course of this research, a few research ideas have emerged which we have not investigated in detail. However, investigating these ideas will provide important insights for future research and development.

*Implementing an AR system with dynamically changing GFOV to initiate better perception:* The results of Experiments 3 and 4 showed that the interplay between the DFOV and GFOV causes less underestimation in the cases of smaller displays than larger displays. However, in our experiments we did not explicitly investigated this phenomenon. It will be interesting to analyze the exact effects of the interplay between DFOV and GFOV by varying them systematically in a dedicated experiment. Once we have identified this effect, we can plan to dynamically change the GFOV based on the distance of the display from user's eyes, to create an accurate depth judgment in outdoor handheld AR systems. However, this research can potentially benefit any AR system.

*Effects of various display technologies used in AR on depth judgment:* We have investigated handheld displays with different sizes and resolutions. In the near future, lightweight and technically inferior HWDs will be introduced to the mass-market. While these HWDs are likely to promote mass adoption of AR, they are expected to be of lower resolution and field of view than the commonly used HWDs in AR. More experimentation is required to understand how depth judgment using these types of displays may differ from currently used handheld displays. Similar studies to investigate differences in depth judgment using optical see-through and video see-through HWDs are required in indoor environments. This investigation will also help us understand the underlying reasons of the contradictions we noticed with Livingston et al. (2009) in Experiment 1. A deeper understanding of the comparative performance of different display technologies will help researchers and users to choose appropriate displays for different AR applications.

*Investigating the effect of motion parallax in handheld AR systems:* We noticed that participants did not move their displays a lot, even though they were allowed to do so. It will be interesting to investigate this behavior by using a position tracker for handheld AR systems, and hence, enabling motion parallax, if this behavior changes. We assume that participants will move their displays more in this situation and will benefit from motion parallax depth cue. Other studies in psychology have repeatedly reported that while motion parallax adds to depth information, its effectiveness as a depth cue is weak compared to other depth cues (Gillam et al., 2011). Motion parallax alone as a depth cue results in distance underestimation, and when coupled with binocular disparity depth estimation increases (Bradshaw et al., 1998). It was also reported that binocular disparity interacts strongly with motion parallax (Bradshaw and Rogers, 1996). Handheld AR systems are inherently monocular. Therefore, we expect that adding motion parallax as a depth cue may lead to further underestimation of depth. A study on AR using HWDs found that motion parallax do not have a strong effect on depth judgment (Jones et al., 2008). How motion parallax affects depth judgment in handheld AR will be interesting to examine. This research will inform future design of handheld or monocular AR systems.

*Investigating AR interfaces with diverse user-groups:* While investigating the literature and performing our own experiments we have realized that in the AR domain most of the evaluations are conducted using mainly male participants of between 20 and 40 years old. It will be helpful to include more female participants in our evaluations. At the same time, it will be useful to investigate the moderating effect gender on the perception of AR systems. Previous research has indicated differences in spatial behavior between males and females (Mikhailova et al., 2012), however the effect on depth judgment was not investigated. We would also like to investigate the difference in perception for different age groups of users. These investigations can help researchers to design interfaces and synthetic cues suitable to a diverse range of users.

## Acknowledgments

We would like to thank Andrew Cunningham, Graeme Jarvis, and other members of the Magic Vision Lab for their support in these experiments and comments on this manuscript. Our sincere thanks to the participants for their voluntary participation. Experiment 5 was partially funded by Nokia Research Centre, Finland. We thank the reviewers for their insightful feedback and Dr. Klaus Gebel for his constructive comments on the presentation of this manuscript.

## Appendix A. Detailed statistical analysis of the experiments

This appendix (Tables A1 to A5) presents measurement, mean (M), and standard deviation (SD) of dependent variables and statistical analyses of these experiments. For all these analyses we used IBM SPSS Statistics package versions 19 and 20.

**Table A1**

Analysis of Experiment 1: participants judged the egocentric distance to a cube using two different visualizations, and in the presence and absence of graphical depth cue. We ran four mixed factorial ANOVAs to analyze four different dependent variables. There was no main effect of visualization on signed error. The graphical cue significantly reduced signed error—$F(1,18)=10.32$, $p=0.005$, $\eta_p^2=0.36$—for both of the visualizations. We noticed a significant interaction effect between visualization and graphical cue—$F(1,18)=7.56$, $p=0.013$, $\eta_p^2=0.27$. In the presence of a graphical depth cue, Melt Vision caused significantly less errors than Edge-overlay X-ray. In the case of absolute error, we found a significant main effect of visualization—$F(1,18)=11.65$, $p=0.03$, $\eta_p^2=0.39$. Melt Vision had significantly less absolute error than the Edge-overlay X-ray visualization. A similar effect was found for accuracy as Melt vision was significantly better than X-ray vision. While cue-on condition improved depth judgment, it was significantly slower than cue-off condition—$F(1,18)=10.03$, $p=0.005$, $\eta_p^2=0.36$. However, in the case of cue-on condition Melt was significantly ($p<0.001$) faster than X-ray. There was a main effect of distance on signed error—$F(4,72)=8.13$, $p<0.001$, $\eta_p^2=0.31$. Error increased consistently with distance.

| Conditions | Dependent variables | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Signed error (± m) | | Absolute error (m) | | Accuracy (%) | | Response time (ms) | |
| | M | SD | M | SD | M | SD | M | SD |
| X-ray + Cue-on | −6.05 | 16.54 | 13.01 | 11.84 | 86.50 | 11.78 | 11,337.60 | 7088.70 |
| Melt + Cue-on | −1.07 | 9.31 | 6.59 | 6.64 | 92.96 | 7.05 | 10,338.42 | 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 |

**Table A2**

Analysis of Experiment 2: participants judged egocentric distance to a pyramid in three different spatial conditions and we varied between two different tracking methods. We did not find a main effect of spatial arrangement and tracking methods on any of the dependent variables. There were main effects of distance on signed error—$F(11,264)=304.95$, $p<.001$, $\eta_p^2=0.93$ and absolute error—$F(11,264)=144.83$, $p<.001$, $\eta_p^2=0.86$. Overall, with increasing distance signed error and absolute error increased.

| Conditions | Dependent variables | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Signed error (± m) | | Absolute error (m) | | Accuracy (%) | | Response time (s) | |
| | M | SD | M | SD | M | SD | M | SD |
| **Spatial conditions** | | | | | | | | |
| On–On | −26.45 | 31.74 | 32.24 | 25.82 | 50.86 | 26.24 | 6.99 | 3.56 |
| On–Off | −30.07 | 30.98 | 33.93 | 26.68 | 48.09 | 22.28 | 6.28 | 3.32 |
| Off–Off | −29.47 | 24.15 | 29.87 | 23.66 | 56.57 | 19.49 | 7.19 | 3.63 |
| **Tracking methods** | | | | | | | | |
| Computer vision | −29.04 | 27.83 | 31.35 | 25.20 | 53.10 | 22.20 | 6.70 | 3.53 |
| Sensors | −28.28 | 30.48 | 32.68 | 25.70 | 50.58 | 23.90 | 6.94 | 3.52 |

**Table A3**

Analysis of Experiment 3: participants judged the egocentric distance to a pyramid using three different display configurations and two different spatial conditions. We ran a series of repeated measure ANOVAs and found that there was a main effect of display configuration on signed error—$F(2,22)=4.29$, $p=0.027$, $\eta_p^2=0.28$. Interestingly, participants significantly underestimated distance least using small-low condition. We have observed a significant main effect of distance on signed error— $F(11,121)=88.76$, $p<.001$, $\eta_p^2=0.89$ and absolute error—$F(11,121)=78.06$, $p<.001$, $\eta_p^2=0.88$. Expectedly, with increasing distance errors increased and accuracy decreased. There was no main effect of spatial conditions on any of the variables.

| Conditions | Dependent variables | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Signed error (± m) | | Absolute error (m) | | Accuracy (%) | | Response time (s) | |
| | M | SD | M | SD | M | SD | M | SD |
| **Display configurations** | | | | | | | | |
| Small-Low | −22.91 | 28.58 | 27.57 | 24.11 | 59.28 | 24.6 | 6.81 | 4.67 |
| Big-Low | −27.24 | 28.08 | 29.57 | 25.6 | 57.32 | 23.8 | 6.45 | 5.23 |
| Big-High | −28.01 | 27.49 | 29.42 | 25.98 | 57.66 | 24.07 | 6.73 | 5.32 |
| **Spatial conditions** | | | | | | | | |
| On–On | −26.41 | 28.13 | 29.09 | 25.35 | 57.66 | 24.07 | 6.73 | 5.32 |
| Off–Off | −25.71 | 28.12 | 28.62 | 25.14 | 58.51 | 24.18 | 6.64 | 4.44 |

**Table A4**

Analysis of Experiment 4: participants judged ordinal depth between two pyramids and then exocentric distance between them. Ordinal error refers to ordinal judgment; signed and absolute errors refer to exocentric distance judgment; and response time refers to the whole task. There was a main effect of display configuration on signed error —$F(1,11)=8.63$, $p=0.013$, $\eta_p^2=0.44$. Consistent with our earlier experiment, participants using an iPhone underestimated the exocentric distance less. Interestingly, a Chi-Square test indicated a significant main effect of display on the number of overestimations and underestimations—$\chi^2(1,N=960)=55.9$, $p<.001$. The iPhone condition resulted in significantly more number of overestimations than iPad condition. In terms of response time, participants were significantly faster using an iPad than an iPhone—$F(1,11)=10.77$, $p=0.007$, $\eta_p^2=0.49$. A Chi-Square test indicated that iPad (361 out of 460) had significantly more number of correct responses than iPhone (333 out of 460) in terms of ordinal depth perception—$\chi^2(1, N=960)=4.08$, $p=.04$. Expectedly, zone also had a significant effect on ordinal perception—$\chi^2(4, N=960)=32.94$, $p<.001$. Errors consistently increased with distance. We did not find any significant effect of spatial arrangement on any of our dependent variables.

| Conditions | Dependent variables | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Ordinal error (count) | | Signed error (± m) | | Absolute error (m) | | Response time (s) | |
| | Correct | Wrong | M | SD | M | SD | M | SD |
| **Display configurations** | | | | | | | | |
| Small-Low (iPhone) | 333 | 147 | −5.43 | 9.43 | 8.47 | 6.82 | 8.92 | 4.65 |
| Big-High (iPad) | 361 | 119 | −8.56 | 8.05 | 9.39 | 7.06 | 6.34 | 3.36 |
| **Spatial conditions** | | | | | | | | |
| On–On | 340 | 140 | −6.96 | 8.89 | 8.96 | 6.86 | 7.59 | 4.16 |
| Off–Off | 354 | 126 | −7.03 | 8.92 | 8.90 | 7.05 | 7.67 | 4.36 |
| **Zones** | | | | | | | | |
| Zone 1 (30 m–40 m) | 156 | 36 | −3.25 | 6.02 | 5.43 | 4.15 | 7.09 | 4.02 |
| Zone 2 (50 m–60 m) | 152 | 40 | −5.13 | 6.09 | 6.70 | 4.30 | 7.23 | 3.60 |
| Zone 3 (70 m–80 m) | 143 | 49 | −6.56 | 7.65 | 8.23 | 5.80 | 7.42 | 3.84 |
| Zone 4 (90 m–100 m) | 131 | 61 | −8.79 | 9.72 | 10.97 | 7.15 | 8.08 | 4.73 |
| Zone 5 (110 m–120 m ) | 112 | 80 | −11.24 | 11.51 | 13.34 | 8.98 | 8.33 | 4.87 |

**Table A5**

Analysis of Experiment 5: see-perceive-score task. Participants subjectively scored their perceived level of information about occluding and occluded regions. Later we calculated the total score by averaging the scores of occluding and occluded regions. We ran one-tailed $t$-tests and statistically significant differences at $p<0.05$ are highlighted in **bold** when Saliency-based X-ray (SB) is superior and in *italics* where Edge-overlay (EO) X-ray is superior. Saliency-based X-ray provides better *foreground* information than Edge-overlay X-ray in all the experimental conditions. In the case of high brightness-low edge, high brightness-high edge, low brightness-low edge, and low brightness-medium edge conditions the differences were significant. In five of the experimental conditions (high brightness-all edges, medium brightness-low and medium edges) Edge-overlay X-ray provided significantly better information about the *background* than Saliency-based X-ray. However, in all other conditions Saliency-based X-ray provided better background information than Edge-overlay X-ray. These differences were significant in the case of medium and low brightness-high edge conditions. Considering both scores for foreground and background information, we found that in the case of high brightness-medium and high edge surfaces, Edge-overlay X-ray performed significantly better. For medium and low brightness-high edge conditions Saliency-based X-ray was significantly better than Edge-overlay X-ray.

| Scored region | Amount of edges | | | | | | Brightness |
|---|---|---|---|---|---|---|---|
| | Low | | Medium | | High | | |
| | EO | SB | EO | SB | EO | SB | |
| **Occluding (foreground)** | **6.44** | **7.68** | 6.20 | 6.92 | **6.48** | **7.56** | High |
| | 5.68 | 6.56 | 7.32 | 8.08 | 5.20 | 6.00 | Medium |
| | **4.92** | **6.08** | **6.24** | **7.24** | 4.72 | 4.48 | Low |
| **Occluded (background)** | *7.44* | *6.24* | *7.4* | *5.24* | *5.68* | *2.64* | High |
| | *7.00* | *5.68* | *6.48* | *5.36* | **4.44** | **6.12** | Medium |
| | 6.68 | 6.76 | 6.84 | 7.20 | **5.32** | **7.12** | Low |
| **Total averaged score** | 7.00 | 7.00 | *6.80* | *6.10* | *6.10* | *5.10* | High |
| | 6.40 | 6.10 | 6.90 | 6.70 | **4.80** | **6.10** | Medium |
| | 5.80 | 6.40 | 6.50 | 7.20 | **5.00** | **5.80** | Low |

## References

Acosta, E.J., Richard, S.J., 1976. The effect of irrelevant information on the stages of processing. J. Motor Behav. 8 (September (3)), 181–187.

Avery, B., Sandor, C., Thomas, B., 2009. Improving spatial perception for augmented reality x-ray vision. In: Proceedings of the IEEE Virtual Reality Conference. IEEE, pp. 79–82.

Azuma, R.T., 1997. A survey of augmented reality. Presence 6, 355–385.

Bane, R., Hollerer, T., 2004. Interactive tools for virtual X-Ray vision in mobile augmented reality. In: Proceedings of the 3rd IEEE/ACM International Symposium on Mixed and Augmented Reality. IEEE Computer Society, pp. 231–239. URL ⟨http://portal.acm.org/citation.cfm?id=1032652.1033720⟩.

Barten, P., 1988. Effect of picture size and definition on perceived image quality. In: Conference Record of the 1988 International Display Research Conference, 1988. pp. 142–145.

Bradshaw, M., Parton, A., Eagle, R., 1998. The interaction of binocular disparity and motion parallax in determining perceived depth and perceived size. Perception 27 (11), 1317–1331.

Bradshaw, M.F., Rogers, B.J., 1996. The interaction of binocular disparity and motion parallax in the computation of depth. Vis. Res. 36 (21), 3457–3468.

Dey, A., Cunningham, A., Sandor, C., 2010. Evaluating depth perception of photorealistic mixed reality visualizations for occluded objects in outdoor environments. In: Proceedings of ACM Symposium on Virtual Reality Software and Technology. Hong Kong, China, November. pp. 211–218.

Dey, A., Jarvis, G., Sandor, C., Reitmayr, G., 2012. Tablet vs. phone: depth perception in handheld augmented reality. In: Proceedings of the International Symposium on Mixed and Augmented Reality. pp. 187–196.

Dey, A., Jarvis, G., Sandor, C., Wibowo, A., Mattila, V.-V., 2011. An evaluation of augmented reality X-ray vision for outdoor navigation. In: Proceedings of International Conference on Artificial Reality and Telexistence. Osaka, Japan, November, pp. 28–32.

DiCarlo, J., Zoccolan, D., Rust, N., 2012. How does the brain solve visual object recognition? Neuron 73 (February (3)), 415–434.

Dünser, A., Grasset, R., Billinghurst, M., 2008. A survey of evaluation techniques used in augmented reality studies. In: ACM SIGGRAPH ASIA 2008 Courses. ACM, New York, NY, USA, pp. 5:1–5:27. URL ⟨http://dl.acm.org/citation.cfm?doid=1508044.1508049⟩.

Ellis, S.R., Menges, B.M., 1998. Localization of virtual objects in the near visual field. Hum. Factors 40 (3), 415–431.

Feiner, S., MacIntyre, B., Hollerer, T., Webster, A., 1997. A touring machine: prototyping 3d mobile augmented reality systems for exploring the urban environment. In: First International Symposium on Wearable Computers, October 1997. Digest of Papers, pp. 74–81.

Foley, J.M., Ribeiro-Filho, N.P., Silva, J.A.D., 2004. Visual perception of extent and the geometry of visual space. Vis. Res. 44 (2), 147–156.

Ganapathy, S., 2013. Design guidelines for mobile augmented reality: User experience. In: Tony Huang, L.A., Livingston, M. (Eds.), Human Factors in Augmented Reality Environments. Springer, USA, pp. 165–180.

Gillam, B., Palmisano, S., Govan, D., 2011. Depth interval estimates from motion parallax and binocular disparity beyond interaction space. Perception 40 (1), 39–49.

Jones, J.A., II, J. E.S., Singh, G., Ellis, S.R., 2011. Peripheral visual information and its effect on distance judgments in virtual and augmented environments. In: McDonnell, R., Thorpe, S.J., Spencer, S.N., Gutierrez, D., Giese, M. (Eds.), APGV. ACM, pp. 29–36.

Jones, J.A., Swan, J., Singh, G., Kolstad, E., 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, pp. 9–14.

Kruijff, E., Swan, J., Feiner, S., 2010. Perceptual issues in augmented reality revisited. In: 9th IEEE International Symposium on Mixed and Augmented Reality (ISMAR), October 2010, pp. 3–12.

Kuhl, S.A., Thompson, W.B., Creem-Regehr, S.H., 2009. HMD calibration and its effects on distance judgments. ACM Trans. Appl. Percept. 6 (September (3)), 19:1–19:20, URL ⟨http://doi.acm.org/10.1145/1577755.1577762⟩.

Lappin, J., Shelton, A., Rieser, J., 2006. Environmental context influences visually perceived distance. Percept. Psychophys. 68 (4), 571–581, URL ⟨http://dx.doi.org/10.3758/BF03208759⟩.

Levin, C., Haber, R., 1993. Visual angle as a determinant of perceived interobject distance. Percept. Psychophys. 54, 250–259.

Livingston, M.A., Ai, Z., Swan, J., 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, pp. 55–62.

Livingston, M.A., Dey, A., Sandor, C., Thomas, B.H., 2013. Pursuit of x-ray vision for augmented reality. In: Tony Huang, L.A., Livingston, M. (Eds.), Human Factors in Augmented Reality Environments. Springer, USA, pp. 67–107.

Livingston, M.A., Karsch, K., Ai, Z., Gibson, G.O., 2011. User interface design for military AR applications. Virtual Real. 15 (2–3), 175–184.

Livingston, M.A., Swan, J., Gabbard, J. L., Hollerer, T.H., Hix, D., Julier, S.J., Baillot, Y., 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, p. 56.

Loomis, J.M., DaSilva, J.A., Philbeck, J.W., Fukusima, S.S., 1996. Visual perception of location and distance. Curr. Dir. Psychol. Sci. 5 (3), 72–77.

Loomis, J.M., Knapp, J.M., 2003. Visual Perception of Egocentric Distance in Real and Virtual Environments. No. 11. Erlbaum, pp. 21–46.

Loomis, J.M., Philbeck, J.W., Zahorik, P., 2002. Dissociation between location and shape in visual space. J. Exp. Psychol.: Hum. Percept. Perform. 28 (5), 1202–1212.

McCandless, J., Ellis, S.R., Adelstein, B.D., 2000. Localization of a time-delayed, monocular virtual object superimposed on a real environment. Presence 9 (1), 15–24.

Mikhailova, E., Slavutskaya, A., Gerasimenko, N., 2012. Gender differences in the recognition of spatially transformed figures: behavioral data and event-related potentials (ERPS). Neurosci. Lett. 524 (2), 74–78.

Milgram, P., Takemura, H., Utsumi, A., Kishino, F., 1994. Augmented Reality: a Class of Displays on the Reality-Virtuality Continuum, pp. 282–292.

Naya, Y., Yoshida, M., Miyashita, Y., 2001. Backward spreading of memory-retrieval signal in the primate temporal cortex. Science 291 (5504), 661–664, URL ⟨http://www.sciencemag.org/content/291/5504/661.abstract⟩.

Piekarski, W., Thomas, B., 2001. Tinmith-metro: new outdoor techniques for creating city models with an augmented reality wearable computer. In: ISWC apos01: Proceedings of the 5th International Symposium on Wearable Computers, pp. 31–38.

Pollock, B., Burton, M., Kelly, J., Gilbert, S., Winer, E., 2012. The right view from the wrong location: depth perception in stereoscopic multi-user virtual environments. IEEE Trans. Vis. Comput. Graph. 18 (April (4)), 581–588.

Pratte, M.S., Ling, S., Swisher, J.D., Tong, F., 2013. How attention extracts objects from noise. J. Neurophysiol. 110 (6), 1346–1356, http://dx.doi.org/10.1152/jn.00127.2013. ⟨http://jn.physiology.org/content/early/2013/06/21/jn.00127.2013.abstract⟩.

Predebon, J., 1991. Spatial judgments of exocentric extents in an open-field situation: familiar versus unfamiliar size. Atten. Percept. Psychophys. 50, 361–366, http://dx.doi.org/10.3758/BF03212228.

Rolland, J.P., Ariely, D., Gibson, W., 1995. Towards quantifying depth and size perception in virtual environments. Presence: Teleoper. Virtual Environ. 4 (Winter (1)), 24–49.

Sandor, C., Cunningham, A., Dey, A., Mattila, V., 2010. An augmented reality x-ray system based on visual saliency. In: Proceedings of the IEEE International Symposium of Mixed and Augmented Reality. IEEE, Seoul, Koria, pp. 27–36.

Sandor, C., Cunningham, A., Eck, U., Urquhart, D., Jarvis, G., Dey, A., Barbier, S., Marner, M., Rhee, S., 2009. Egocentric space distortion visualizations for rapid environment exploration in mobile mixed reality. In: Proceedings of the IEEE Virtual Reality Conference 2010. IEEE, Waltham, MA, USA, pp. 47–50.

Sheinberg, D.L., Logothetis, N.K., 2001. Noticing familiar objects in real world scenes: the role of temporal cortical neurons in natural vision. J. Neurosci. 21 (4), 1340–1350, URL ⟨http://www.jneurosci.org/content/21/4/1340.abstract⟩.

Simon, J., Sly, P.E., Vilapakkam, S., 1981. Effect of compatibility of s-r mapping on reactions toward the stimulus source. Acta Psychol. 47 (1), 63–81.

Singh, G., II, J.E.S., Jones, J.A., Ellis, S.R., 2012. Depth judgments by reaching and matching in near-field augmented reality. In: Coquillart, S., Feiner, S., Kiyokawa, K. (Eds.), VR. IEEE, pp. 165–166.

Steinicke, F., Bruder, G., Kuhl, S., 2011. Realistic perspective projections for virtual objects and environments. ACM Trans. Graph. 30 (October (5)), 112-1–112-10, URL ⟨http://doi.acm.org/10.1145/2019627.2019631⟩.

Stix, G., 1992. See-through view. virtual reality may guide physician's hands. Sci. Am. 267 (3), 166.

Stoffels, E., Molen, M.V.D., Keuss, P., 1985. Intersensory facilitation and inhibition: immediate arousal and location effects of auditory noise on visual choice reaction time. Acta Psychol. 58 (1), 45–62.

Sugase-Miyamoto, Y., Matsumoto, N., Kawano, K., 2011. Role of temporal processing stages by inferior temporal neurons in facial recognition. Front. Psychol. 2 (141).

Swan, J., Gabbard, J., 2005. Survey of user-based experimentation in augmented reality. In: Proceedings of 1st International Conference on Virtual Reality, HCI International, July 2005. pp. 1–9.

Swan, J., Jones, A., Kolstad, E., Livingston, M., Smallman, H., 2007. Egocentric depth judgments in optical, see-through augmented reality. IEEE Trans. Vis. Comput. Graph. 13 (May–June (3)), 429–442.

Swan, J.I., Livingston, M.A., Smallman, H.S., Brown, D., Baillot, Y., Gabbard, J.L., 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, pp. 19–26.

Tan, D.S., Czerwinski, M.P., Robertson, G.G., 2006. Large displays enhance optical flow cues and narrow the gender gap in 3d virtual navigation. human factors: the. J. Hum. Factors Ergon. Soc. 48.

Tsuda, T., Yamamoto, H., Kameda, Y., 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, pp. 62–69.

Ullman, S., 2009. Recognizing visual classes and individual objects by semantic hierarchies. In: Dickinson, S.J., Leonardis, A., Schiele, B., Tarr, M.J. (Eds.), Object Categorization. Cambridge University Press, San Francisco, CA, USA.

van Krevelen, D.W.F., Poelman, R., 2010. A survey of augmented reality technologies, applications and limitations. Int. J. Virtual Real. 9 (June (2)), 1–20.

Vanrullen, R., Thorpe, S.J., 2001. The time course of visual processing: from early perception to decision-making. J. Cogn. Neurosci. 13 (May (4)), 454–461, URL ⟨http://dx.doi.org/10.1162/08989290152001880⟩.

Waller, D., Richardson, A.R., 2008. Correcting distance estimates by interacting with immersive virtual environments: effects of task and available sensory information. J. Exp. Psychol.: Applied 14 (1), 61–72.

Wither, J., 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, pp. 92–99.