Egocentric Distance Judgments in Full-Cue Video-See-Through VR Conditions are No Better than Distance Judgments to Targets in a Void

Koorosh Vaziri* Maria Bondy Amanda Bui Victoria Interrante†
Department of Computer Science and Engineering of University of Minnesota

Figure 1: Left to Right: (1) Custom 3D-printed visor with adjustable cameras and neutral-density lens filters, mounted on an optical-see-through head-mounted display. (2) Equipment in use, with portable backpack computer. (3) Gradual reduction of scene detail: raw camera view, Sobel filtered camera view (non-photorealistic style), and full background subtraction (target only).

ABSTRACT

Understanding the extent to which, and conditions under which, scene detail affects spatial perception accuracy can inform the responsible use of sketch-like rendering styles in applications such as immersive architectural design walkthroughs using 3D concept drawings. This paper reports the results of an experiment that provides important new insight into this question using a custom-built, portable video-see-through (VST) conversion of an optical-see-through head-mounted display (HMD). Participants made egocentric distance judgments by blind walking to the perceived location of a real physical target in a real-world outdoor environment under three different conditions of HMD-mediated scene detail reduction: full detail (raw camera view), partial detail (Sobel-filtered camera view), and no detail (complete background subtraction), and in a control condition of unmediated real world viewing through the same HMD.

Despite the significant differences in participants' ratings of visual and experiential realism between the three different video-see-through rendering conditions, we found no significant difference in the distances walked between these conditions. Consistent with prior findings, participants underestimated distances to a significantly greater extent in each of the three VST conditions than in the real world condition. The lack of any clear penalty to task performance accuracy not only from the removal of scene detail, but also from the removal of *all* contextual cues to the target location, suggests that participants may be relying nearly exclusively on context-independent information such as angular declination when performing the blind-walking task. This observation highlights the limitations in using blind walking to the perceived location of a target on the ground to make inferences about people's understanding of the 3D space of the virtual environment surrounding the target. For applications like immersive architectural design, where we seek to verify the equivalence of the 3D spatial understanding derived from virtual immersion and real world experience, additional measures of spatial understanding should be considered.

Keywords: Virtual reality, spatial perception, non-photorealistic rendering

1 INTRODUCTION

Accurate spatial perception is important in many application areas, including but not limited to: architecture, design, engineering, medicine, and manufacturing. However, egocentric distances have historically been found to be underestimated when people are immersed in a virtual environment using a head-mounted display [36]. Although this problem has been shown to be less severe with newer HMDs [3, 6, 18, 33], evidence of spatial compression continues to be regularly observed, and while multiple work-arounds have been proposed to counter egocentric distance underestimation in VR, e.g. [15,23,25,37,38,46], each has its own potential drawbacks, and the root cause of the problem remains unclear. Our present experiment aims to shed additional light on this long-standing question by considering the impact of rendering detail on spatial perception accuracy from a new perspective: (1) by removing detail from a real world view rather than adding detail to a computer-rendered model; (2) by varying the level of the virtually-presented scene detail along a spectrum that extends all the way from ground-truth realism to a void with no content at all; and (3) by addressing this question in the context of a natural outdoor environment devoid of the prominent linear perspective cues that typically dominate the perception of interior spaces.

Our focus on scene detail / rendering realism is relevant for a variety of reasons. While one might assume that a maximally photorealistic rendering style would naturally be the most optimal choice for an immersive VR experience, it has been shown that when current capabilities in modeling and rendering fall short of enabling perfect realism, a deliberately non-photorealistic (NPR) rendering style may be more advantageous, e.g. [8, 27]. The judicious use of sketch-like rendering styles can also be useful to selectively focus attention [5], mitigate emotion contagion [52], or convey uncertainty about scene details [48]. In architecture, research has shown that using a sketch-like rendering style to present a proposed building model engages clients more deeply and invites them to suggest changes and modifications more freely than when a more traditional CAD-style rendering is used [39]. Our present investigations are particularly motivated by a desire to better understand the extent to which users are able to accurately assess the fundamental 3D spatial qualities of a designed environment when immersed in a virtual model created at an early stage in the design process, before details such as surface treatments have been determined.

---

*e-mail: umn@kvaziri.com

†*Senior Member, IEEE*, e-mail: interran@umn.edu

2021 IEEE Virtual Reality and 3D User Interfaces (VR) | 978-1-6654-1838-6/21/$31.00 ©2021 IEEE | DOI: 10.1109/VR50410.2021.00056

2642-5254/21/$31.00 ©2021 IEEE
DOI 10.1109/VR50410.2021.00056

2 BACKGROUND AND RELATED WORK

2.1 Perception

Egocentric distance underestimation in VR has been the subject of much prior research, where causes and possible solutions have been explored in depth, e.g. [13,16,19,24,31,53]. For a comprehensive review, please see [36].

The first researchers to specifically investigate the impact of rendering quality/photorealism on spatial perception in VR were Gooch and Willemsen [10,54]. In one study [54], they asked 12 participants to perform distance estimation via multiple replications of blind-walking to an orange Frisbee at three different distances – 2m, 3.5m and 5m – in each of three different rendering conditions: (1) a real hallway; (2) a panoramic stereo photograph of the same real hallway, taken from a vantage point approximately matched to the participant's eye height and presented in an HMD; and (3) a 3D rendered model of the same hallway. They found that performance in the real-world condition was close to ideal, but that participants significantly underestimated distances in the image-based and computer-rendered conditions equivalently. In another study [10], they asked 12 participants to judge egocentric distances with the same direct-blind walking method in two different versions of the same hallway: (1) real-world viewing; and (2) a 3D model of the same hallway rendered using a line-drawing style, in which edges and creases were depicted using black lines over a white background. Similarly, they discovered that participants made accurate judgments in the real-world condition, but significantly underestimated distances in the line-drawing style view.

Thompson et al. [50] replicated these experiments in a different venue using visually-directed triangulated walking towards targets at distances of 5m, 10m and 15m in a between-subjects experimental design with 48 total participants, 12 in each of four viewing conditions: (1) a real world view of a lobby; (2) an approximately eye-height-matched stereo panoramic photo of the same lobby, viewed in an HMD; (3) an intentionally "low quality" noise-texture-mapped 3D model that structurally matched the real world lobby environment; and (4) a wire-frame version of the same 3D model, rendered using black lines over a white background. Their results also found that distances were estimated accurately in the real-world condition but significantly underestimated to a qualitatively similar extent in each of the three HMD viewing conditions. Based on these results, they concluded that the distance estimation errors observed in the VR conditions were not due to a lack of realism in the graphics rendering: although participants were consistently underestimating distances in every VR condition, reducing the quality of the computer graphics did not reduce performance any further.

In later work, Interrante et al. [13] found that participants made equivalently accurate judgments of distances in HMD-based VR as in the real world (RW) when they entered an unfurnished real world environment and then put on an HMD that immersed them in an exact-matching, high quality photorealistic 3D replica of that same real world environment, created by texture-mapping photographs of the empty real room onto matching 3D geometry. Follow up studies by Interrante et al. [14] confirmed that participants' accurate responses were not driven by their metric memory of the previously-viewed real world space. These findings showed that there is nothing fundamentally inherent to the use of an HMD that forces distances to be mis-perceived. However, Phillips et al. [35] found that participants significantly underestimated distances in the same virtual replica room when it was rendered in a line-drawing style by replacing the original photographic textures with textures consisting of black feature lines drawn over a white background, despite being explicitly told that that they would be immersed in an exact full-scale VR replica of their concurrently-occupied real environment, which they saw themselves in before donning the HMD. These findings suggest that degrading visual realism *could* potentially impede the accuracy of people's egocentric distance judgments, when doing so destroyed a compelling illusion that the VR rendering was a real world view. A follow-up study by Phillips and Interrante [34] found that superimposing the same black lines over the original photographic texture was enough to break the illusion of reality and evoke underestimation errors, confirming that the conditions associated with enabling accurate spatial perception in VR are indeed fragile.

In related work, Kunz et al. [20] found a significant effect of response method in two experiments comparing distance perception accuracy between a highly photorealistic and less-realistically-rendered version of a 3D virtual environment replicating a real space not previously seen by the viewer. They found no impact of rendering quality when distance judgments were made using action-based direct blind walking, but they did find a significant effect of rendering quality when distance judgments were made via verbal report. Among the theories they advanced to explain these findings is the possibility that participants were selectively attending to different information depending on the task they needed to perform.

Recently, Vaziri et al. [51] suggested a new approach to studying the potential impact of rendering realism on distance perception accuracy in VR. To surmount the limitations on visual and experiential realism imposed by current technologies – that no matter how well a virtual environment is modeled and rendered, it's still very difficult to trick people into thinking that what they are seeing in an HMD is actually a real view of their surrounding real environment – Vaziri et al. studied the impact of converting a live real world view to a sketch-like style, akin to the approach used by Legge et al. [22] to study how deficits in visual spatial acuity affect egocentric distance judgments in real world experiments. Using a similar approach as the AR-Rift [47], Vaziri et al. [51] built a custom-made video-see-through visor for an optical-see-through (OST) HMD, enabling them to compare distance estimation accuracy between: (1) a live view of the real world seen through the OST display; (2) a view of the raw stereo images provided by two cameras mounted at the user's own stereo disparity on a removable visor attached to the front of the same HMD; and (3) an NPR line-drawing style view created by applying a Sobel filter to the real-time camera feeds. In a within-subjects experiment across three different real indoor hallways, counterbalanced between subjects across the three different viewing conditions, they found that participants underestimated egocentric distances significantly more in the two video-see-through (VST) conditions than when using a real world view, but they did not find significant differences in distance perception accuracy between the VST viewing conditions. However, all of their viewing conditions featured prominent linear perspective cues, provided by the architecture of the hallways used for testing, and the strength of those cues could have obviated the effect of textural details within the mediated views. Our present experiment extends the work of Vaziri et al. in part by considering the impact of scene detail loss on distance perception accuracy in a natural outdoor environment devoid of prominent linear perspective cues.

We are not aware of previous efforts to study egocentric distance perception in real outdoor environments using VST technology; there is however a large literature on real world distance perception outdoors [26], and in outdoor environments simulated using HMDs worn indoors (e.g [2,6]). Previous work has also compared distance perception accuracy between virtual and augmented reality viewing conditions [17,28]. Researchers have historically found that egocentric distances are accurately perceived (on average) in flat, open, outdoor real world environments [40]. This accuracy can be disrupted, however, by the presence of intervening, non-occluding obstacles [11], and by the proximity of obstacles behind the target [55]. On a common horizontal groundplane, targets that are farther away will appear higher in the visual field. Ooi et al. [32] and Messing and Durgin [30] have shown, in the real world and VR respectively, that people can use the angular declination to a target on the ground, relative to the visual horizon, to estimate its distance.

2.2 Video-See-Through Hardware and Design

In 1992, Bajura et al. [1] proposed putting cameras in front of an HMD to enable a mixed reality experience, and many other video-see-through systems have been developed since then [9,43–45,49]. One of the major issues in building a VST system is the parallax (apparent displacement) that can occur if the camera images are captured from any position other than that of the viewer's actual eyes. Not only does parallax cause discomfort, it can also interfere with the proper perception of distances and sizes [12,56]. The ultimate goal of any VST system is to be parallax-free in every direction. Because one cannot replace the viewers' eyeballs with cameras, a complete parallax-free solution must use mirrors or prisms to redirect each camera's line-of-sight so that their centers of projection match the locations of the eyeballs. Additionally, the cameras need to match the field-of-view (FOV) of the HMD and lens distortion should be corrected [7]. State et al. [43] developed a distortion-free (orthoscopic) and parallax-free VST HMD using off-the-shelf components and a custom 3D-printed mounting system to hold the cameras and mirrors. However there are considerable technical difficulties associated with achieving a fully parallax-free VST system that can also support a wide field of view.

2.3 NPR Software

Many artistic image-based filters have been developed, as surveyed by Kyprianidis et al. [21], including cartoon shaders and painterly effects. Simple Sobel [41] and Canny [4] filters remain popular however, especially for mobile platforms where performance and resources are limited and efficiency is in high demand. The Sobel operator defines the value at a pixel (i,j) from the intensity differences between the horizontally and vertically surrounding pixels. This is a local operation done in the image frame and is prone to frame-to-frame and stereo incoherence; nevertheless we felt that it would be adequate for our purposes.

3 HARDWARE DESIGN

3.1 VST HMD with Adjustable IPD Bracket

As in Vaziri et al. [51], we mounted two Logitech C615 web-cams, stripped from their enclosures, onto a 3D-printed video-see-through visor attachment where they could be moved left and right to match each participant's inter-pupillary distance (IPD). The visor was attached to the front of an nVis ST50 optical-see-through head-mounted display equipped with a foam attachment that blocked the peripheral view of the real world. Although this system was carefully designed to be parallax-free in the horizontal and vertical directions, there is some parallax in depth due to the forward displacement of the cameras from the eyes. As detailed in [51], after a modest field-of-view correction this parallax has a negligible effect on targets placed at distances over 10 ft. away. Since all of the targets in our experiment are placed at distances beyond 10 ft., we were satisfied that the parallax in the depth direction wouldn't cause a visual displacement in the apparent target location. Finally, we had to modify the power module of our HMD so that it could use the USB ports of our backpack computer rather than relying on an AC power source. The total weight of our VST HMD assembly (shown on the left in Fig. 1) was about 4.0 lbs. or 1.8kg including the cord.

3.2 Portable Backpack Computer

To run the experiment, we used an HP Omen VR backpack computer with a pair of spare batteries, allowing us to run participants back-to-back by using one set while the other re-charged. Mounted on the backpack harness with batteries installed, the HP Omen weighed 10.4 lbs. or 4.7kg. While still heavy, this equipment is significantly lighter than the custom computer used in [51]. The total weight of the equipment worn by the participants, including the backpack computer and VST HMD assembly, was 14.4 lbs. or 6.5kg.

3.3 Custom Neutral-Density Filters

Because our experiment was conducted outdoors in an open-field location with varying lighting conditions, controlling the light was an important concern. The hour of day, density of cloud cover, progression of the seasons, and many other variables affected the overall light intensity in this environment. The two tiny web-cams we used to capture the real-time views simply did not have a fast enough shutter to be able to function properly in broad day-light, and to the extent that they did, the resulting pictures lacked important details in the shadows. To overcome this limitation, we made some custom neutral-density filters to attach to the lenses of our cameras to reduce the amount of incoming light.

Our custom-made neutral-density filters (shown mounted on the camera lenses, on the left side of Fig. 1) were constructed using simple plumbing accessories and automotive window tinting films. To enable a variable reduction of the available light, we built two different sets of filters using tinting films with 20% and 5% tints, capable of blocking 80% and 95% of the incoming light respectively. We measured the light intensity of the outdoor environment using a lightmeter app before each trial to ensure that the most appropriate filter was attached to the cameras before moving the participants to their next starting location.

3.4 Altered Orange Cones

For our target-only viewing condition, we needed to be able to robustly remove the entire background, leaving only the target visible. As explained above, our experiment was conducted outdoors under varying light conditions. We discovered that normal plastic orange cones are highly reflective with unpredictable specular reflection patterns, as can be seen in Fig. 2. From some viewing angles, parts of the cone could appear nearly pure white to our cameras instead of the standard orange color, and the non-uniform exposure caused problems for our segmentation algorithm, making the object appear patchy with missing parts in some views. To make matters worse, the uneven reflectance disrupted the stereo-coherence between the left and right images, as shown in the third image of Fig. 2.

To ensure a robust and uniform reflection, we experimented with different fabric colors and ultimately made a custom orange cone by covering an existing 24cm (9 7/16 in) plastic cone with a neon-orange felt fabric making the final cone 26cm (10 15/64 in) tall, as shown on right side of Fig. 2. This fabric color was extremely unique and nowhere to be found in our outdoor environments, making it the best choice for all of our views and not just the target-only view where the color-segmentation was needed (Fig. 1). We constructed multiple prototypes that differed slightly due to their hand-made nature, and from these we picked the one that gave the most uniform result and used it as the target for the entire experiment.

3.5 Software

We used C++, OpenMP parallelization, and OpenCV to simultaneously obtain and render the two web-cam streams for our views. For our NPR line-drawing style view, we used the OpenCV Sobel algorithm and achieved similar performance as [51] – greater than 30 fps on our HP Omen computer.

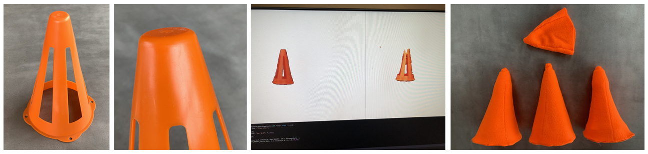

Figure 2: Left to Right: (1) Plastic orange cone. (2) Unpredictable specular reflections. (3) Stereo-incoherence of bare cone. (4) Final altered orange cones covered with Neon Orange felt fabric.

For the background subtraction in our cone-only view, we used color segmentation in the LAB color-space. First, we converted each camera RGB frame into LAB channels, threw away the L-channel, negated both the A and B channels, and added them together. Then, we applied a ~30% binary threshold to this channel, followed by a 3x3 kernel Gaussian blur to reduce salt-and-pepper noise, and 5x5 kernel erosion and dilation morphological operations ["in-order" using the OpenCV "morphologyEx" function (cv::MORPH_OPEN)] to reduce the noise even further, and finally applied another 30% threshold. To create a mask, we negated this channel, and finally applied "reversed-order" erosion and dilation morphs with same 5x5 kernel (cv::MORPH_CLOSE). We used the resulting mask to copy the pixels from the source frame that corresponded to the position of the orange-cone pixels, while the other pixels were set to white. The final result was just an orange cone on a completely white background as shown on the right-sides of both Fig. 1 and Fig. 3.

3.6 System Latency

Latency can be a contributing factor to cybersickness [42], and even small delays have been shown to negatively affect people's sense of presence in VR [29]. Because we had to transfer our camera frames using standard USB 2.0 ports, we had to take extra steps to minimize the overall system latency. To reduce processing time, we down-sampled the captured frames to half of their original size before processing them and then up-sampled back to the original size. We ended up with an end-to-end latency of about 110-140 milliseconds, comparable to [51]. To determine end-to-end latency, we printed timestamps on a separate computer screen and used an external camera to take a side-by-side picture of the HMD lenses and the computer monitor. The difference in timestamps could then be read to compute the latency.

4 USER STUDY

The overarching goal of our experiment was to gain further insight into the factors potentially influencing spatial perception accuracy in non-photorealistically rendered immersive virtual environments. Our ultimate objective was to better understand the extent to which, and conditions under which, conceptual sketch-style rendering might be able to support accurate spatial understanding in VR. While multiple prior studies have suggested that people may be able to perceive 3D space as well in sparsely-detailed virtual environments as in photorealistically rendered ones, the reasons for these findings remain unclear. In the limit, when nothing at all can be seen, there can be no understanding of the surrounding 3D environment. At what point, and under what conditions, does loss of detail become catastrophic to spatial understanding? All of the relevant studies that have been done so far, including [10, 50, 51, 54], were conducted in rectilinear indoor spaces where linear perspective cues dominate. In the absence of such cues, what information do people need to see in order to be able to perceive distances accurately and to achieve an overall accurate spatial understanding of the 3D structure of their surrounding virtual environment?

Figure 3: Left: A photo of the L_3 physical location. Right: Viewing mode M_{Cone} rendered at this location.

Figure 4: Top: Photographs of the L_1 and L_2 physical locations. Bottom: representative views at these locations rendered using M_{NPR}.

To attempt to address these questions, we designed an experiment that compared participants' distance perception accuracy in the context of a real-world outdoor environment devoid of strong linear perspective cues from receding parallel lines, under a full range of conditions of computer-mediated visual quality degradation, from an unprocessed camera view to a view in which all scene information except the target itself was removed. We also included a live real-world view condition to provide a baseline control for the effects of wearing a backpack computer and a restricted field of view HMD.

4.1 Method

Using a within-subjects design, we asked participants to make egocentric distance judgments using direct blind-walking under three different VST viewing conditions: (1) unprocessed real-time camera view (M_{Cam}); (2) line-drawing-style NPR view (M_{NPR}): and (3) just the orange cone rendered on a white background (M_{Cone}), followed by a real-world control condition. The three VST viewing conditions were experienced in a counter-balanced order between participants. Specifically, there were six possible combinations of presentation orders of the three viewing conditions, and equal numbers of participants experienced each combination. The real world condition (M_{RW}) was always done last to avoid any potential carry-over effects from exposure to the real scene or real target [57]. The real world view was experienced under the same physical conditions as the VR views: while wearing the backpack computer and looking through the same HMD with the VST visor removed. Each participant experienced each of the different viewing modes in a different location, denoted by L_1, L_2, and L_3, in a large outdoor recreational area. Locations L_1 and L_2 were counterbalanced between the raw camera and NPR viewing modes, while L_3 was always used in the target-only and real world conditions. Specifically, half of the participants in each of the six different viewing order combinations experienced the M_{Cam} viewing mode in location L_1 and the other half experienced the M_{NPR} viewing mode in that location. Although the M_{Cone} and M_{RW} trials were conducted in the same location (L_3), the location was irrelevant in the M_{Cone} condition because no background could be seen. Fig. 4 shows photographs of the L_1 and L_2 locations, along with screen captured images of representative participant views in these locations rendered using M_{NPR}. Fig. 3 shows a photograph of the L_3 location and a representative participant view at this location rendered using M_{Cone}. All three locations were qualitatively similar, but varied slightly in some features.

The three VST viewing modes and three physical locations were presented in the following counter-balanced way: the first twelve participants experienced the combinations (L_1/M_{Cam}), (L_2/M_{NPR}), (L_3/M_{Cone}) in shuffled order, and the next twelve participants experienced (L_1/M_{NPR}), (L_2/M_{Cam}), (L_3/M_{Cone}) in shuffled order. With this design, we aimed to enable a within-subjects comparison of distance estimation accuracy between rendering styles while avoiding carry-over effects from prior exposure to the same environment under different viewing conditions. Pooling the data across subjects, we also had the ability to verify the absence of a separate significant effect of the different locations on distance judgment accuracy between the M_{Cam} and M_{NPR} conditions. Participants wore the HMD and the backpack computer in all of the experimental conditions, even when turned off in the M_{RW} condition. They also wore disposable ear plugs to mask audio input. The custom adjustable-IPD camera mount was attached in the M_{Cam}, M_{NPR}, and M_{Cone} conditions only. The experiment was conducted with the approval of our university's Institutional Review Board (IRB).

4.1.1 Participants

Through a combination of personal contacts and posted flyers, a total of 29 candidates from our local University community were originally scheduled to fill 24 slots for this experiment. Among these people, two never showed up, and two failed our preliminary stereo-vision test. Of the remaining 25, the study was finished with the first 24 participants, but after all the data was collected, statistical analysis revealed that one participant's responses were more than two standard-deviations different from the mean, so we recruited one last participant as their replacement. The final data is represented by 7 females and 17 males between the ages of 18 and 31 ($\mu = 21.1 \pm 2.8$), who were demographically varied. Each participant was compensated with a $30 gift card to an online retailer.

4.1.2 Procedure

Candidates arrived one at a time at the University Hall and were given written instructions describing the experiment procedure and asked to sign a consent form. We first screened for low vision by asking the candidate to read three lines of letters starting from the top of a Snellen 10ft. Optometric Chart, from a distance of 10 ft., to ensure that all of our participants had a visual spatial acuity of 20/60 or better. If the candidate could wear their corrective lenses in the HMD, we allowed them to wear those lenses during this test. If anyone had failed to pass this vision test, they would have been disqualified from participating in our study, but nobody did.

Next, we measured each candidate's inter-pupillary distance with a ruler and verified the values with a 3D iPhone app. Then, the candidate put on the backpack computer and inserted the ear plugs, and we adjusted the horizontal positions of the cameras on the VST mount to match both their IPD and the locations of their pupils by ensuring that the cameras were aligned with the centers of their eyes. After this, we screened for stereo vision ability by randomly showing the candidate three custom-made random-dot stereograms of increasing complexity on the HMD's screens. Fig. 5 illustrates the images shown: an apple, an open winged bat, and a man playing basketball. A 25-pixel shift was used when creating each stereogram. For each stimulus, we asked the participant to describe with words what they saw. Two candidates failed this test and were thanked and terminated at this stage. As a token of appreciation, terminated candidates were offered free refreshments on their way out.

The candidates who passed both screening tests became official participants in our study. They were asked to remove the HMD, and two experimenters walked with them, carrying all of the supplies and equipment for the participant, to a place that was nearer to location where the experiment would be conducted but from which the open-field areas were still out of sight. We then asked the participant to put on masked safety glasses that only allowed them to see their own two feet, and we guided them, by holding on to the end of a wooden stick, along the rest of the way into and across the recreational area to our main base, which was a park bench. There, we asked the participant—while standing with their back to the field—to remove the safety blindfold glasses and sit down on the park bench seat without looking around. This park bench was strategically positioned so that the participant couldn't see any of the field behind them while seated, and we also specifically asked them to refrain from looking around during the experiment to ensure that they were not exposed to any other vistas within the environment apart from what they could see at the start of each trial.

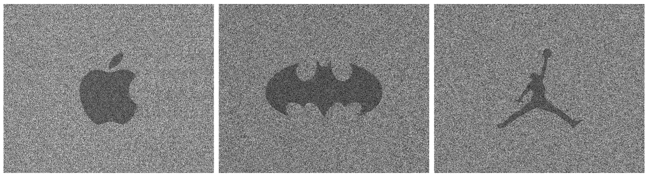

Figure 5: Stereograms shown in random order to test participants' stereo vision (modified here for visualization purposes).

Before each block of trials, and again at the end of the experiment, we asked the participant to fill out a standard Simulator Sickness Questionnaire (SSQ). After they filled out the SSQ, we offered them water and snacks, and asked them to let us know when they were ready to begin. When they acknowledged readiness, we helped them to put on the backpack computer while still seated facing away from the fields, and then to put on the HMD and ear plugs. After this, we asked them to confirm the stereo vision test once again by identifying the object shown in one of the previously-described stereogram pictures (chosen randomly by the experimenter). Two of the participants who had originally passed the stereo-vision test developed difficulty seeing the stereograms after completing one or more blocks of trials, but we let them continue in the experiment. After the stereo vision test, we blocked the participant's view by displaying a black screen on the HMD and then guided them to their designated location by having them hold on to one end of the prop while an experimenter held the other end and slowly walked.

In each environment, participants were asked to perform blind-walking distance judgments to targets indicated by our altered and customized orange cone covered with neon-orange colored felt fabric, placed (one at a time) at six different distances in front of them: 4m, 4.5m, 5m, 6m, 6.5m, and 7m. The distances of 4.5m and 6.5m were shown first, in that order, and were treated as training trials with their results being recorded in the same fashion as the other trials but ignored in the data analysis. The remaining four trials were presented in random order, predetermined using a computer program. Participants did not receive any feedback about their performance at any time during or after the experiment. The starting point for each trial was arbitrarily varied within each location area, and the target was placed manually by measuring out the appropriate distance in front of the participant while they were blindfolded, from wherever they stood, using a soft measuring tape to ensure that the participant couldn't hear the sound of the tape through their ear plugs. The starting location was then marked with a golf-tee on the grass lawn. Once the two experimenters were out of the participant's field of view, the participant was instructed to lift up their blindfold flapper in the M_{RW} viewing mode (Fig. 6) or to open their eyes in the M_{Cam}, M_{NPR}, and M_{Cone} viewing modes while the appropriate graphics were made visible on the HMD by the principal experimenter. Participants then had to take visual aim at the target, say "ready," then close their eyes and either replace the blindfold flapper in the M_{RW} view themselves or have the principal experimenter set the graphics to a black screen in other viewing modes, before walking with their eyes still closed to where they thought the target was located. While their view was blocked, the orange cone was removed from their path by the assisting experimenter. At the end of each trial, an aluminum bracket was placed adjacent to the participant's toes at their stopping

Figure 6: Left: HMD with flapper to block the participant's view in the real-world viewing condition (M_{RW}). Right: M_{RW} trial at L_3 location.

location, and their walked distance from the golf tee to the bracket was measured using a tape measure. While still blindfolded, the participant was then led by the principal experimenter to a slightly different spot within the same location to start their next trial, using simple verbal instructions.

After each block of trials, and with the help of the prop, the participants were escorted—blindfolded and with their eyes closed—back to the park bench base where we repeated the seating procedure described above to ensure that they couldn't see the field after taking off the HMD. During a short break, they were offered some refreshments and asked to fill out another SSQ and a short survey, in which they rated—on a 7-point scale—the visual and functional realism of the environment they had just experienced, as well as their level of "presence" in that environment. Immediately after the block of cone-only trials (M_{Cone}), we also asked the participant to show us—by using their two hands—how tall they thought the cone was, and we recorded their response by measuring the distance between their hands with a tape measure. This was done so that we could later check for any potential effect of cone size mis-judgments on distance errors in that condition where the cone appeared alone.

At the end of the entire experiment, which was always after the last of the real-world viewing trials (M_{RW}), we asked the participants to fill out an exit survey with two open-ended questions asking about their overall experience and any final thoughts. The entire experiment took about 120 minutes on average for each participant.

4.2 Results

We analyzed the impact of viewing mode and distance-shown on distance-walked using a 2-way repeated measures ANOVA. Mauchly's test showed that the assumption of sphericity was met for the within-subjects factor *view* (W = 0.697, p = 0.166), but violated for the within-subjects factor *distance-walked* (W = 0.189, $p < 0.001$). After Greenhouse-Geisser correction, we found both a significant main effect of distance-shown: (F(1.49, 34.33) = 130.601, $p < 0.001$), and a significant main effect of view: (F(2.52, 58.01) = 3.156, $p = 0.039$) on distance walked. Post-hoc tests showed that the distances walked were significantly different between each distance shown, and that participants walked farther in the real world condition than in each of the VST conditions, but there were no significant differences in distance walked between any of the VST conditions. These findings are summarized in Fig. 7 and Fig. 8. The average relative error (RE) and standard deviation (SD) in distances walked across all viewing modes was μ = -9.7% ± 30.5. The average RE/SD in M_{Cam} was μ = -13.1% ± 26.9, in M_{NPR} was μ = -12.3% ± 27.0, in M_{Cone} was μ = -10.9% ± 33.0, and in M_{RW} was μ = -2.4% ± 23.0.

We also ran a three-way ANOVA on the data from the three VST viewing conditions (error ∼ view x position x block order) as a combined sanity check for any unanticipated potential impact of trial location (towards the trees vs towards the river) or any potential impact of the order in which the VST viewing conditions were experienced. We did not include the real world trials in this analysis because the M_{RW} block was always done last. As expected, we found no significant main effect of block order (F(2, 273) = 1.23, $p = 0.29$), nor of position (F(1, 273) = 0.198, $p = 0.66$), or of view, consistent with the analysis described above. We did notice a marginally significant two-way interaction between block order and view (F(4, 273) = 2.31, $p = 0.059$), and a statistically significant three-way interaction (F(2, 273) = 5.72, p ¡ 0.01), but a close inspection of the data did not reveal any clear trends. Fig. 9 shows the raw data used in this analysis.

Hypothesizing that participants' assumptions about the size of the target cone might have affected their judgments of distance in the cone-only viewing condition M_{Cone}, we tested for a correlation between the average signed relative error in each participant's cone size estimate and the average signed relative error in the distances they walked in that condition. We found a very weak ($r^2 = 0.2$) but statistically significant (F(1,22) = 5.34, $p = 0.03$) positive correlation, shown in Fig. 10. Nevertheless, as can also be seen in Fig. 10, more people under-walked than over-walked, and more people over-estimated the cone size than under-estimated it. There is no evidence to suggest that the handful of participants who severely over-walked in the cone-only condition did so because they mis-perceived the cone size.

As described in Sect. 3.3, because of the outdoor lighting and varying weather conditions, we had to use neutral density (ND) filters to reduce and adjust the light input, so that our slow-shutter cameras could function properly. Out of the 72 total blocks of trials done in the M_{Cam}, M_{NPR}, and M_{Cone} viewing modes, we used the 5% (95% block) ND filters 49 times, and the 20% (80% block) ND filters 23 times. Further drilling into the data reveals that the 5% filters were used 18 times in the M_{Cam} condition, 17 times in M_{NPR} and 14 times in M_{Cone}. The 20% filters were used 6 times in M_{Cam}, 7 times in M_{NPR}, and 10 times in M_{Cone}. On average, the 5% ND filter was used 2.13 times more than the 20% ND filter. Unfortunately, there were not enough data points to enable a robust analysis of any potential impact of ND filter use on the distances walked.

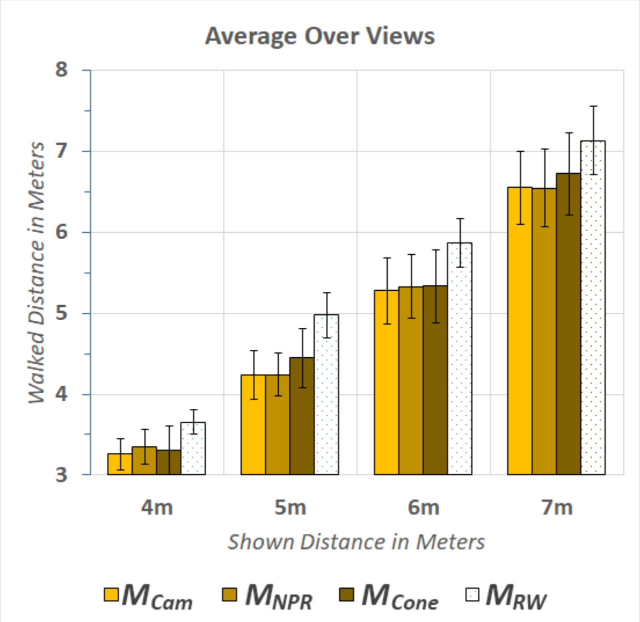

Figure 7: A chart showing the average of the distances walked by each participant in each viewing condition, for each distance shown. The error bars show one standard error on each side.

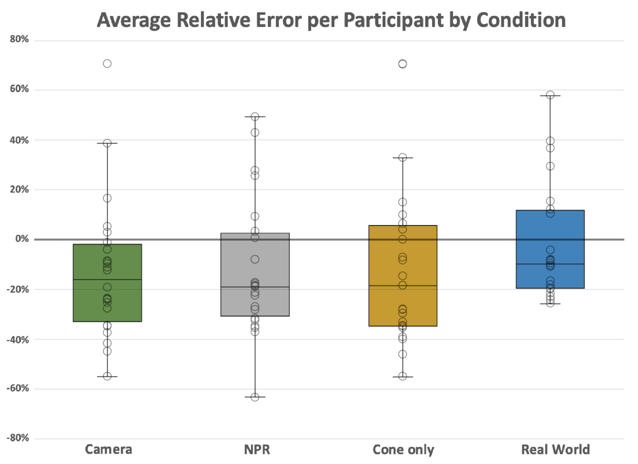

Figure 8: A box-and-whiskers plot of the distribution of the average relative errors in the distance judgments made by each participant in each condition. Each circle represents the average of the relative errors made by a single participant over all of their distance judgments in the corresponding condition, and the box plot encodes the statistics of the distribution of those points.

General linear hypothesis testing of the survey results using a linear mixed-effects model revealed statistically significant differences between all of the viewing conditions within the categories of visual realism, functional realism, and sense of presence at $p<0.001$ (Fig. 11). After performing Bonferroni correction for multiple comparisons, we found that participants reported significantly higher levels of visual realism in the real-world than when experiencing the NPR and target-only viewing conditions, and in the camera viewing condition over the NPR and cone views, but not between the real world and camera views or the NPR and cone-only views. Participants reported significantly higher levels of functional realism in the real-world condition than in any of the three VST viewing conditions, and in the camera view over the NPR and cone-only views, but not between the NPR and cone-only view. Reported levels of presence were significantly higher in the real world condition than in the NPR and cone-only conditions, and in the camera condition over the cone-only view, but not in any of the other pairwise comparisons. These results are shown in Fig. 11.

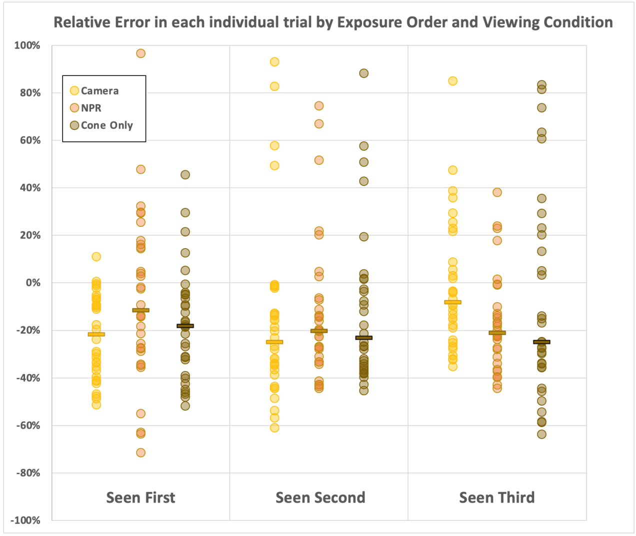

Figure 9: A scatter plot of the average relative errors in the distance judgments made by each individual participant in each viewing and block order condition. The tiny horizontal line icons mark the medians of each distribution. The data on the left represents the results from the first block of trials experienced by each participant, one third of whom experienced each viewing condition as their first exposure. The remaining columns are organized according to the same principle.

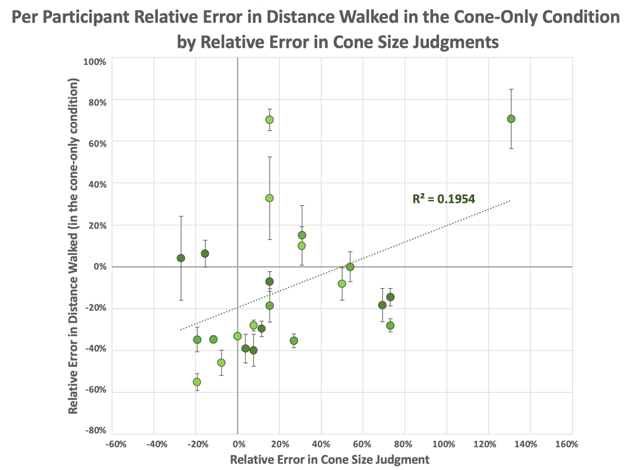

Figure 10: Scatter plot of each participant's relative error in cone size perception vs average relative error in distance judgments, in the cone-only trials. Point shading encodes the block order; darkest green corresponds to participants who experienced the cone-only condition first. Error bars show ± 1 standard error and do not appear along the cone size axis because it was only estimated once.

5 DISCUSSION

The results of this experiment are encouraging with respect to the apparent lack of penalty, also found in previous research [50, 51], for sketch-like viewing scenarios, in that a severe degradation of scene detail did not cause a significant decrease in the accuracy of participants' distance judgments, despite the fact that our experiment was conducted outdoors in an open grassy field lacking the strong linear perspective cues characteristic of indoor environments.

However, the finding that task performance in the cone-only condition was qualitatively as good as in the other conditions is somewhat concerning, considering that nothing was visible at all in that view except the target itself. It makes sense that in the absence of contextual cues, people would base their judgments first, foremost,

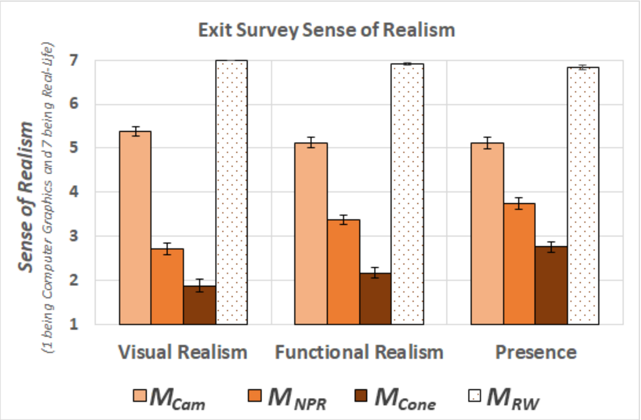

Figure 11: Sense of Realism exit-survey ratings. Error bars show one standard error on each side.

and perhaps solely, on the position of the target relative to their own eye height, in conjunction with the implicit assumption that the target and their own feet are resting on a common horizontal ground plane. It is intriguing to consider that the same thing might also be happening in the other conditions as well. This finding is in line with previous research by Ooi et al. [32]. This discovery raises the possibility that studies using blind walking to targets on the ground to explore the impact of auxiliary environmental factors on spatial perception in VR, such as the quality of the computer graphics or the presence or absence or absolute size of a nearby avatar, may be destined to fail if participants are only focusing on the target when making their judgments and largely ignoring the surround.

The fact that 7% of our population (2 out of 29) failed the initial stereo vision test highlights the importance of using rigorous stereo-testing protocols to objectively and independently validate participants' stereo abilities, as opposed to relying on self-reports of "normal vision". Our finding that 8% of the remaining participants (2 of 25) developed difficulties in resolving stereo images when retested between trial blocks, even though they had previously passed the preliminary testing procedures, suggests a need to additionally consider the possibility (and possible consequences) of progressive eye fatigue over the course of a long experiment.

Finally, we recognize that multiple compromising factors may potentially have affected our findings. Consistent across all of our tested conditions, the HMD we used is older and heavier than current non-OST consumer headsets, there was non-trivial latency from the camera feed, parallax due to the offset between the camera origins and the eyes, a rough open-field terrain, and a gradual change of seasons. Additionally, there were some variations in weather between some blocks of trials for some participants. While we tried our best to minimize the impact of varying light conditions by having two options for the neutral-density filters, the participant experience in full sun with hard shadows might have been different from the experience under overcast skies with no shadows; however way more data points would be needed to explore those possibilities.

6 CONCLUSIONS

This paper makes two important contributions:

First, we extend the results of Vaziri et al. [51] to outdoor environments. In doing so, we clarify that their main result – similarity in distance judgment accuracy between conditions of a highly realistic camera-provided view and a non-realistic NPR-style view – cannot be explained by users' reliance on the prominent linear perspective cues common to each condition, provided by the essential structure of the built environment.

Second, we provide a possible explanation for the findings of Thompson et al. [50], and others, who failed to find a significant impact of the quality of the virtual environment rendering style on people's action-based judgments of egocentric distances to targets on the ground on those environments. Specifically, our discovery that such distances are not only unaffected by a degradation in the "quality" of the portrayal of the environment surrounding the target, but are in fact unaffected by the complete absence of any portrayal of the surrounding environment at all, informs future work in this area by clarifying that – at least under the specific conditions tested - participants may be attending solely to the target itself, and not to the surrounding environment, when completing the blind walking task. This possibility of a singular focus on angular declination to the target location is consistent with prior findings that experimental manipulations that have the potential to interact with angular declination-based judgments, such as manipulating the height of the virtual cameras, increasing the available vertical field of view, or reducing the weight of the HMD, have been found to affect performance on blind walking-based distance estimation tasks in VR, while manipulations that do not interact with processes involved in inferring the angular declination to the target, such as improving the accuracy of surface shading models or enriching the virtual environment with familiar-sized objects, have had minimal effects on such judgments. A better theoretical basis for understanding when, why and how the visual stimulus around a target on the ground may or may not affect people's judgments of its distance takes us one small step closer to better understanding the essential basis of the longstanding problem of spatial mis-perception errors in VR.

ACKNOWLEDGMENTS

This research was supported by the National Science Foundation through grants II-NEW 1305401 and CHS: Small 1526693, and by the Linda and Ted Johnson Digital Design Consortium Endowment.

REFERENCES

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

[2] B. Bodenheimer, J. Meng, H. Wu, G. Narasimham, B. Rump, T. P. Mcnamara, T. H. Carr, and J. J. Rieser. Distance estimation in virtual and real environments using bisection. In *Proc. ACM Symposium on Applied Perception in Graphics and Visualization (APGV)*, pp. 35–40, 2007.

[3] L. E. Buck, M. K. Young, and B. Bodenheimer. A comparison of distance estimation in hmd-based virtual environments with different hmd-based conditions. *ACM Transactions on Applied Perception*, 15(3):Article 21, 2018.

[4] J. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6):679–698, Nov. 1986.

[5] F. Cole, D. DeCarlo, A. Finkelstein, K. Kin, K. Morley, and A. Santella. Directing gaze in 3d models with stylized focus. In *Eurographics Symposium on Rendering*, pp. 377–387, 2006.

[6] S. H. Creem-Regehr, J. K. Stefanucci, W. B. Thompson, N. Nash, and M. McCardell. Egocentric distance perception in the oculus rift (dk2). In *Proc. ACM Symposium on Applied Perception (SAP)*, p. 8, 2015.

[7] E. K. Edwards, J. P. Rolland, and K. P. Keller. Video see-through design for merging of real an virtual environments. In *Proc. Virtual Reality Annual International Symposium (VRAIS)*, pp. 223–233, 1993.

[8] J. Fischer, D. Bartz, and W. Straßer. Stylized augmented reality for improved immersion. In *Proc. IEEE Virtual Reality (VR)*, pp. 195–202, 2005.

[9] H. Fuchs, M. A. Livingstone, R. Raskar, D. Colucci, K. Keller, A. State, J. Crawford, P. Rademacher, S. H. Drake, and A. A. Meyer. Augmented reality visualization for laproscopic surgery. In *Proc. Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pp. 934–943, 1998.

[10] A. A. Gooch and P. Willemsen. Evaluating space perception in npr immersive environments. In *Proc. Non-Photorealistic Animation and Rendering (NPAR)*, pp. 105–110, 2002.

[11] Z. J. He, B. Wu, T. L. Ooi, G. Yarbrough, and J. Wu. Judging egocentric distance on the ground: occlusion and surface integration. *Perception*, 33(7):789–806, 2004.

[12] R. T. Held and M. S. Banks. Misperceptions in stereoscopic displays: a vision science perspective. In *Proc. ACM Symposium on Applied Perception in Graphics and Visualization (APGV*, pp. 23–32, 2008.

[13] V. Interrante, B. Ries, and L. Anderson. Distance perception in immersive virtual environments, revisited. In *Proc. IEEE Virtual Reality (VR)*, pp. 3–10, 2006.

[14] V. Interrante, B. Ries, J. Lindquist, M. Kaeding, and L. Anderson. Elucidating factors that can facilitate veridical spatial perception in immersive virtual environments. *Presence: Teleoperators and Virtual Environments*, 17(2):176–198, 2008.

[15] A. J. Jones, J. E. Swan II, and M. Bolas. Peripheral stimulation and its effect on perceived spatial scale in virtual environments. *IEEE Transactions on Visualization and Computer Graphics*, 19(4):701–710, 2013.

[16] J. A. Jones, D. M. Krum, and M. T. Bolas. Vertical field-of-view extension and walking characteristics in head-worn virtual environments. *ACM Transactions on Applied Perception*, 14(2):Article 9, 2016.

[17] J. A. Jones, J. E. Swan II, G. Singh, E. Kolstad, and S. R. Ellis. The effects of virtual reality, augmented reality, and motion parallax on egocentric depth perception. In *Proc. ACM Symposium on Applied Perception in Graphics and Visualization (APGV*, pp. 9–14, 2008.

[18] J. W. Kelly, L. A. Cherep, and Z. D. Siegel. Perceived space in the htc vive. *ACM Transactions on Applied Perception*, 15(1):Article 2, 2017.

[19] S. A. Kuhl, W. B. Thompson, and S. H. Creem-Regehr. Hmd calibration and its effects on distance judgments. *ACM Transactions on Applied Perception*, 6(3):Art. 19, Aug. 2009.

[20] B. R. Kunz, L. Wouters, D. Smith, W. B. Thompson, and S. H. Creem-Regehr. Revisiting the effect of quality of graphics on distance judgments in virtual environments: A comparison of verbal reports and blind walking. *Attention, Perception, & Psychophysics*, 71(6):1284–1293, Aug. 2009.

[21] J. E. Kyprianidis, J. Collomosse, T. Wang, and T. Isenberg. State of the "art": a taxonomy of artistic stylization techniques for images and video. *IEEE Trans. Visualization and Computer Graphics*, 19(5):866–885, May 2013.

[22] G. E. Legge, R. Gage, Y. Baek, and T. M. Bochsler. Indoor spatial updating with reduced visual information. *PLoS ONE*, 11(3), March 2016.

[23] M. Leyrer, S. A. Linkenauger, H. H. Bülthoff, and B. J. Mohler. Eye height manipulations: a possible solution to reduce underestimation of egocentric distances in head-mounted displays. *ACM Transactions on Applied Perception*, 12(1):Article 1, 2015.

[24] B. Li, J. Walker, and S. A. Kuhl. The effects of peripheral vision and light stimulation on distance judgments through hmds. *ACM Transactions on Applied Perception*, 15(2):Article 12, 2018.

[25] B. Li, R. Zhang, and S. Kuhl. Minification affects action-based distance judgments in oculus rift hmds. In *Proc. ACM Symposium on Applied Perception (SAP)*, p. 4, 2014.

[26] J. M. Loomis and J. M. Knapp. Visual perception and egocentric distance in real and virtual environments. In L. J. Hettinger and M. W. Haas, eds., *Virtual and Adaptive Environments: Applications, Implications, and Human Performance Issues*, chap. 2, pp. 21–46. Lawrence Erlbaum Associates, Mahwah, New Jersey, 2004.

[27] R. McDonnell, M. Breidt, and H. H. Bülthoff. Render me real?: investigating the effect of render style on the perception of animated virtual humans. *ACM Transactions on Graphics*, 31(4):Article 91, 2012.

[28] D. Medeiros, M. Sousa, D. Mendes, A. Raposo, and J. Jorge. Perceiving depth: optical versus video see-through. In *Proc. ACM Symposium on Reality Software and Technology*, pp. 237–240, 2016.

[29] M. Meehan, S. Razzaque, M. C. Whitton, and F. P. Brooks, Jr. Effect of latency on presence in stressful virtual environments. In *Proc. IEEE Virtual Reality (VR)*, pp. 141–148, 2003.

[30] R. Messing and F. H. Durgin. Distance perception and the visual horizon in head-mounted displays. *ACM Transactions on Applied Perception*, 2(3):234–250, 2005.

[31] B. J. Mohler, S. H. Creem-Regehr, and W. B. Thompson. The influence of feedback on egocentric distance judgments in real and virtual environments. In *Proc. ACM Symposium on Applied Perception in Graphics and Visualization (APGV)*, pp. 9–14, 2006.

[32] T. L. Ooi, B. Wu, and Z. J. He. Distance determined by the angular declination below the horizon. *Nature*, 414(6860):197–200, 2001.

[33] A. Peer and K. Ponto. Evaluating perceived distance measures in room-scale spaces using consumer-grade head mounted displays. In *Proc. IEEE Symp. on 3D User Interfaces (3DUI)*, pp. 83–86, 2017.

[34] L. Phillips and V. Interrante. A little unreality in a realistic replica environment degrades distance estimation accuracy. In *Proc. IEEE Virtual Reality (VR)*, pp. 235–236, 2011.

[35] L. Phillips, B. Ries, V. Interrante, M. Kaeding, and L. Anderson. Distance perception in npr immersive virtual environments, revisited. In *Proc. ACM Symposium on Applied Perception in Graphics and Visualization (APGV)*, pp. 11–14, 2009.

[36] R. S. Renner, B. M. Velichkovsky, and J. R. Helmert. The perception of egocentric distances in virtual environments-a review. *ACM Computing Surveys*, 46(2):Article No. 23, Nov. 2013.

[37] A. R. Richardson and D. Waller. Interaction with an immersive virtual environment corrects users' distance estimates. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 49(3):507–517, 2007.

[38] B. Ries, V. Interrante, M. Kaeding, and L. Anderson. The effect of self-embodiment on distance perception in immersive virtual environments. In *Proc. Virtual Reality Software and Technology (VRST)*, pp. 167–170, 2008.

[39] J. Schumann, T. Strothotte, S. Laser, and A. Raab. Assessing the effect of non-photorealistic rendered images in cad. In *Proc. ACM SIGCHI*, pp. 35–41, 1996.

[40] M. J. Sinai, T. L. Ooi, and Z. J. He. Terrain influences the accurate judgement of distance. *Nature*, 395(6701):497–500, 1998.

[41] I. Sobel and G. Feldman. A 3x3 isotropic gradient operator for image processing. 1968.

[42] K. M. Stanney, R. R. Mourant, and R. S. Kennedy. Human factors issues in virtual environments: A review of the literature. *Presence: Teleoperators and Virtual Environments*, 7(4):327–351, Aug. 1998.

[43] A. State, K. P. Keller, and H. Fuchs. Simulation-based design and rapid prototyping of a parallax-free, orthoscopic video see-through head-mounted display. In *Proc. Int. Symp. Mixed and Augmented Reality (ISMAR)*, pp. 28–31, 2005.

[44] A. State, M. A. Livingstone, W. F. Garrett, G. Hirota, M. C. Whitton, E. D. Pisano (MD), and H. Fuchs. Technologies for augmented-reality systems: realising ultrasound-guided needle biopsies. In *Proc. ACM SIGGRAPH*, pp. 439–446, 1996.

[45] A. Steed, Y. W. Adipradana, and S. Friston. The ar-rift 2 prototype. In *Proc. IEEE Virtual Reality (VR)*, pp. 231–232, 2017.

[46] F. Steinicke, G. Bruder, K. Hinrichs, M. Lappe, B. Ries, and V. Interrante. Transitional environments enhance distance perception in immersive virtual reality systems. In *Proc. ACM Symposium on Applied Perception in Graphics and Visualization (APGV)*, pp. 19–26, 2009.

[47] W. Steptoe, S. Julier, and A. Steed. Presence and discernability in conventional and non-photorealistic immersive augmented reality. In *Proc. IEEE International Symposium on Mixed and Augmented Reality*, pp. 213–218, 2014.

[48] T. Strothotte, M. Masuch, and T. Isenberg. Visualizing knowledge about virtual reconstructions of ancient architecture. In *Computer Graphics International*, pp. 36–43, 1999.

[49] A. Takagi, S. Yamazaki, Y. Saito, and N. Taniguchi. Development of a stereo video see-through hmd for ar systems. In *Proc. International Symposium on Augmented Reality (ISAR)*, pp. 68–77, 2000.

[50] W. B. Thompson, P. Willemsen, A. A. Gooch, S. H. Creem-Regehr, J. M. Loomis, and A. C. Beall. Does the quality of the computer graphics matter when judging distances in visually immersive environments? *Presence: Teleoperators and Virtual Environments*, 13(5):560–571, Oct. 2004.

[51] K. Vaziri, P. Liu, S. Aseeri, and V. Interrante. Impact of visual and experiential realism on distance perception in vr using a custom video see-through system. In *Proc. ACM Symposium on Applied Perception (SAP)*, p. 8, 2017.

[52] M. Volonte, B. S. V., H. Chaturvedi, N. Newsome, E. Ebrahimi, T. Roy, S. B. Daily, and T. Fasolino. Effects of virtual human appearance fidelity on emotion contagion in affective inter-personal simulations. *IEEE Transactions on Visualization and Computer Graphics*, 22(4):1326–1335, 2016.

[53] P. Willemsen, M. B. Colton, S. H. Creem-Regehr, and W. B. Thompson. The effects of head-mounted display mechanical properties and field of view on distance judgments in virtual environments. *ACM Transactions on Applied Perception*, 6(2):Art. 8, Feb. 2009.

[54] P. Willemsen and A. A. Gooch. Perceived egocentric distances in real, image-based, and traditional virtual environments. In *Proc. IEEE Virtual Reality (VR)*, pp. 275–276, 2002.

[55] J. K. Witt, J. K. Stefanucci, C. R. Riener, and D. R. Proffitt. Seeing beyond the target: Environmental context affects distance perception. *Perception*, 36(12):1752–1768, 2004.

[56] A. J. Woods, T. Docherty, and R. Koch. Image distortions in stereoscopic video systems. In *Proc. SPIE 1915, Stereoscopic Displays and Applications IV*, pp. 36–48, 1993.

[57] C. Ziemer, J. Plumert, J. Cremer, and J. Kearney. Estimating distance in real and virtual environments: Does order make a difference? *Attention,Perception & Psychophysics*, 71(5):1095–1106 91, 2009.