Taylor & Francis
Taylor & Francis Group

International Journal of Human–Computer Interaction

ISSN: 1044-7318 (Print) 1532-7590 (Online) Journal homepage: http://www.tandfonline.com/loi/hihc20

Visually Perceived Distance Judgments: Tablet-Based Augmented Reality Versus the Real World

J. Edward Swan II, Liisa Kuparinen, Scott Rapson & Christian Sandor

INTERNATIONAL JOURNAL OF HUMAN–COMPUTER INTERACTION
http://dx.doi.org/10.1080/10447318.2016.1265783

Taylor & Francis
Taylor & Francis Group

Visually Perceived Distance Judgments: Tablet-Based Augmented Reality Versus the Real World

J. Edward Swan II[a], Liisa Kuparinen[b], Scott Rapson[c], and Christian Sandor[d]

[a]Department of Computer Science and Engineering, Mississippi State University, Starkville, Mississippi, USA; [b]Department of Computer Science and Information Systems, University of Jyväskylä, Jyväskylä, Finland; [c]School of Information Technology & Mathematical Sciences, University of South Australia, Adelaide, Australia; [d]Interactive Media Design Lab, Graduate School of Information Science, Nara Institute of Science and Technology, Nara, Japan

ABSTRACT

Does visually perceived distance differ when objects are viewed in augmented reality (AR), as opposed to the real world? What are the differences? These questions are theoretically interesting, and the answers are important for the development of many tablet- and phone-based AR applications, including mobile AR navigation systems. This article presents a thorough literature review of distance judgment experimental protocols, and results from several areas of perceptual psychology. In addition to distance judgments of real and virtual objects, this section also discusses previous work in measuring the geometry of virtual picture space and considers how this work might be relevant to tablet AR. Then, the article presents the results of two experiments. In each experiment, observers bisected egocentric distances of 15 and 30 m in tablet-based AR and in the real world, in both indoor corridor and outdoor field environments. In AR, observers bisected the distances to virtual humans, while in the real world, they bisected the distances to real humans. This is the first reported research that directly compares distance judgments of real and virtual objects in a tablet AR system. Four key findings were: (1) In AR, observers expanded midpoint intervals at 15 m, but compressed midpoints at 30 m. (2) Observers were accurate in the real world. (3) The environmental setting—corridor or open field—had no effect. (4) The picture perception literature is important in understanding how distances are likely judged in tablet-based AR. Taken together, these findings suggest the depth distortions that AR application developers should expect with mobile and especially tablet-based AR.

1. Introduction

Recently, a considerable number of augmented reality (AR) applications for tablet computers have been developed. Applications for tablet AR span a wide range of areas, including enhancing paintings in art galleries (van Eck & Kolstee, 2012), furniture layout (Sukan, Feiner, Tversky, & Energin, 2012), visualization of cultural heritage (Haugstvedt & Krogstie, 2012), as components of multi-user systems that include other types of AR devices and computer displays (Thomas, Quirchmayr, & Piekarski, 2003), and AR browsers (Kooper & MacIntyre, 2003; MacIntyre, Hill, Rouzati, Gandy, & Davidson, 2011; Mobilizy, 2016; SPRXmobile, 2016). More recently, mobile AR map-related navigation applications have been developed (Kamilakis, Gavalas, & Zaroliagis, 2016; Morrison et al., 2009; Nurminen, Järvi, & Lehtonen, 2014). Navigation is an important and ubiquitous use case, and previous research has indicated that current AR applications have user experience and usability problems with navigation, finding points of interest, and other tasks related to navigation (Ko, Chang, & Ji, 2013; Olsson & Salo, 2012). Furthermore, problems have been found with mobile AR navigation applications specifically (Rehrl, Häusler, Leitinger, & Bell, 2014). These facts motivate the work described in this article, which studies user understanding of locations and distances in tablet AR.

Compared to maps, either paper or electronic, the main benefits of an AR browser are ease of use and low mental load. For example, a seminal experiment by Shepard and Metzler (1971) showed that for mental rotations, reaction time is linearly proportional to the angle of rotation from the original position. This type of mental rotation is not required for any AR display, as graphics are by definition always correctly aligned with the environment. And, although computerized map applications such as location-based navigation systems typically automatically align the map with the user's current heading, even in this case the user's mental load for matching map locations to environment locations is larger than it is for AR displays, because map users still need to mentally transform the map's bird's-eye view to their first person view. Indeed, an experiment by Tonnis, Sandor, Klinker, Lange, and Bubb (2005) directly compared AR to a correctly aligned schematic map and found that reaction times were significantly higher in the map condition. In addition, with location-based navigation systems that provide directions, there is the persistent problem that users rely too much on turn-by-turn

CONTACT J. Edward Swan II swan@acm.org Department of Computer Science and Engineering, Mississippi State University, 321 Butler Hall, P.O. Box 9637, Starkville, MS 39762-9637, USA.

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/hihc.

directions and ignore the real environment, which hampers spatial knowledge acquisition and leaves users lost and disoriented if the mobile map fails (Huang, Schmidt, & Gartner, 2012). In contrast, AR navigation systems result in spatial knowledge acquisition (Huang et al., 2012), and are much safer when driving (Medenica, Kun, Paek, & Palinko, 2011). However, a map is better suited for overviews and route planning (Lynch, 1960). Therefore, in the cartography community, AR is seen as a promising method for conveying route information, which complements a map's bird's-eye view.

However, AR browsers face another challenge: although with maps it is easy to understand relative distances to points of interest, this is more challenging with AR displays. And, while we believe it is generally desirable for AR users to easily understand distances to points of interest, this is especially valuable when the points are not directly visible, and therefore no real-world depth cues are available (Dey & Sandor, 2014; Kytö, Mäkinen, Häkkinen, & Oittinen, 2013).

As a first step toward addressing these issues, and motivated by AR map-based applications for navigation, in the work reported here we have investigated the visually perceived distance of directly visible virtual objects, in both indoor and outdoor environments. Furthermore, while most previous AR distance perception work has investigated head-mounted displays (HMDs), in this work we have examined AR displays with a handheld form-factor, such as tablets and phones, as these platforms are much more widely used than HMDs.

We therefore describe two experiments that compare visually perceived distance in tablet AR to the real world. Our initial hypothesis was that visually perceived distance would differ between tablet AR and the real world, but we did not know how it would differ. However, there are two bodies of existing work that seem relevant. First, an AR application operating on a tablet is similar in many ways to a framed photograph or picture drawn with accurate linear perspective. A large body of existing work has shown that observers can understand depth and layout in pictures, even when the observer's eye point is quite far removed from the camera's center of projection (Pirenne, 1970; Rogers, 1995; Vishwanath, Girshick, & Banks, 2005), although distances in pictures tend to be compressed relative to the real world (Cutting, 2003; Rogers, 1995). Second, depth perception has been extensively studied in virtual environments seen through HMDs (Swan, Jones, Kolstad, Livingston, & Smallman, 2007; Thompson, Fleming, Creem-Regehr, & Stefanucci, 2011), and has also been studied in large-format displays (Klein, Swan, Schmidt, Livingston, & Staadt, 2009; Ziemer, Plumert, Cremer, & Kearney, 2009). This large body of work has found that judged distances are initially underestimated, but rapidly become more accurate with practice and feedback (Jones, Swan, Singh, & Ellis, 2011; Waller & Richardson, 2008). Although some of these studies examined depth perception in HMD AR, viewing an AR scene on a tablet may be perceptually quite different than viewing AR through an HMD, and therefore it is uncertain how this previous work will apply to tablet AR.

In this article, we take a two-step approach toward understanding depth judgments in tablet AR. First, we have extensively examined the relevant literature relating to both picture perception as well as previous depth judgment studies in AR and Virtual Reality (VR). Here, we summarize and present this work. Second, from the set of previously described depth judgment techniques, we have chosen the *bisection* task, and used this task to conduct two experiments in which we compare depth judgments in tablet AR to the real world, in both indoor corridor and outdoor field environments. The real-world part of our experiments is a replication of a method reported by Lappin, Shelton, and Rieser (2006). In addition, Bodenheimer et al. (2007) have performed a very similar experiment in HMD-based VR. A key insight from this work is the importance of the picture perception literature in understanding how distances are likely to be judged in tablet AR devices.

2. Literature Review

In this section, we first briefly review the long history of attempts to measure visually perceived distance, with a particular focus on the distance judgment tasks that have been developed. We then discuss the geometry of virtual picture space, and describe the important fact that geometric distortions in pictures are typically not perceived. Next, we discuss the more recent efforts to measure visually perceived distance in VR and AR. We conclude with a discussion of direct versus relative distance perception, and also carefully define some of the major terms that have been used to express distance judgments.

2.1. Measuring Visually Perceived Distance

Human distance perception has been extensively studied for well over 100 years (Cutting & Vishton, 1995), and although it is not yet considered to be fully understood, these many years of effort have left a rich legacy of experimental methods and techniques. A central challenge in evaluating distance perception is that *perception*, as a component of conscious experience, cannot be measured directly, and therefore experimental methods involve some observer judgment that can be quantified. Among the most widely used judgments have been *verbal reports*, where observers report the distance from themselves to a target object in terms of meters or some other measurement unit; *matching tasks*, where observers adjust the position of an indicator in one direction to match the distance to a target object in another direction; *bisection tasks*, where observers adjust the position of an indicator to the middle of the distance between themselves and a target object; and *blind action*, where observers perform an action without vision, such as blind walking or blind reaching, to a previously seen target (Thompson et al., 2011).

In addition, Cutting and Vishton (1995), considering basic evolutionary tasks such as walking, running, and throwing, have divided perceptual space into three distance categories, centered on the observer: personal space, action space, and vista space. *Personal space* encompasses arm's reach and slightly beyond; within personal space objects are grabbed and manipulated with the hands. *Action space* can be quickly reached when walking or running, objects can be accurately thrown, and conversations held. Finally, *vista space* is all distances beyond action space; it is the space that a walking or running observer will soon encounter, and contains objects that the observer might be moving toward or away from. Depending on many variables, such as the height of the observer and their experience with the task at hand, the boundary between personal and action space is within 1–3 m, and the boundary between action and vista space is anywhere from about 20 to perhaps 40 m. However, the boundaries between these spaces are not perceptually sharp; each space gradually fades into the next. The idea behind this categorization is that distance perception evolved for different perceptual purposes within each distance category, and therefore we should expect distance perception to operate somewhat differently in each category. For example, within personal space we are most concerned with reaching and grabbing, within action space we are most concerned with moving our body and throwing, while within vista space we are most concerned with planning future movements. In terms of studying distance perception, this line of thinking leads us to anticipate that the structure of perceived space will differ according to distance category (Cutting, 1997).

Within action space, over the past 20 years blind walking has become the dominant method for measuring distance judgments (Thompson et al., 2011). In blind walking, an observer views a target object, and then walks to the object's location with occluded vision. At least two factors explain blind walking's popularity: First, it has been repeatedly found that observers can perform this task with remarkable accuracy in full-cue environments, with little systematic bias (Waller & Richardson, 2008). In addition, blind walking provides an absolute indication of perceived distance, which can be objectively measured in the real world. However, blind walking has rarely been studied for distances over 20 m (Loomis & Philbeck, 2008), and it is clear that the method has some maximum distance limit, likely within action space.

In contrast, methods where the observer remains stationary, such as verbal reports and bisection, can be used to study the entire range of distances, from personal to vista space. In particular, verbal reports have been used to study distances as far as 9 km (Da Silva, 1985). However, many investigations have established that, while verbal reports are generally well fit with linear functions, the slope of the function varies and in general is less than 1.0, meaning that verbal reports typically indicate systematically compressed distances. Furthermore, many concerns have been raised about verbal reports being influenced by cognitive knowledge that is not perceptual in nature (Loomis & Philbeck, 2008). Finally, because verbal reports do not involve positioning a physical object, the indicated distance cannot be objectively measured. Over the past 30 years, these concerns have motivated a search for alternative judgment methods.

Bisection has also been used to study a range of distances, with many studies examining distances up to hundreds of meters (Da Silva, 1985). However, for any distance judgment method, an important question is whether the structure of perceived space, as indicated by that method, is accurate or reveals systematic errors. After all, it is a common experience that humans are able to manipulate their limbs and maneuver their bodies with great dexterity and accuracy, at least within personal and action space. For bisection, this question has been asked by a large number of scientists over many decades. In an important early experiment, Gilinsky (1951) found that bisected intervals were systematically compressed. However, Gilinsky's results came from only two observers, and many later experiments, encompassing hundreds of observers and distances ranging from 0.4 to 296 m, found that observers generally bisect real-world distances accurately (Bodenheimer et al., 2007; Da Silva, 1985; Purdy & Gibson, 1955; Rieser, Ashmead, Talor, & Youngquist, 1990). Despite these results, an important recent experiment by Lappin et al. (2006), on which we have based the work reported here, found bisection results that differ from this large body of work in two important respects: First, they found a significant effect of environment, where observers bisected the same distance differently in different environmental contexts. Second, they found that bisected intervals were generally expanded, which contradicts the repeated finding of either accurate or compressed distance judgments for most other judgment methods, replicated over many decades (Cutting & Vishton, 1995).

2.2. The Geometry of Virtual Picture Space

An AR application running on a tablet or phone is similar to a photograph or picture drawn with accurate linear perspective. Any such picture is like a window into a virtual, three-dimensional picture space that exists on the other side of the picture's surface. Since the development of the theory of linear perspective during the Middle Ages, it has been known that a drawing or painting in accurate perspective must be drawn from a center of projection (CoP), while in photography the camera's position determines the CoP. When an observer's eye point is located at the CoP, the eye receives the same light field as the original camera (Figure 1a), and the observed picture space is geometrically correct (Vishwanath et al., 2005).

Figure 1 illustrates what happens to the geometry of this three-dimensional picture space when the eye point is no longer located at the CoP (Sedgwick, 1991). When the observer's eye point is farther from the picture surface than the CoP, the pixels on the picture surface project farther into picture space (Figure 1b), and therefore objects are geometrically elongated in depth and farther from the observer. When the eye point moves closer to the picture surface than the CoP, the opposite effect happens (Figure 1c), and objects are geometrically compressed and closer to the observer. Lateral movements of the eye point away from the CoP cause objects to geometrically shear in the opposite direction (Figure 1d). In general, moving the eye point away from the CoP causes the geometry of picture space to undergo some combination of shearing and elongation or compression (Sedgwick, 1991; Vishwanath et al., 2005).

However, it is common experience that these geometric distortions are typically not perceived, even when viewing a picture or photograph from many different locations (Rogers, 1995). Indeed, the usefulness of photography, cinema, and perspective drawings is largely based on this perceptual invariance (Cutting, 1987), and over many years, a number of hypotheses for why and how this perceptual invariance operates have been examined (Vishwanath et al., 2005). Nevertheless, when the observer's eye point is moved far enough from the CoP, these geometric distortions can become

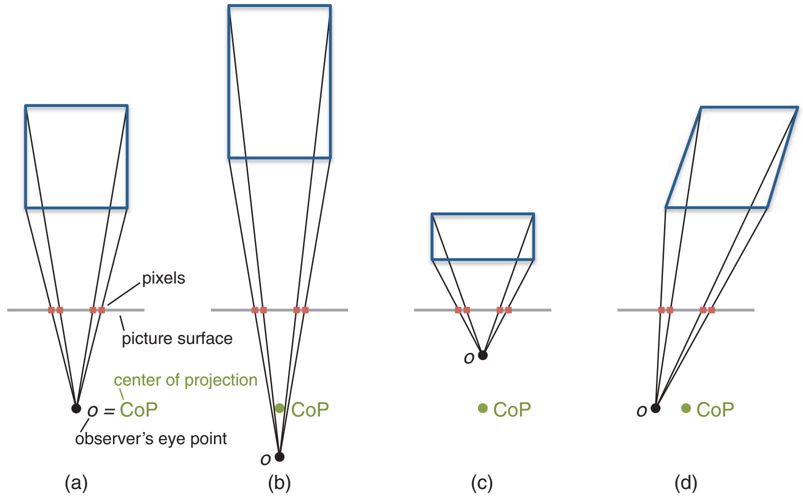

Figure 1. Top-down view of the projection from a picture surface into virtual picture space. (a) The observer's eye point is positioned at the picture's center of projection. (b) The observer is farther from the picture surface than the center of projection. (c) The observer is closer than the center of projection. (d) The observer is to the left of the center of projection.

visible even for pictures drawn in correct perspective (Todorović, 2009), as well as in photography that uses extreme wide angle or telephoto lenses (Pirenne, 1970; Vishwanath et al., 2005). Rogers (1995), in a comprehensive review, finds that displacing the eye from the CoP can introduce perceptual distortions in the geometrically predicted directions (Figure 1), but the strength of these distortions varies widely with setting and task.

Tablets or phones typically have a wide-angle camera, which shows more of the world than would be seen if the tablet were an empty frame (Kruijff, Swan, & Feiner, 2010). Therefore, aligning the eye point with the CoP (Figure 1a) requires positioning the eye very close to the display surface. For example, for the iPad3 that we used in the experiments reported in this article, the CoP is 18.5 cm from the screen. As most users cannot focus this close, Figure 1b illustrates the typical viewing situation for tablet AR, where the eye point is farther than the CoP. This means that object distances will be geometrically expanded; however, as discussed above, this expansion may not be perceived. In addition, many studies have indicated that distances are compressed in pictures, even when the light field matches that of a real-world scene (Figure 1a), and furthermore the degree of compression increases as depicted distance increases (Cutting, 2003; Rogers, 1995). Therefore, the picture perception literature does not clearly predict how depth will be perceived in tablet AR.

2.3. Visually Perceived Distance in Virtual and Augmented Reality

Over the past 20 years, distance perception has been intensively studied in virtual reality (VR); this large body of work has been surveyed by Thompson et al. (2011), Waller and Richardson (2008), and Swan et al. (2007). Most of this research has examined distance perception at action space distances when the virtual environment is seen through an HMD. A consistent and repeated finding is that distances in VR are underestimated relative to the real world. Waller and Richardson (2008) give a compelling meta-analysis of this literature: they analyzed 28 egocentric distance judgment experiments from a variety of laboratories, which used comparable viewing conditions and observer judgments; 14 of these experiments studied VR judgments while the other 14 studied real-world judgments. They found that the VR distance judgments averaged 71% of the correct distance, while the real-world distance judgments averaged 99.9% of the correct distance. However, these VR results require observers to be carefully isolated from the real world. A number of studies have also found that, when observers are allowed to move around in and interact with a VR environment, and receive feedback from their movements, their distance judgments improve and rapidly become veridical (Jones et al., 2011; Mohler, Creem-Regehr, & Thompson, 2006; Waller & Richardson, 2008).

None of the experiments cited by Waller and Richardson (2008) used bisection. However, two experiments have used bisection to study distance perception in HMD VR: Bodenheimer et al. (2007) and Williams, Johnson, Shores, and Narasimham (2008). Both found that bisected intervals were compressed in VR, although Bodenheimer et al. also found expanded intervals at closer distances, and in the same experiment found accurately bisected intervals in the real world.

A small number of experiments have examined how distance perception operates in AR. Most of this work has used blind walking tasks to study action space distances, and presented virtual objects through an HMD. Swan et al. (2007) found that distance in AR was underestimated relative to the real world, but to a lesser degree than has typically been found for VR. Jones, Swan, Singh, Kolstad, and Ellis (2008) then directly compared AR, VR, and a real-world control condition in the same experiment, and found underestimation in VR, but no underestimation in AR. Contradicting these findings, Grechkin, Nguyen, Plumert, Cremer, and Kearney (2010) found similar amounts of underestimation in AR and VR. However, Jones et al. (2011) explained these contradictory findings by demonstrating that when observers can move while seeing visual flow information from the real world, their AR and VR distance judgments rapidly become accurate and indistinguishable from similar judgments in the real world. However, when observers cannot move while seeing the real world, as was the case in Grechkin et al. (2010), their AR and VR distance judgments remain underestimated. Overall, an important implication of this thread of work is that, because AR users naturally see virtual objects in a real-world context, the VR distance underestimation phenomena is unlikely to exist for HMD AR systems involving walking users.

All of these experiments (Grechkin et al., 2010; Jones et al., 2008, 2011; Swan et al., 2007) involved *optical see-through AR*, where observers view the world through the optical combiners of the HMD. A small number of additional studies examined *video see-through AR*, where observers wear a VR HMD and view the world through an attached video camera. Messing and Durgin (2005) used a blind walking task and a monocular HMD, and found that distances were underestimated to a similar degree to what has typically been found for VR. In contrast, Kytö et al. (2013) used a stereo camera and HMD, and studied the effect of stereo viewing and *auxiliary augmentations*—additional virtual objects placed in close proximity to real objects—on distance judgments of virtual objects. They found that both stereo viewing and auxiliary augmentations improved verbal report and ordinal depth judgment tasks. Kytö, Mäkinen, Tossavainen, and Oittinen (2014) then found similar improvements for matching tasks. However, to fully examine the effect of optical versus video see-through AR on depth judgments, it would be necessary to directly compare both conditions as part of the same experiment. To date, the authors are not aware of any experiments where this has been done.

Distance perception in tablet- and phone-based AR has been examined by Dey, Sandor, and their colleagues (Dey, Cunningham, & Sandor, 2010; Dey, Jarvis, Sandor, & Reitmayr, 2012; Dey & Sandor, 2014; Sandor, Cunningham, Dey, & Mattila, 2010). These evaluations, which used verbal report to examine action to vista space distances, introduced several novel depth visualization methods and verified their effectiveness. In addition, Dey et al. (2012) systematically varied screen size and resolution, and found that a larger screen significantly improves metric distance perception, while a smaller, high resolution screen significantly improves ordinal distance judgments.

2.4. Direct Versus Relative Distance Perception

As discussed above, blind walking is considered to provide a *direct* measure of perceived distance. Bisection, in contrast, provides a measure of perceived distance that is *relative* to the location of a target object (Bingham & Pagano, 1998; Rieser et al., 1990). It is worth more deeply considering the difference between direct and relative measures, as well as what each might mean in terms of perception. Consider Figure 2. Here, observer o is viewing target t. Assume that the observer uses a task such as blind walking to make a direct distance judgment, such as j_u or j_o. As shown in Figure 2, the interval oj_u falls short of the actual distance ot, while oj_o is longer than ot. In this article, we term the interval oj_u an *underestimated* distance judgment, and the interval oj_o an *overestimated* distance judgment. Furthermore, if j_o represents the mean and distribution of many distance judgments, then we term the distance tj_o to be the *constant error* (CE) of j_o, which measures the mean accuracy of the judgments over time. We further term the distribution of many judgments the *variable error* (VE), which measures the *precision* of the judgments over time.

Now, consider instead that the observer determines the bisection b of the interval ot between themselves and the target. This is a relative distance judgment, which does not measure the metric distance ot, but does say something[1] about how the observer perceives the distance ot. Let b_c and b_e represent the mean and distribution of many such bisection judgments. In this article, we term the interval ob_c a *compressed* distance judgment, because ob_c is shorter than the actual midpoint interval om. Likewise, we term the interval ob_e an *expanded* distance judgment,[2] because ob_e is longer than om. Constant and variable errors also apply to collections of these relative distance judgments.

However, now consider further what the compressed interval ob_c means perceptually. In order to match ob_c with b_ct, the observer must see the space between o (themselves) and b_c as

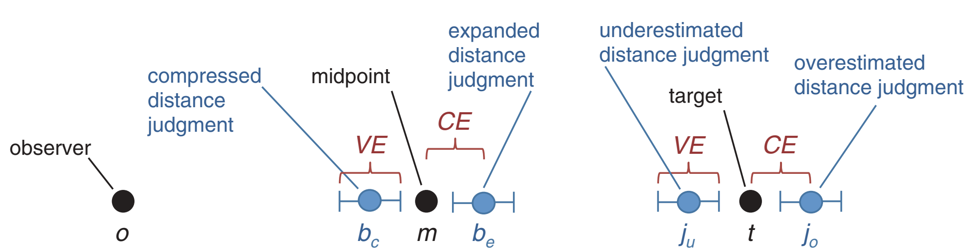

Figure 2. Direct versus relative distance perception.

[1]In particular, the bisected distance gives *ob/ot*, the ratio of the interval *ob* to *ot*; it does not give a metric value for either *ob* or *ot* (Bingham & Pagano, 1998). However, this is only absolutely true when there is no other information to establish the scale of the scene, such as, for example, glowing objects on an otherwise featureless black plane. The complex, real-world environments where we expect tablet AR applications to be used contain objects of known size, such as people, architecture, cars, trees, and so forth, and these have been shown to confer metric scaling information on the scene (Bingham, 1993).

[2]In other experiments that have used bisection, constant compression error has been referred to as *foreshortened* (Bodenheimer et al., 2007; Gilinsky, 1951; Lappin et al., 2006; Rieser et al., 1990), while constant expansion error has been referred to as *anti-foreshortened* (Bodenheimer et al., 2007; Lappin et al., 2006).

being expanded, or longer than it really is, and the space between b_c and t as compressed. Likewise, in order to match ob_e with b_et the observer must see the space between o and b_e as compressed, or shorter than it really is, and the space between b_e and t as expanded. Therefore, if we wanted to speak in terms of what the observer perceives, we could justify reversing the sense of *compressed* and *expanded* in our terminology. However, in this article we will use the terms as defined above, and understand that we are referring to the size of the intervals ob_c and ob_e, and not to the perceptual experience of viewing them.

3. Experiment I

We now describe two experiments[3] that we conducted, which used bisection and the method of Lappin et al. (2006) to study how depth judgments operate in tablet AR. The two experiments differed slightly in how they implemented the bisection method, and they were conducted in different locations.

The purpose of Experiment I was to study how visually perceived distance operates in tablet AR. As discussed above, Lappin et al. (2006) used bisection to measure the visually perceived distance of targets at 15 and 30 m in three different environments: the lobby of a building, an interior corridor, and an outdoor field. In their method a target person stood either 15 or 30 m away, and observers instructed an adjustment person to move to the perceived midpoint between themselves and the target person. On half of the trials, the adjustment person started next to the observer and walked toward the target person (Figure 2: from o toward t), while on the remaining trials the adjustment person started next to the target person and walked toward the observer (Figure 2: from t toward o).

In Experiment I, we closely replicated Lappin et al. (2006) in tablet AR and real-world conditions, with the exception that the adjustment person always started next to the observer and walked toward the target person (Figure 2: from o toward t). This reduced the total number of trials per observer; but later, in Experiment II, we had the adjustment person walk in both directions. In the AR condition the observer only saw the target person on the AR device, while in the real-world condition, the observer saw a real target person. In addition, in the AR condition we attached the tablet to a tripod. Although this differs from typical AR phone usage, where we expect users to hold the phone in their hands, the tripod allowed us to fully replicate and extend Lappin et al.'s (2006) procedure, and it also allowed us to keep the experimental settings as consistent as possible between trials. We ran Experiment I in two different environments: an open field and an interior corridor.

Before running this experiment, we anticipated finding differences in the visually perceived distance to virtual and real targets. These differences would appear as a constant error in the perceived midpoint position that varied by condition. However, we did not know the direction—compression or expansion—in which the constant error would vary. In addition, because the virtual targets were only presented pictorially, we anticipated finding less precisely positioned midpoints for the virtual targets. This would appear as a variable error that is larger for the virtual than for the real targets.

3.1. Method

Apparatus

For an AR tablet, we used an iPad3 (Figure 3), with a resolution of 2048 × 1536 pixels displayed on a 9.7″ screen at 264 dpi. We developed a simple AR system to display a virtual target person in the scene captured by the tablet's camera. The iPad3 camera captures video frames at 1080p resolution.

In order to calibrate a tablet- or phone-based AR system, one must know the field of view (FOV) of the device's camera to a high degree of accuracy. Although the iPad3's data sheet lists the camera frustum as 54° vertical by 40.5° degrees horizontal, we independently measured the FOV in our laboratory by imaging a series of test grids mounted at different distances, which yielded 56° vertical by 43.5° horizontal. As previously mentioned (Section 2.2), this FOV means that the center of projection was located 18.5 cm from the iPad3's screen, about the same distance as the iPad3's width. Overall, we believe that we achieved very comparable quality between the real and virtual targets (see Figure 4; the virtual target person is the farthest in 4b, compare to the real target in 4c).

Our AR system used OpenGL ES2 to render the virtual target person and their shadow. The virtual target person was a photograph of one of the article authors; we calibrated the height of the virtual target by having that author stand next to their virtual self at many distances, including the 15 and 30 m examined in the experiments. We used a billboard to render the virtual target person, and we generated the shadow by warping the billboard texture onto the ground plane and turning it black. The experimenter could interactively adjust the shadow's opacity, direction, and length in order to match real shadows in the experimental environment. Figure 4a shows how well the shadows matched.

We provided orientation tracking by implementing the method described by Kim, Reitmayr, and Woo (2013). In order for the tracking algorithm to track feature points across video frames, the pixels that make up each feature point have to remain the same color as the iPad is moved. Therefore, we had to turn off the camera's automatic exposure control, which normally adapts to changing luminance by adjusting the exposure frame by frame. Although this did not cause problems indoors, we found that outdoor settings were too bright for the tablet's camera. Therefore, in the field environment we additionally mounted a neutral density filter in front of the iPad's camera, which reduced the luminance to an acceptable level.

As discussed in Section 3, we attached the AR tablet to a tripod. For each observer, we adjusted the height of the mounted tablet so that it was at a consistent position relative to the height of their face. While we did not base this adjustment on a precise measurement of the observer's eye height, for all standing observers, looking straight ahead, the top of the tablet was between the tip of their nose and their forehead.

[3]Some preliminary results were reported in a poster abstract (Kuparinen, Swan, Rapson, & Sandor, 2013).

Figure 3. Experimental task and environments: Observers bisected the distance between themselves and a target person by directing an adjustment person to stand at the midpoint. Observers saw both real targets (a, far figure) and virtual targets (b, c, d). Over two experiments, observers experienced three different environments: A frozen lake (a, b), an open field (c), and a corridor (d).

The tripod was mounted perpendicular to the ground and did not tilt, and so was parallel to the observer's face. Observers stood at a tape mark, which we positioned so that they stood a comfortable distance from the tablet; the screen was approximately 55 cm in front of their eyes. We also recorded the experiment by mounting a video camera on another tripod, which we placed a few meters behind the observer.

Environmental Settings

We used two environmental settings, both located on the campus of the University of South Australia: an open field and a corridor. Of the eight observers in the field environment, we ran six in the field shown in Figures 3c and 4a, which was ~40 m wide by ~150 m long. We later ran two additional field observers, but at that time the first field had become a construction zone, so we used a second field that was considerably larger than the first. Both fields were in remote locations that were not commonly accessed by students or employees; none of the observers reported previously visiting either field. The corridor, shown in Figures 3d, 4b, and 4c, was ~2 m wide by ~50 m long, and lined with office doors. The corridor is located in a campus building, and of the eight observers who experienced the corridor condition, three had previously visited the building and were generally familiar with the corridor.

Experimental Design

Within each condition, observers judged targets at two distances, 15 and 30 m, with two repetitions per distance. Before the second repetition, observers moved to a second predefined location, in order to reduce any reliance on environmental cues. Each observer thus made eight judgments: 2 conditions (AR, real) × 2 locations × 2 distances (30, 15 m), which were counterbalanced and nested in the order listed here. We distributed 16 observers between the two environments so that there were 8 observers in each environment, and therefore condition and distance varied within observers while environment varied between observers.

Procedure

Before the experiment, we explained the procedure to the observers. We asked observers to "rely on your inner sense of distance", and to not count steps or rely upon landmarks in the environment. Follow-up discussions with observers suggested that they had not used these kinds of strategies. Observers did not practice the bisection task before the experiment. The procedure took about 25 minutes.

Two experimenters conducted the experiment with each observer: an *adjustment person* and a *target person*. Observers generally stood so their back faced the test area, and only turned around when it was time to conduct a trial.

At the beginning of a real-world trial, the target person positioned themselves at the correct distance from the observer. During the trial the target person stood still. The adjustment person began walking from the observer toward the target person. To allow the observer to see both people clearly, the adjustment person positioned themselves so that, from the perspective of the observer, their

Figure 4. AR view: (a) Field scene, showing a real person and their shadow (right) next to a virtual person and shadow (left). (b) Corridor scene, showing a virtual target person (far figure) and a real adjustment person (near figure). (c) A photograph of the same scene as (b), with a real target person (far figure). The figures differ because 4b is a screenshot from an iPad video feed, while 4c was taken with a high-quality digital camera.

horizontal offset from the target person was about half a meter; see Figures 3a, 4b, and 4c. When the observer believed the adjustment person was half of the distance to the target person, they asked them to stop. The adjustment person stopped and faced the observer, and then encouraged the observer to fine-tune their position by offering to take small steps forward or backward.

For the AR trials, the procedure was as similar as possible to the real-world trials. The target person first positioned themselves at the correct distance from the observer, and the adjustment person adjusted the shadow of the virtual target person so that their shadow visually matched the angle and length of the actual target person's shadow (Figure 5a). The virtual target person was a static image that did not move. After the shadow adjustment, the target person left the test area, and stood out of view while the observer performed the bisection task with the adjustment person. As in the real-world trials, the virtual target person was a different person than the adjustment person, and therefore differed in height.

Observers

We recruited 16 observers (9 male, 7 female) from the students and staff at the University of South Australia. Their ages ranged between 22 and 65, with $M = 34.5$ and $SD = 13.3$, where M is the mean and SD the standard deviation. We rewarded their participation with lemonade and chocolate bars.

environment experiment	observer	direction	Constant Error (meters) AR 30 m	AR 15 m	real 30 m	real 15 m	Variable Error (SD/M) (%) AR 30 m	AR 15 m	real 30 m	real 15 m
corridor	1	away	-1.45	-0.30	-0.25	-0.80	1.6	9.8	2.4	8.4
I	2	away	-2.75	1.25	-0.90	-0.20	11.0	4.0	11.0	3.9
	3	away	0.55	0.20	1.45	1.00	6.8	0.0	10.7	0.0
	4	away	-3.10	2.05	-0.70	1.00	8.3	3.7	13.8	1.7
	5	away	-2.00	0.10	-0.80	-0.15	6.5	1.9	12.0	1.0
	6	away	-2.95	0.80	-1.00	0.70	11.1	15.3	8.1	5.2
	7	away	1.50	1.30	-0.55	0.20	2.6	3.2	0.5	1.8
	8	away	-2.10	0.55	-0.55	-0.95	1.1	2.6	3.4	1.1
field	9	away	-1.70	1.65	-1.10	-0.25	13.8	3.9	1.0	4.9
I	10	away	-4.25	-0.25	-1.30	-0.55	2.0	2.9	1.0	1.0
	11	away	-1.75	0.40	1.10	1.10	0.5	7.2	3.5	0.0
	12	away	-4.95	0.15	-1.05	0.05	9.1	4.6	5.6	2.8
	13	away	-2.80	0.20	-0.55	-0.20	10.4	0.0	0.5	5.8
	14	away	-5.25	0.40	-1.90	-0.45	13.8	7.2	6.5	3.0
	15	away	-0.65	-0.35	1.65	0.80	1.5	1.0	3.0	3.4
	16	away	-1.05	0.80	-0.15	0.40	1.5	1.7	2.4	1.8
ice	17	away	-4.25	0.15	-1.15	0.25	4.6	0.9	1.5	2.7
II		toward	-2.50	0.50	-2.15	-0.35	9.1	5.3	9.4	3.0
	18	away	0.50	1.55	3.35	1.55	4.6	2.3	8.9	2.3
		toward	3.30	1.70	3.35	1.60	10.0	6.1	4.2	1.6
	19	away	-0.30	0.30	-0.75	-0.30	14.4	1.8	3.5	5.9
		toward	-2.15	0.05	-1.80	-0.35	3.9	2.8	9.6	6.9
	20	away	0.60	1.95	-0.90	-0.10	5.4	0.7	4.0	1.9
		toward	1.15	1.30	-0.65	-0.25	3.9	1.6	8.4	2.9
	21	away	-2.80	0.65	-2.10	-0.65	2.3	4.3	5.5	1.0
		toward	-1.95	0.65	-1.10	-0.30	11.4	2.6	0.0	3.9
	22	away	-0.55	0.75	0.65	0.50	3.4	0.9	4.1	5.3
		toward	-1.30	0.85	0.15	0.70	1.0	4.2	1.4	1.7
	23	away	1.95	2.15	0.10	1.00	6.3	11.0	3.7	1.7
		toward	-0.55	1.15	0.45	0.55	10.3	7.4	1.4	4.4
	24	away	-0.55	2.35	1.25	0.70	17.1	12.2	3.0	5.2
		toward	-1.50	0.50	-0.30	0.60	3.1	8.8	2.9	5.2

Figure 5. Constant error and variable error for each observer.

3.2. Results for Each Observer

Figure 5 shows the results for each observer, from both Experiment I (observers 1–16) and Experiment II (observers 17–24). The left-hand section of Figure 5 shows *constant error* in meters, assessed as $M(CE)$, where:

$$CE = judged\,midpoint - correct\,midpoint.$$

As discussed in Section 2.4, $CE < 0$ represents a compressed midpoint judgment; a green bar extending to the left graphically depicts the amount of compression. Likewise, $CE > 0$ represents an expanded midpoint judgment; an amber bar extending to the right depicts the amount of expansion.

The right-hand section of Figure 5 shows *variable error*, where:

$$VE = SD(judged\,midpoints)/M(judged\,midpoints).$$

Variable error is thus a Weber fraction, given by the coefficient of variation *SD/M*; it is reported as a percentage of the mean, and is therefore a scale-free measure of the precision of each observer's judgments in each condition.

In the graphs depicting results (Figure 6), we express constant error as $M(CE/midpoint)(\%)$, averaged over all experimental conditions and expressed as a percentage of the correct midpoint. We express variable error as $RMS(SD/M)(\%)$, *RMS*-averaged[4] between observers, and within each observer calculated as *SD/M* for each experimental condition, as shown in Figure 5.

3.3. Results

Figures 6a and 6b show constant and variable errors from Experiment I, listing them according to the factors of *condition* (AR, real), *environment* (corridor, field), and *target distance* (30, 15 m). Using these factors as a model, we conducted a repeated-measures ANOVA on both constant and variable errors.

Figure 6a shows the constant error. There is a strong condition by distance interaction ($F_{1,14} = 31.4$, $p < 0.001$), as well as a main effect of distance ($F_{1,14} = 27.7$, $p < 0.001$). In the AR condition, observers compressed midpoints at 30 m (−14.5%), and expanded midpoints at 15 m (+7.5%). In the real condition, the data do not show an effect of distance (30 m: −2.7%; 15 m: +1.4%). A priori

[4]The appropriate measure of central tendency for the coefficient of variation is the root mean square (*RMS*), not the mean (*M*).

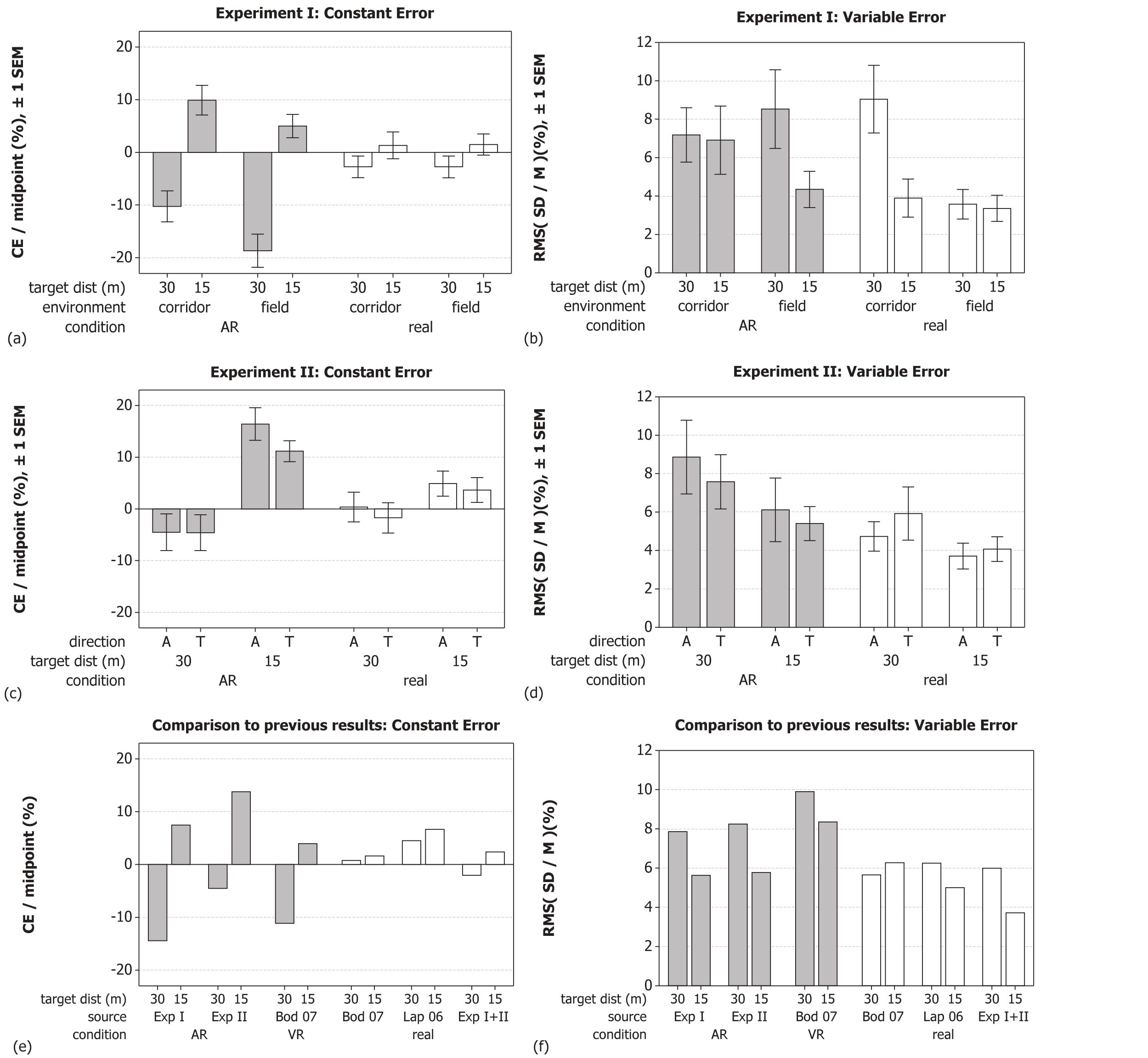

Figure 6. Experimental results: (a, b) Experiment I. (c, d) Experiment II. (e, f) Comparison of our results (source = Exp I, Exp II, Exp I+II) to those reported by Bodenheimer et al. (2007) and Lappin et al. (2006). In all graphs, grey bars represent virtual targets and white bars represent real targets.

paired F tests show that in AR the compressed midpoints at 30 m differ significantly from zero ($F_{1,15} = 23.0$, $p < 0.001$), as do the expanded midpoints at 15 m ($F_{1,15} = 10.1$, $p = 0.006$). However, in the real world, neither midpoint differs significantly from zero (30 m: $F_{1,15} = 2.7$; 15 m: $F_{1,15} = 0.4$). Interestingly, despite testing two very different environments, the data has no main effects or interactions with environment.

Figure 6b shows the variable error. There is a three-way interaction between condition, environment, and distance ($F_{1,14} = 4.6$, $p = 0.05$), as well as a marginal main effect of distance ($F_{1,14} = 4.1$, $p = 0.062$). This is caused by contrary effects for the two conditions: In AR, observers were relatively precise at 15 m in the field (4.3%), compared to their precision in the other three conditions (7.6%), while in the real world, observers were relatively less precise at 30 m in the corridor (9.0%), compared to the other three conditions (3.6%). This is a curious effect, and examining Figure 5 shows that it is not the result of a single, exceptional observer, but reflects the influence of the majority of observers.

3.4. Discussion

The purpose of Experiment I was to study how visually perceived distance operates in tablet AR. As we anticipated, constant error reveals differences in the visually perceived distance of AR and real-world targets. In the real-world condition, observers were accurate, but in the AR condition observers expanded intervals at 15 m and compressed them at 30 m. In addition, constant error did not indicate any effect of environment, and while the design did not have a large amount of power to detect this between-observers effect, the lack of an environment effect is consistent with both Lappin et al. (2006) and Bodenheimer et al. (2007), who also found no constant error differences between field and corridor environments.

We also anticipated that the AR targets would show more variable error than the real targets, and this effect is part of the three-way interaction between condition, environment, and distance. Furthermore, variable error was greater at 30 m compared to 15 m. Finally, in the real world, the interaction suggests more variable error in the corridor than the field, consistent with Lappin et al. (2006).

4. Experiment II

As discussed above, in Experiment I, the adjustment person always started next to the observer and walked toward the target person. However, in Lappin et al. (2006), the target person alternated between starting next to the observer and walking toward the target person (Figure 2: from *o* toward *t*), and starting next to the target person and walking toward the observer (Figure 2: from *t* toward *o*). Although in Experiment I we had the target person walk in one direction to reduce the total number of trials per observer, Experiment I leaves open the possibility that observers might respond differently depending on the direction that the target person walks. Therefore, the purpose of Experiment II was to replicate Experiment I, but with a modified experimental method where the adjustment person walked both toward and away from the observer. Other than this change, we followed the same procedures as Experiment I. We ran Experiment II on a frozen lake, replicating the open field environment of Experiment I.

Before running this experiment, we anticipated AR results generally similar to Experiment I. However, in the AR condition we anticipated the possibility of smaller constant and variable errors when the adjustment person walked toward the observer, because in that case the observer could see the actual, real-world starting position of the adjustment person, and therefore could potentially bisect a real-world interval. In the real condition, we anticipated results similar to Experiment I.

4.1. Method

In Experiment II, we used exactly the same procedures as Experiment I, except for what is noted here.

Environmental Setting

We only used a single environment in Experiment II. Our goal was to replicate the open field environment of Experiment I. However, we conducted this experiment in Finland in early spring, when every field was covered with snow. Therefore, we used a frozen lake, near the University of Jyväskylä, for the experimental setting. The lake, shown in Figures 3a and 3b, is similar to the field environments from Experiment I in that the textured lake surface provided a similar visible texture gradient. In addition, we felt that the frozen lake was interesting because it is among the flattest possible environmental settings.

Experimental Design

In Experiment II the adjustment person walked both away from and toward the observer, so we added the factor *direction* to the design. Therefore, each observer made 16 judgments: 2 conditions (AR, real) × 2 locations × 2 distances (30, 15 m) × 2 directions (away, toward), which were counterbalanced and nested in the order listed here.

Procedure

The procedures in Experiment II were identical to those in Experiment I, with the exception that observers judged each distance twice, with the adjustment person walking in opposite directions.

Observers

We recruited 8 observers (4 male, 4 female) from the staff of the Department of Computer Science and Information Systems at the University of Jyväskylä. The ages of the observers ranged between 30 and 50, with $M = 36.3$ and $SD = 6.1$. Participation was voluntary and not rewarded.

4.2. Results

Figures 6c and 6d show the constant and variable errors from Experiment II.[5] Here, in addition to listing the results according to the factors of *condition* (AR, real) and *target distance* (30, 15 m), the factor *direction* indicates whether the adjustment person walked away (A) from or toward (T) the observer. Using these factors as a model, we conducted a repeated-measures ANOVA on both constant and variable error.

Figure 6c shows the constant error. As in Experiment I, there is a strong condition by distance interaction ($F_{1,7} = 53.4$, $p < 0.001$), as well as a main effect of distance ($F_{1,7} = 47.8$, $p < 0.001$). In the AR condition, observers expanded midpoints at 15 m (+13.8%), but, unlike Experiment I, observers did not compress midpoints at 30 m (−4.5%). In the real condition, as in Experiment I, the data do not show an effect of distance (30 m: −0.7%; 15 m: +4.3%). A priori paired F tests show that the expanded midpoints in AR at 15 m differ significantly from zero ($F_{1,7} = 22.4$, $p = 0.002$), but no other midpoint does (AR, 30 m: $F_{1,7} = 1.2$; real, 30 m: $F_{1,7} = 0.03$; real, 15 m: $F_{1,7} = 1.7$). The data has no main effects or interactions with direction. Figure 6d shows the variable error. There is a marginal main effect of distance ($F_{1,7} = 4.7$, $p = 0.067$), where observers were more precise at 15 m (4.9%) than at 30 m (7.0%). The data has no main effects or interactions with condition or direction.

4.3. Discussion

The purpose of Experiment II was to replicate Experiment I, and in addition have the adjustment person walk both toward and away from the observer. As discussed above, we anticipated AR results generally similar to Experiment I, but with

[5]In Experiment II, one observer had an outlying data value, with CE = +6.3 m, when the other values in the experimental cell ranged from CE = −0.1 to CE = +0.4 m. The video of the trial revealed that the adjustment person did not hear the observer's first instruction to stop. We replaced this value with the median of the remaining values in the cell.

smaller constant and variable errors. The pattern of results is indeed similar: For constant error, we again found expanded midpoints at 15 m, while for variable error the results are very similar between the two experiments. However, the data is equivocal regarding whether constant and variable errors became smaller in Experiment II: The only change in error magnitude is for constant error at 30 m, where midpoints were significantly compressed in Experiment I, but not in Experiment II. In addition, in the AR condition we anticipated smaller constant and variable errors during the trials when the adjustment person started at the location of the target person and walked toward (T) the observer, because during those trials the observer could see the actual, real-world starting position of the adjustment person. However, we found no effect of direction, and so the data does not support this hypothesis; this finding also suggests that only testing one direction in Experiment I did not affect the results. Finally, as predicted, in the real condition both constant and variable errors were similar between the two experiments, and observers continued to accurately bisect targets.

5. Comparison with Previous Results

As previously discussed, these experiments closely replicated the method and design of Lappin et al. (2006). In addition, Bodenheimer et al. (2007), studying virtual reality in an HMD, also closely replicated Lappin et al. This suggests utility in more closely comparing our results to these publications, and we perform this comparison in Figures 6e and 6f. For the AR condition we list both of our experiments separately, but for the real-world data we combined the results.

Figure 6e compares constant error. Over both experiments, in the AR condition the pattern for constant error is that observers expanded midpoints at 15 m and compressed them at 30 m. Likewise, Bodenheimer et al. (2007) also found expanded midpoints at 15 m and compressed midpoints at 30 m for VR targets. Given how different the two virtual environments are—HMD VR and tablet AR—the similarity of this pattern is striking. In addition, as previously mentioned, in the real world a major finding of Lappin et al. (2006) was an overestimation effect for bisection, at both 15 and 30 m. However, we did not replicate this effect; in both experiments we found accurate real-world results, and so did Bodenheimer et al. (2007). These findings are consistent with the hypothesis that in real-world settings bisection is generally accurate, as others have also reported (Da Silva, 1985; Purdy & Gibson, 1955; Rieser et al., 1990).

Figure 6f compares variable errors. Over both experiments, in the real world we found an overall variable error of 5.1%, which is very close to the 5.9% reported by Lappin et al. (2006) and the 6.0% reported by Bodenheimer et al. (2007). However, our AR variable error of 7.0% is somewhat less than the overall 9.2% that Bodenheimer et al. report finding in VR. Furthermore, for virtual targets both we and Bodenheimer et al. found more variable error at 30 than at 15 m. Overall, these experiments suggest that observers are consistently 2–3% less precise when the target is virtual instead of real, and for virtual targets are about 2% less precise at 30 as opposed to 15 m.

6. General Discussion

The purpose of the work reported in this article was to study how visually perceived distance operates in tablet AR. As discussed in Section 1, we were especially motivated by AR map-based applications, where it is desirable for users to understand distances to points of interest. We used bisection, and replicated the method of Lappin et al. (2006). In Experiment I we slightly deviated from Lappin et al.'s method, in that the adjustment person always walked toward the target. However, in Experiment II the adjustment person walked in both directions, and therefore Experiment II fully replicated Lappin et al.'s method. Over both experiments, in AR our primary finding is a pattern of expanded midpoints at 15 m and compressed midpoints at 30 m (Figure 6e). The expansion at 15 m was significantly different than zero over both experiments, but the compression at 30 m only significantly differed from zero in Experiment I. In addition, bisections were also more variable in AR than in the real world. These results contrast with accurate results in the real world, and so we conclude, unsurprisingly, that perceived distance operates differently in tablet AR and the real world.

The pattern of expanded midpoints at 15 m and compressed midpoints at 30 m can be explained by the geometry of virtual picture space and how that geometry is perceived (Section 2.2; Figure 1). In both experiments, the observers' eyes were farther than the tablet's center of projection—the eyes were about 55 cm away, for a center of projection located 18.5 cm in front of the tablet. As shown in Figure 1b, this results in expanded geometry, which can explain the expansion of midpoints at 15 m. In addition, many previous studies have indicated that perceived pictorial distance is increasingly compressed as depicted distance increases (Cutting, 2003; Rogers, 1995), and this can explain the compression of midpoints at 30 m in Experiment I. If this explanation is correct, then we can make two predictions that can be tested in future experiments: (1) We predict additional midpoint expansion for targets closer than 15 m, and additional compression of targets farther than 30 m. And, at some measurable point between 15 and 30 m, midpoints will change from expansion to compression. In addition, (2) if viewing the tablet from an eye point that is further than the camera's center of projection is driving expanded midpoints for targets at 15 m, then modifying the observer's eye point or the camera's center of projection should modify this expansion in a predictable direction (Sedgwick, 1991; Vishwanath et al., 2005).

Finally, as discussed in Section 5, Bodenheimer et al. (2007) found the same pattern of constant error—expansion at 15 m and compression at 30 m—as we did (Figure 6e), despite using HMD VR instead of tablet AR. Could the reasoning given above also explain Bodenheimer et al.'s results? In both cases—our work and Bodenheimer et al.'s—observers saw a pictorial representation of the scene in accurate linear perspective. Furthermore, in both cases the visual scene was truncated, with the observers losing the foreground information from their feet to the bottom of the scene, and it is believed that this truncation is a source of compression and flattening of pictorial depth (Rogers, 1995). However, unlike our experiments, in Bodenheimer et al. observers saw the scene in stereo and from the correct center of projection, and so the similarity of the pattern of results may well be coincidental.

7. Conclusions and Future Work

In this article, we first presented a comprehensive literature review, which reviewed previous work in measuring distance judgments in the real world, in pictures, and in HMD-based VR and AR. To our knowledge, this literature review is the first in the AR field to consider the substantial previous work in picture perception, a topic that seems particularly relevant for tablet-based AR.

We then reported the results of two experiments, which applied a bisection method to study distance judgments in tablet AR. Our bisection method was based on one reported by Lappin et al. (2006) in the real world, and in HMD-based VR by Bodenheimer et al. (2007). In addition to analyzing our results in terms of previous work, we graphically compared our results to AR, VR, and real-world distance judgments from both Lappin et al. (2006) and Bodenheimer et al. (2007).

The novelty of this research is that we are the first to directly compare distance judgments of real and virtual objects in a tablet-based AR system. The results of our investigations are highly significant, as they inform AR application developers of the distortions in depth judgments that they can expect users to make. One of the key insights of our research is the importance of the picture perception literature in understanding how distances are likely to be judged in tablet-based AR devices. These devices fundamentally differ from HMD-based VR and AR in that the observer simultaneously views both virtual picture space and the display surface itself. This makes viewing tablet-based AR similar to viewing a photograph, which can be viewed from many different locations without picture space distortions being perceived (Rogers, 1995).

As discussed in Section 1, in this work we are motivated by numerous AR application areas, especially AR map-based applications for navigation, where it is important for users to understand distances to points of interest. As current AR map and navigation application have problems with spatial perception (Rehrl et al., 2014), the results of this research present important findings on how to better take the user's distance estimations into account when designing AR navigation applications.

Our results suggest a number of useful future experiments and interaction methods:

Handheld Augmented Reality: Because the primary goal of this work was to replicate the bisection method of Lappin et al. (2006) in the real world, while extending the method to work with tablet-based AR, we mounted the tablet on a tripod. This gave us experimental control and repeatability, at some cost in ecological validity: Although a mounted AR display is ecologically valid for some head-up AR applications, such as air traffic control tower tasks (Axholt, Peterson, & Ellis, 2008), the most common use case for tablets and especially phones is that they are handheld. When used this way, user movement introduces motion parallax into the tablet scene (Cutting & Vishton, 1995). Extending the experiment to include a handheld condition, perhaps with specific movements to introduce controllable amounts of motion parallax, would explore the effect of motion parallax on depth judgments.

Additional Distances: Because we replicated the method of Lappin et al. (2006), both experiments only examined targets at 15 and 30 m. However, as discussed in some detail in Section 6, there is much to be learned by replicating the experiment at a wide range of distances, from closer than 15 m to farther than 30 m.

Additional Environments: Also replicating Lappin et al. (2006), we examined only two environments, an indoor corridor and an outdoor field, as well as a frozen lake. Cleary this is a very small sample of the many possible environmental configurations that could be tested.

Blind Walking: As discussed in Section 2.1, blind walking has been extensively used to study distance perception at action space distances, both in the real world and in HMD VR and AR. This suggests using blind walking to study distance perception in tablet AR at action space distances of ~1 to perhaps 15 or 20 m. Blind walking could also be combined with bisection; for example Sinai, Ooi, and He (1998) used both blind walking and perceptual matching to study perceived depth in the same experiment. In addition to the theoretical interest of these experiments, tablet AR has been proposed for applications that operate in action space, such as paintings in art galleries (van Eck & Kolstee, 2012) and furniture layout (Sukan et al., 2012).

Eye Height: In this experiment, although we mounted the tablet on a tripod, we adjusted the height of the tablet according to the height of the observer's face and eyes. Eye height has been found to effect distance judgments in both real and HMD VR environments (Leyrer, Linkenauger, Bülthoff, Kloos, & Mohler, 2011; Ooi & He, 2007), which indicates that in HMD VR and AR, eye height must be modeled accurately for the correct perception of distances and layout. However, as previously discussed in Section 2.2, observers can understand depth and layout in pictures, even when the observer's eye point is quite different from the camera's center of projection (Cutting, 1986; Rogers, 1995). An experiment which systematically varies tablet height relative to eye height could test the importance of eye height on visually perceived distance in tablet AR.

Connectedness: In addition, in AR it seems intuitive that if a virtual object is *connected* to a known real-world location, then observers will more accurately perceive the distance to that virtual object. For example, a virtual sign on a real building could be seen as painted on the building—connected to the building—and therefore perceived as being the same distance as the building. Another kind of virtual-to-real connection involves *shadows*, which connect virtual objects to the ground plane (Figure 4a), and result in more accurate depth perception in AR (Sugano, Kato, & Tachibana, 2003), as well for general 3D computer graphics (Hubona, Wheeler, Shirah, & Brandt, 1999). In addition, as mentioned in Section 2.3 above, Kytö et al. (2013, 2014) have shown that the judged distance of an unconnected virtual object can be improved by showing *auxiliary augmentations*, which are additional connected virtual objects. Kytö et al. (2013) also showed improved depth judgments for *X-ray vision*, where the unconnected virtual object exists behind an opaque surface. Additional designs and experiments could test the effect of different kinds of connection on visually perceived distance in tablet AR.

Depth Cursors: It has long been known that observers can judge the distance of a familiar object more accurately than an unfamiliar, abstract object, because familiar objects allow the use of familiar size as a distance cue (Cutting & Vishton, 1995). Therefore, in this work we used a model of an actual person as a target object. However, our target object is an analogue for an AR browser's *depth cursor*: a user interface element that indicates locations in depth. In the general history of user interface design, there is a long tradition of using abstract shapes for cursors (e.g., Zhai, Buxton, & Milgram, 1994), and this continues for current implementations of AR browsers (Kooper & MacIntyre, 2003; MacIntyre et al., 2011; Mobilizy, 2016; SPRXmobile, 2016) and evaluations in the research community (Dey et al., 2012). We hypothesize that familiar, non-abstract objects, such as our virtual target person, may make more effective AR depth cursors than abstract objects, but this should be directly tested in future experiments. In addition, it may be the case that, because mobile AR users are perceptually adapted to their own body's height, they will perceive the location of a depth cursor which is modeled on their own height more accurately than one which has a different height, or is some abstract shape without a clearly understandable real-world height. Perhaps this height matters more than whether or not the depth cursor looks like a person. We believe there is utility in further investigating these ideas.

Acknowledgments

The authors acknowledge Ville Pekkala and Rebekah Rousi for assistance with data collection. They also thank the anonymous reviewers for many helpful suggestions.

Funding

This material is based upon work supported by the US National Science Foundation under grants 1018413 and 1320909, to J. E. Swan II. In addition, this work was supported by the Nokia Foundation and the University of Jyväskylä International Mobility Grant and Finnish Doctoral Program in User-Centered Information Technology, to L. Kuparinen.

References

Axholt, M., Peterson, S., & Ellis, S. R. (2008). User boresight calibration precision for large-format head-up displays. In B. Fröhlich, K. Ernst, & M. Hachet (Eds.), Presented at the *VRST '08: Proceedings of the 2008 ACM Symposium on Virtual Reality Software and Technology* (pp. 141–148). New York, NY: ACM. doi:10.1145/1450579.1450610

Bingham, G. P. (1993). Perceiving the size of trees: Form as information about scale. *Journal of Experimental Psychology: Human Perception and Performance*, *19* (6), 1139–1161.

Bingham, G. P., & Pagano, C. C. (1998). The necessity of a perception-action approach to definite distance perception: Monocular distance perception to guide reaching. *Journal of Experimental Psychology: Human Perception and Performance*, *24* (1), 145–168.

Bodenheimer, B., Meng, J., Wu, H., Narasimham, G., Rump, B., McNamara, T. P., Carr, T. H., & Rieser, J. J. (2007). Distance estimation in virtual and real environments using bisection. In R. Fleming & M. Langer (Eds.), *Proceedings of the 4th Symposium on Applied Perception in Graphics and Visualization* (pp. 35–40). ACM Press. doi:10.1145/1272582.1272589

Cutting, J. E. (1986). The shape and psychophysics of cinematic space. *Behavior Research Methods, Instruments, & Computers*, *18* (6), 551–558. doi:10.3758/BF03201428

Cutting, J. E. (1987). Rigidity in cinema seen from the front row, side aisle. *Journal of Experimental Psychology: Human Perception and Performance*, *13* (3), 323–334.

Cutting, J. E. (1997). How the eye measures reality and virtual reality. *Behavior Research Methods, Instruments, and Computers*, *29* (1), 27–36. doi:10.3758/BF03200563

Cutting, J. E. (2003). Reconceiving perceptual space. In H. Hecht, R. Schwartz, & M. Atherton (Eds.), *Looking into pictures: An interdisciplinary approach to pictorial space* (pp. 215–238). Cambridge, MA: MIT Press.

Cutting, J. E., & Vishton, P. M. (1995). Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth. In W. Epstein & S. J. Rogers Eds., *Perception of space and motion. Handbook of perception and cognition* (pp. 69–117). San Diego, CA: Elsevier. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/B9780122405303500055

Da Silva, J. A. D. (1985). Scales for perceived egocentric distance in a large open field: Comparison of three psychophysical methods. *The American Journal of Psychology*, *98* (1), 119–144. doi:10.2307/1422771

Dey, A., Cunningham, A., & Sandor, C. (2010). Evaluating depth perception of photorealistic mixed reality visualizations for occluded objects in outdoor environments. In T. Komura & Q. Peng (Eds.), *Proceedings of the 17th ACM Symposium on Virtual Reality Software and Technology* (pp. 211–218). New York, NY: ACM Press. doi:10.1145/1889863.1889911

Dey, A., Jarvis, G., Sandor, C., & Reitmayr, G. (2012). Tablet versus phone: Depth perception in handheld augmented reality. In M. Gandy, K. Kiyokawa, & G. Reitmayr (Eds.), *Presented at the International Symposium on Mixed and Augmented Reality (ISMAR)* (pp. 187–196). Piscataway, NJ: IEEE. doi:10.1109/ISMAR.2012.6402556

Dey, A., & Sandor, C. (2014). Lessons learned: Evaluating visualizations for occluded objects in handheld augmented reality. *International Journal of Human–Computer Studies*, *72* (10–11), 704–716. doi:10.1016/j.ijhcs.2014.04.001

Gilinsky, A. S. (1951). Perceived size and distance in visual space. *Psychological Review*, *58* (6), 460–482. http://doi.org/10.1037/h0061505

Grechkin, T. Y., Nguyen, T. D., Plumert, J. M., Cremer, J. F., & Kearney, J. K. (2010). How does presentation method and measurement protocol affect distance estimation in real and virtual environments? *ACM Transactions on Applied Perception*, *7* (4), 1–18. doi:10.1145/1823738.1823744

Haugstvedt, A.-C., & Krogstie, J. (2012). Mobile augmented reality for cultural heritage: A technology acceptance study. In M. Gandy, K. Kiyokawa, & G. Reitmayr (Eds.), *Presented at the International Symposium on Mixed and Augmented Reality (ISMAR)* (pp. 247–255). Piscataway, NJ: IEEE. doi:10.1109/ISMAR.2012.6402563

Huang, H., Schmidt, M., & Gartner, G. (2012). Spatial knowledge acquisition with mobile maps, augmented reality and voice in the context of GPS-based pedestrian navigation: Results from a field test. *Cartography and Geographic Information Science*, *39* (2), 107–116. doi:10.1559/15230406392107

Hubona, G. S., Wheeler, P. N., Shirah, G. W., & Brandt, M. (1999). The relative contributions of stereo, lighting, and background scenes in promoting 3D depth visualization. *ACM Transactions on Computer–Human Interaction*, *6* (3), 214–242. doi:10.1145/329693.329695

Jones, J. A., Swan, J. E., Singh, G., & Ellis, S. R. (2011). Peripheral visual information and its effect on distance judgments in virtual and augmented environments. In S. Creem-Regehr, & K. Myszkowski (Eds.), *Proceedings of the ACM SIGGRAPH Symposium on Applied Perception in Graphics and Visualization* (APGV) (pp. 29–36). New York, NY: ACM Press. doi:10.1145/2077451.2077457

Jones, J. A., Swan, J. E., Singh, G., Kolstad, E., & Ellis, S. R. (2008). The effects of virtual reality, augmented reality, and motion parallax on egocentric depth perception. In L. T. De Paolis, A. Mongelli (Eds.), *Augmented Reality, Virtual Reality, and Computer Graphics: Third International Conference, AVR 2016, June 15-18, 2016. Proceedings, Part I* (pp. 388–396). Lecce, Italy: ACM Press. doi:10.1145/1394281.1394283

Kamilakis, M., Gavalas, D., & Zaroliagis, C. (2016). Mobile user experience in augmented reality vs maps interfaces: A case study in public transportation. In L. T. De Paolis & A. Mongelli (Eds.), *Augmented Reality, Virtual Reality, and Computer Graphics*, *Lecture Notes in Computer Science (LNCS)* (Vol. 9768, pp. 388 -396). Cham, Switzerland: Springer International Publishing. doi:10.1007/978-3-319-40621-3_27

Kim, H., Reitmayr, G., & Woo, W. (2013). IMAF: In situ indoor modeling and annotation framework on mobile phones. *Personal and Ubiquitous Computing*, *17* (3), 571–582. doi:10.1007/s00779-012-0516-3

Klein, E., Swan, J. E., Schmidt, G. S., Livingston, M. A., & Staadt, O. G. (2009). Measurement protocols for medium-field distance perception in large-screen immersive displays. In D. Reiners, A. Steed, R. Lindeman (Eds.), *Proceedings of IEEE Virtual Reality 2009 (IEEE VR 2009)* (pp. 107–113) Piscataway, NJ: IEEE. doi:10.1109/VR.2009.4811007

Ko, S. M., Chang, W., & Ji, Y. G. (2013). Usability principles for augmented reality applications in a smartphone environment. *International Journal of Human–Computer Interaction*, *29* (8), 501–515. doi:10.1080/10447318.2012.722466

Kooper, R., & MacIntyre, B. (2003). Browsing the real-world wide web: Maintaining awareness of virtual information in an AR information space. *International Journal of Human–Computer Interaction*, *16* (3), 425–446. doi:10.1207/S15327590IJHC1603_3

Kruijff, E., Swan, J. E., & Feiner, S. (2010). Perceptual issues in augmented reality revisited. In J. Park, V. Lepetit, & T. Höllerer (Eds.), *Presented at the International Symposium on Mixed and Augmented Reality (ISMAR)* (pp. 3–12). Piscataway, NJ: IEEE. doi:10.1109/ISMAR.2010.5643530

Kuparinen, L., Swan, II, J. E., Rapson, S., & Sandor, C. Depth perception in tablet-based augmented reality at medium- and far-field distances. In J. Geigel & J. K. Stefanucci (Eds.) *Poster Compendium, Proceedings of ACM SIGGRAPH Symposium on Applied Perception (SAP 2013)*, August 2013. New York, NY: ACM. doi:10.1145/2492494.2514654

Kytö, M., Mäkinen, A., Häkkinen, J., & Oittinen, P. (2013). Improving relative depth judgments in augmented reality with auxiliary augmentations. *ACM Transactions on Applied Perception*, *10* (1), 1–21. doi:10.1145/2422105.2422111

Kytö, M., Mäkinen, A., Tossavainen, T., & Oittinen, P. (2014). Stereoscopic depth perception in video see-through augmented reality within action space. *Journal of Electronic Imaging*, *23* (1), 011006. doi:10.1117/1.JEI.23.1.011006

Lappin, J. S., Shelton, A. L., & Rieser, J. J. (2006). Environmental context influences visually perceived distance. *Perception & Psychophysics*, *68* (4), 571–581.

Leyrer, M., Linkenauger, S. A., Bülthoff, H. H., Kloos, U., & Mohler, B. (2011). The influence of eye height and avatars on egocentric distance estimates in immersive virtual environments. In D. Gutierrez & M. Giese (Eds.), *Proceedings of the ACM SIGGRAPH Symposium on Applied Perception in Graphics and Visualization* (pp. 67–74). New York, NY: ACM Press. doi:10.1145/2077451.2077464

Loomis, J. M., & Philbeck, J. W. (2008). Measuring spatial perception with spatial updating and action. In M. Behrmann, R. L. Klatzky, & B. Macwhinney (Eds.), *Embodiment, ego-space, and action* (pp. 1–43). New York, NY: Psychology Press.

Lynch, K. (1960). *The image of the city*. Cambridge, MA: The MIT Press.

MacIntyre, B., Hill, A., Rouzati, H., Gandy, M., & Davidson, B. (2011). The Argon AR web browser and standards-based AR application environment. In G. Reitmayr, J. Park, & G. Welch (Eds.), *Presented at the International Symposium on Mixed and Augmented Reality (ISMAR)* (pp. 65–74). Piscataway, NJ: IEEE. doi:10.1109/ISMAR.2011.6092371

Medenica, Z., Kun, A. L., Paek, T., & Palinko, O. (2011). Augmented reality vs. street views: A driving simulator study comparing two emerging navigation aids. In O. Juhlin & Y. Fernaeus (Eds.), *Presented at the Human–Computer Interaction with Mobile Devices and Services (MobileHCI '11)* (pp. 265–274). New York, NY: ACM Press. doi:10.1145/2037373.2037414

Messing, R., & Durgin, F. H. (2005). Distance perception and the visual horizon in head-mounted displays. *ACM Transactions on Applied Perception*, *2* (3), 234–250. http://doi.org/10.1145/1077399.1077403

Mobilizy. (2016, June). *Wikitude*. Retrieved from http://www.layar.com

Mohler, B. J., Creem-Regehr, S. H., & Thompson, W. B. (2006). The influence of feedback on egocentric distance judgments in real and virtual environments. In R. Fleming & S. Kim (Eds.), *Proceedings of the 3rd Symposium on Applied Perception in Graphics and Visualization* (pp. 9–14). New York, NY: ACM Press. doi:10.1145/1140491.1140493

Morrison, A., Oulasvirta, A., Peltonen, P., Lemmelä, S., Jacucci, G., Reitmayr, G., Näsänen, J., & Juustila, A. (2009). Like bees around the hive: A comparative study of a mobile augmented reality map. In K. Hinckley, M. R. Morris, S. Hudson, & S. Greenberg (Eds.), *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1889–1898). New York, NY: ACM Press. doi:10.1145/1518701.1518991

Nurminen, A., Järvi, J., & Lehtonen, M. (2014). A mixed reality interface for real time tracked public transportation. In D. Gorteman & S. Hietanen (Eds.), *Presented at the 10th ITS European congress*, Paper SP0053. Brussels, Belgium: ERTICO.

Olsson, T., & Salo, M. (2012). Narratives of satisfying and unsatisfying experiences of current mobile augmented reality applications. In E. H. Chi & K. Höök (Eds.), *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2779–2788). New York, NY: ACM Press. doi:10.1145/2207676.2208677

Ooi, T. L., & He, Z. J. (2007). A distance judgment function based on space perception mechanisms: Revisiting Gilinsky's (1951) equation. *Psychological Review*, *114* (2), 441–454. doi:10.1037/0033-295X.114.2.441

Pirenne, M. H. (1970). *Optics, painting & photography*. London, England: Cambridge University Press.

Purdy, J., & Gibson, E. J. (1955). Distance judgment by the method of fractionation. *Journal of Experimental Psychology*, *50* (6), 374–380. doi:10.1037/h0043157

Rehrl, K., Häusler, E., Leitinger, S., & Bell, D. (2014). Pedestrian navigation with augmented reality, voice and digital map: Final results from an in situ field study assessing performance and user experience. *Journal of Location Based Services*, *8* (2), 75–96. doi:10.1080/17489725.2014.946975

Rieser, J. J., Ashmead, D. H., Talor, C. R., & Youngquist, G. A. (1990). Visual perception and the guidance of locomotion without vision to previously seen targets. *Perception*, *19* (5), 675–689.

Rogers, S. (1995). Perceiving pictorial space. In W. Epstein & S. Rogers (Eds.), *Perception of space and motion: Handbook of perception and cognition* (pp. 119–163). San Diego, CA: Academic Press. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/B9780122405303500067

Sandor, C., Cunningham, A., Dey, A., & Mattila, -V.-V. (2010). An augmented reality X-ray system based on visual saliency. In V. L. Park & T. Höllerer (Eds.), *Presented at the International Symposium on Mixed and Augmented Reality (ISMAR 2010)* (pp. 27–36). Piscataway, NJ: IEEE. doi:10.1109/ISMAR.2010.5643547

Sedgwick, H. A. (1991). The effects of viewpoint on the virtual space of pictures. In S. R. Ellis, M. K. Kaiser, & A. C. Grunwald (Eds.), *Pictorial communication in virtual and real environments* (pp. 460–479). London, England: Taylor & Francis.

Shepard, R., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, *171* (972), 701–703.

Sinai, M. J., Ooi, T. L., & He, Z. J. (1998). Terrain influences the accurate judgement of distance. *Nature*, *395* (6701), 497–500. http://doi.org/10.1038/26747

SPRXmobile. (2016, June). *Layar reality browser*. Retrieved from http://www.layar.com

Sugano, N., Kato, H., & Tachibana, K. (2003). The effects of shadow representation of virtual objects in augmented reality. In B. MacIntyre, D. Schmalstieg, & H. Takemura (Eds.), *Presented at the Second IEEE and ACM International Symposium on Mixed and Augmented Reality* (pp. 76–83). Piscataway, NJ: IEEE Comput. Soc. doi:10.1109/ISMAR.2003.1240690

Sukan, M., Feiner, S., Tversky, B., & Energin, S. (2012). Quick viewpoint switching for manipulating virtual objects in hand-held augmented reality using stored snapshots. In M. Gandy, K. Kiyokawa, & G. Reitmayr (Eds.), *Presented at the International Symposium on Mixed and Augmented Reality (ISMAR)* (pp. 217–226). Piscataway, NJ: IEEE. doi:10.1109/ISMAR.2012.6402560

Swan, J. E., Jones, A., Kolstad, E., Livingston, M. A., & Smallman, H. S. (2007). Egocentric depth judgments in optical, see-through augmented reality. *IEEE Transactions on Visualization and Computer Graphics*, *13* (3), 429–442. doi:10.1109/TVCG.2007.1035

Thomas, B. H., Quirchmayr, G., & Piekarski, W. (2003). Through-walls communication for medical emergency services. *International Journal of Human–Computer Interaction*, *16* (3), 477–496. doi:10.1207/S15327590IJHC1603_6

Thompson, W. B., Fleming, R., Creem-Regehr, S., & Stefanucci, J. K. (2011). *Visual perception from a computer graphics perspective*. Boca Raton, FL: CRC Press.

Todorović, D. (2009). The effect of the observer vantage point on perceived distortions in linear perspective images. *Attention, Perception, & Psychophysics*, *71* (1), 183–193. doi:10.3758/APP.71.1.183

Tonnis, M., Sandor, C., Klinker, G., Lange, C., & Bubb, H. (2005). Experimental evaluation of an augmented reality visualization for directing a car driver's attention. In R. Azuma, O. Bimber, & K. Sato (Eds.), *Presented at the International Symposium on Mixed and Augmented Reality (ISMAR)* (pp. 56–59). Piscataway, NJ: IEEE. doi:10.1109/ISMAR.2005.31

van Eck, W., & Kolstee, Y. (2012). The augmented painting: Playful interaction with multi-spectral images. In S. White, H. B.-L. Duh, & J. D. Bolter (Eds.), *Presented at the International Symposium on Mixed and Augmented Reality (ISMAR)* (pp. 65–69). Piscataway, NJ: IEEE. doi:10.1109/ISMAR-AMH.2012.6483990

Vishwanath, D., Girshick, A. R., & Banks, M. S. (2005). Why pictures look right when viewed from the wrong place. *Nature Neuroscience*, *8* (10), 1401–1410. doi:10.1038/nn1553

Waller, D., & Richardson, A. R. (2008). Correcting distance estimates by interacting with immersive virtual environments: Effects of task and available sensory information. *Journal of Experimental Psychology. Applied*, *14* (1), 61–72. doi:10.1037/1076-898X.14.1.61

Williams, B., Johnson, D., Shores, L., & Narasimham, G. (2008). Distance perception in virtual environments. In S. H. Creem-Regehr & K. Myszkowski (Eds.), *Presented at the ACM Symposium on Applied Perception in Graphics and Visualization* (p. 193). New York, NY: ACM Press. doi:10.1145/1394281.1394322

Zhai, S., Buxton, W., & Milgram, P. (1994). The "Silk Cursor": Investigating transparency for 3D target acquisition. In B. Adelson, S. Dumais, & J. Olson (Eds.), *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 459–464). New York, NY: ACM Press. doi:10.1145/191666.191822

Ziemer, C. J., Plumert, J. M., Cremer, J. F., & Kearney, J. K. (2009). Estimating distance in real and virtual environments: Does order make a difference? *Attention, Perception, & Psychophysics*, *71* (5), 1095–1106. doi:10.3758/APP.71.5.1096

About the Authors

J. Edward Swan II is a Professor of Computer Science and Engineering at Mississippi State University. His Ph.D. (1997) in computer science is from Ohio State University, where he studied computer graphics and human–computer interaction. He is currently studying perception in augmented and virtual reality, including methods for representing depth.

Liisa Kuparinen's expertise includes information systems science and human–computer interaction. Her Ph.D. thesis "Designing and evaluating user interfaces of mobile map services: The viewpoint of supporting user's location-awareness" is in the pre-examination phase. After years as a doctoral student, Kuparinen recently began a professional position as a UX Designer.

Scott Rapson received his Bachelor of Electrical and Mechatronics Engineering from the University of South Australia in 2015. During his studies, he worked on Augmented Reality in the Magicvisionlab. He currently works as a software engineer in Adelaide, Australia.

Christian Sandor is an Associate Professor at Nara Institute of Science and Technology, Japan. He previously worked at Canon's Leading-Edge Technology Research Headquarters in Tokyo, Japan. Before joining Canon, he obtained a doctorate in computer science from the Munich University of Technology, Germany.