An Experimental Comparison of Perceived Egocentric Distance in Real, Image-Based, and Traditional Virtual Environments using Direct Walking Tasks

Peter Willemsen *Amy Ashurst Gooch*

UUCS-02-009

School of Computing
University of Utah
Salt Lake City, UT 84112 USA

February 12, 2002

Abstract

In virtual environments, perceived egocentric distances are often underestimated when compared to the same distance judgments in the real world. The research presented in this paper explores two possible causes for this reduced distance perception in virtual environments: (1) real-time computer graphics rendering, and (2) immersive display technology. Our experiment compared egocentric distance judgments in three complex, indoor environments: a real hallway with full-cue conditions; a virtual, stereoscopic, photographic panorama; and a virtual, stereoscopic computer model. Perceived egocentric distance was determined by a directed walking task in which subjects walk blindfolded to the target. Our results show there is a significant difference in distance judgments between real and virtual environments. However, the differences between distance judgments in virtual photographic panorama environments and traditionally rendered virtual environments are small, suggesting that the display device is affecting distance judgments in virtual environments.

An Experimental Comparison of Perceived Egocentric Distance in Real, Image-Based, and Traditional Virtual Environments using Direct Walking Tasks

Peter Willemsen   Amy Ashurst Gooch

School of Computing
University of Utah
50 S. Central Campus Dr., Room 3190
Salt Lake City, UT 84112-9205
www.cs.utah.edu
{willemsn,gooch}@cs.utah.edu

Abstract

In virtual environments, perceived egocentric distances are often underestimated when compared to the same distance judgments in the real world. The research presented in this paper explores two possible causes for this reduced distance perception in virtual environments: (1) real-time computer graphics rendering, and (2) immersive display technology. Our experiment compared egocentric distance judgments in three complex, indoor environments: a real hallway with full-cue conditions; a virtual, stereoscopic, photographic panorama; and a virtual, stereoscopic computer model. Perceived egocentric distance was determined by a directed walking task in which subjects walk blindfolded to the target. Our results show there is a significant difference in distance judgments between real and virtual environments. However, the differences between distance judgments in virtual photographic panorama environments and traditionally rendered virtual environments are small, suggesting that the display device is affecting distance judgments in virtual environments.

CR Categories: I.3.7 [Computing Methodologies]: Computer Graphics—3D Graphics

Keywords: perception, immersive environments, virtual environments, virtual reality, egocentric distance

1 Introduction

The perceived distance to objects in virtual environments when viewed through head-mounted display (HMD) systems have generally been shown to be underestimated. The same distance judgments conducted in the physical world have been shown to be quite accurate for distances up to about 25 meters [Knapp 1999]. The exact reason for this disparity and compression of space in a virtual environment is unknown. In this paper, we report results from experiments designed to explore the cause of this perceived distance compression.

There are two likely causes for the perceived compression of space in virtual environments. Either information is missing from the scene's rendering, or the underestimation of distance stems from the display system. In real-time rendering environments, the graphics pipeline is impoverished; reflective highlights, inter-reflections, and global illumination are missing or inaccurate. In addition, the richness of geometric detail found in the real world cannot yet be fully modeled in real-time computer graphics. Similarly, current display technology is insufficient for reproducing the resolution and field of view of the human vision system. In HMDs, users may not feel immersed in the environment, due to factors such as accommodation to the fixed distance of the screens. The inability to perceive depth correctly in virtual environments is most likely due to a combination of both the display systems and the rendered graphics.

We conducted experiments comparing human depth perception in the real world to performance in virtual environments experienced through a head-mounted display system. The environments consisted of (1) real world hallway environment; (2) virtual, stereoscopic, real-time geometric model of the hallway environment; (3) virtual, stereoscopic, photographic panoramas of the hallway. The real environment provides us with a degree of certainty that the experimental methodology is correct in that the experimental results should be consistent with previous human performance in real environments. For the traditional virtual environment, we created a polygonal, scaled version of the real world hallway environment. The panorama environment was based on photographic images of the hallway which allowed us to explore distance perception in the presence of pictorial cues such as shadows, inter-reflections, global illumination, etc., which are often left out of real-time computer-generated worlds due to computational and rendering constraints. These additional cues were presented to the subjects by capturing stereoscopic panoramas of the same hallway junction, similar in spirit to the image-based virtual environments created by Chen [1995].

If subjects perceive distance more accurately in the photographic panorama environment over the computer-generated geometric environment, our experiment suggests that the underestimation of depth may be a graphics rendering issue. However, if subjects do not perform differently in the photographic panorama environment versus the computer-generated geometric environment, then the perceived distance compression is most likely a result of deficiencies in the display system.

In Section 2 we describe basic research in spatial vision, as well as research specifically addressing differences in spatial vision between real and virtual environments. We describe the structure of the experiment and the equipment used in Section 3. The remainder of the paper details the results of our experiment (Section 4) and conclusions we can draw from this research (Section 5).

2 Related Work

In the past, depth perception research has been conducted mainly by the psychology community. However, it is becoming increasingly important to research in the computer graphics field.

The perception of depth is controlled by several physiological and pictorial cues including accommodation, convergence, stereopsis, familiar size, relative size, eye height, aerial perspective, texture, shading, and motion parallax [Cutting and Vishton 1995; Palmer 1999]. These cues provide both absolute and relative depth information. Our experiment focused on the perception of absolute egocentric distances in virtual environments. Absolute egocentric distance refers to an actual metric distance between an observer and an external target.

Prior spatial vision research has shown that visually directed actions such as blind walking to a previously viewed target are good probes for investigating how physical space maps to perceived visual space [Loomis et al. 1992; Philbeck and Loomis 1997; Philbeck et al. 1997; Rieser et al. 1990]. In such studies, subjects are first allowed to view targets and then asked to walk without vision to the location of the perceived target. Results from these studies, conducted in real world hallways and outdoor spaces, with full cue conditions, show that subjects are accurate at judging distances between 1 and 25 meters.

More recent work has been conducted in virtual environments to compare the results of real world perception with perception in immersive virtual environments. As in the physical world, visually directed actions, such as blind walking, have been used as probes for perceived distance in the virtual world [Knapp 1999; Loomis and Knapp in press; Witmer and Kline 1998]. Results from these and other distance estimate studies [Lampton et al. 1995], with one exception [Witmer and Sadowski 1998], have shown a large underestimation in perceived distance in the virtual environment that varied from about 50% to 70% of the true distances.

Several computer graphics researchers have called for the validation and verification of virtual environments [Stanney 1995; Zeltzer and Pioch 1996] with respect to the physical world, but only a few researchers have tried to evaluate perception in virtual environments. Some researchers have investigated training transfer [Koh et al. 1999], while others have tried to account for how the feeling of immersion is affected by visual depth cues such as accommodation [Ellis and Menges 1997; Neveu et al. 1998; Surdick et al. 1997]. Recent research evaluated the effectiveness of stereopsis, cast shadows, and diffuse inter-reflections in signaling imminent contact in a manipulation task [Hu et al. 2000]. Much of the virtual environment perception investigations have been conducted with impoverished models [Eggleston et al. 1996; Henry 1992; Rademacher et al. 2001] and lower resolution display systems [Henry 1992; Rademacher et al. 2001].

Our research contributes by providing data important to understanding depth perception in complex virtual environments. We explore the causes for the perceived underestimation of distance in complex real-world, non-synthetic environments viewed through head-mounted displays. To our knowledge, no one has published research that has investigated perceived distances in an image-based virtual environment such as the one we tested with our photographic panoramas.

Figure 1: Sample panorama of T-junction hallway used in experiments. The target is located in the right most hallway.

(a) Panoramic (b) Polygonal

Figure 2: Composite of panorama and polygonal environments. The target image is inlaid.

3 Experiment Description

The goal of this experiment is to provide insight into the underestimation of egocentric distance judgments in virtual environments. We probed subjects' perception of egocentric distance using a directed walking task in which subjects walked without vision to a previously viewed target. The target was placed on the ground at distances of 2m, 3.5m, and 5m. These distances were chosen due to the physical limits of our tracked lab space. All subjects were college age, were given a stereogram eye test, and had normal or corrected to normal vision. Our subject pool consisted of 6 females and 6 males. Each subject performed a total of 36 trials in three conditions: (1) real hallway, (2) virtual 360 degree photographic panorama of the real hallway displayed at the highest available resolution in our HMD, and (3) fully polygonal, computer graphics rendering of the same hallway displayed in our HMD. All environmental conditions were presented binocularly. Three training trials were conducted for each condition followed by three trials at each of the three distances. No feedback was provided during any phase of the experiment. The order in which the environment conditions were presented was randomized over the subject pool. Target distance ordering was also randomized.

Prior to each experimental condition, subjects practiced blind walking. Training consisted of the subjects walking blindfolded in a hallway for approximately 5 minutes.

For each trial, subjects were shown a target located on the floor. The target used in all of these experiments was a bright orange frisbee. Subjects were instructed to obtain a good image of the target and their local surroundings. Each subject was told that a "*good image*" is obtained if when the subject closed their eyes, he or she would still be able to "*see*" the environment, and most importantly, the target. Cognitive maps such as these have been extensively explored in the perception literature, as well as some virtual environment studies [Billinghurst and Weghorst 1995]. Subjects were allowed to rotate their head about their neck, but were instructed not to move their head from side to side or back and forth, in order to reduce depth cues resulting from motion parallax. Subjects viewed the current environment until they felt they had obtained a good image of the target and the scene. Once a good image was achieved, subjects either blindfolded themselves or the HMD screen was cleared to black. Subjects then started walking toward the target location. All subjects were directed to walk purposefully and decisively to the target location and stop walking when they felt their feet were where the target had been observed. The same instructions were provided for each environment.

Environment Comparison	F	p
Panoramic Image vs. Real	14.079	0.004
Computer Geometry vs. Real	24.806	0.001
Panoramic vs. Geometry	0.801	0.392

Table 1: ANOVA Analysis between Environment Conditions.

3.1 Equipment and Software

The equipment used in this experiment consisted of an nVision Datavisor HiRes HMD with interlaced 1280x1024 resolution, full color, and a 52 degree diagonal field of view. The angular resolution of the HMD is on the order of 2 arc minutes per pixel. The display was configured with 100% stereo overlap between the two eyes.

Both computer generated environments were rendered on an SGI Onyx2 R12000 with two IR2 rendering pipelines. One rendering pipeline was used for each eye to provide stereopsis. A fixed interpupil distance of 6cm was used for the geometric and panoramic virtual environments. We recorded the interpupil distance for each subject and the average interpupil distance was 6.22cm, with the range being defined by $[5.25cm, 7.0cm]$. However, it is important to note that precise control of effective interpupular distance in HMDs is difficult [Wann et al. 1995]. We lack precise control over image position and accommodation, and these parameters cannot be determined by polling the HMD. The geometrically modeled virtual environment ran at no less than 22 frames per second. The polygon count in this model was approximately 28,000 polygons utilizing 35 texture files. Our software uses OpenGL and Sense8's WorldToolKit libraries to load a VRML model created in Alias|Wavefront's Maya Complete. The panoramic environment was created from photographs taken with a Voigtländer BESSA-L camera body equipped with a 12mm lens providing 112 degrees vertical by 90 degrees horizontal field of view in portrait mode. The panorama environment ran at no less than 40 frames per second.

A series of panoramic images was created, one for each eye, and at a range of eye heights spaced at 5cm intervals. The stereo pair of images nearest to the subject's eye height was used for generating the panorama of the hallway. One of the panoramas is shown in Figure 1.

Panoramas were generated using the Panorama Factory software package. The individual photographs were digitized and stitched together to form a cylindrical panorama. Our virtual environment software mapped left and right eye panoramic images onto separate cylinders for display to the left or right eye of the subject. One side effect of this process is the lack of visual information at the bottom and top of the cylinder. The subjects perceived this artifact as a black hole below and above them. Although subjects were surprised to see holes in the panoramic environments and no representation of their feet in either virtual environment, subjects were able to continue without pause through the experiment. We are uncertain as to how much of an effect this artifact may have had on the subjects and we plan to study this issue further.

During the portion of the experiment conducted in the real world hallway, the subjects' positions and the target locations were recorded with a Leica Reflectorless Total Station. In the portion of the experiment in which subjects wore the HMD, subject positions were recorded using an IS600-Mark2 Intersense tracker. Subjects were afforded head rotation, but translations did not update the rendering the subjects viewed.

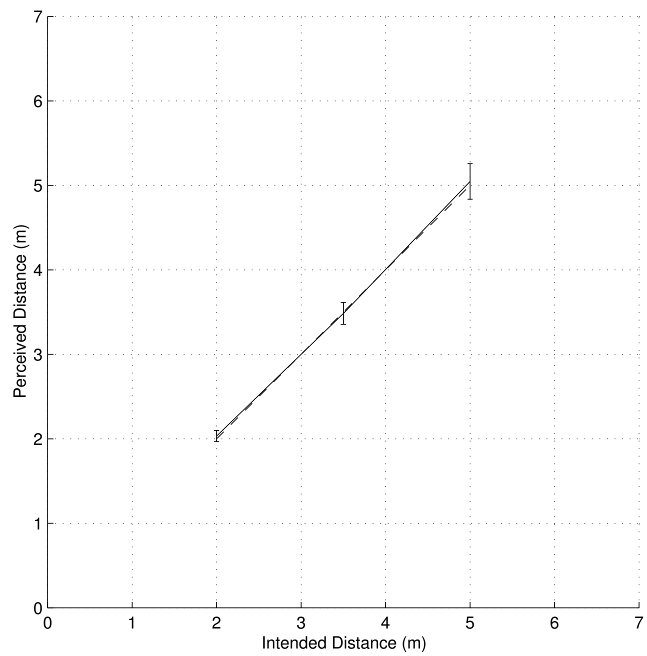

Figure 3: Average intended versus perceived distance in the real hallway. Error bars represent ±1 SEM. The dashed line represents ideal performance.

4 Results

We analyzed the data using a 3 (environment) x 3 (distance) x 2 (sex) repeated measures ANOVA, with environment and distance as within-subjects variables and sex as a between-subjects variable. The results from our experiments indicate that there is a significant difference ($p_{pano} = 0.004$ and $p_{cg} = 0.001$) in performance between the real world condition and both of the virtual environment conditions. The ANOVA summary is provided in Table 1. Figures 3, 4, and 5 show the perceived egocentric distance averages for the real hallway environment, the panoramic virtual environment, and the polygonal virtual environment, respectively. Perceived distance data was averaged across all subjects. In these plots, error bars represent ±1 standard error of the mean (SEM). The plot in Figure 6 provides a direct comparison of the data.

The distance judgments based on actual images (i.e. panoramic environment) were slightly better than those based on the computer geometry, but the differences were small and not statistically significant ($p = 0.392$). This leads us to believe that the quality of graphics does not matter as much as previously thought and that the display is playing a large roll in the compression of distance our subjects perceived. While we did confirm results obtained by previous research that distance perception in virtual environments is foreshortened compared to distance perception in real environments, it is worth noting that the distance judgments made in the computer geometry condition were more accurate than those reported by most other researchers [Knapp 1999; Witmer and Kline 1998; Witmer and Sadowski 1998].

5 Conclusions

Our research has examined two likely causes for the compression of perceived distances in virtual environments. The lack of correspondence between intended and perceived distances is either

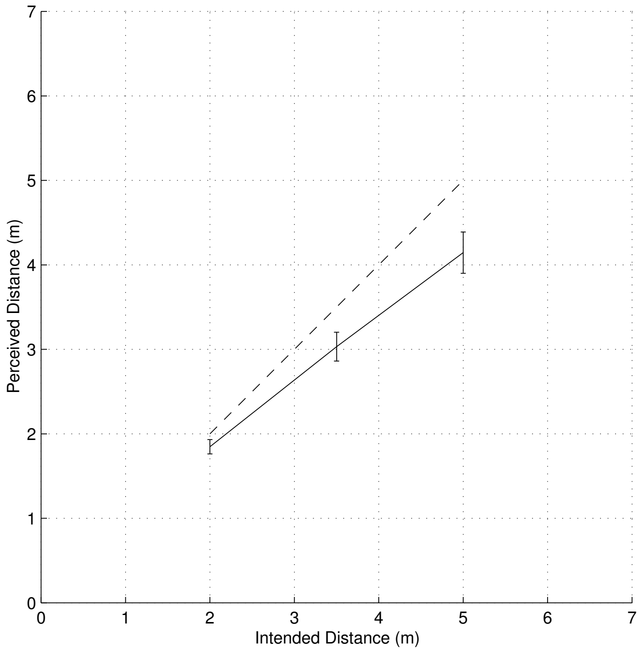

Figure 4: Average intended versus perceived distance in the stereoscopic, panorama environment. Error bars represent ±1 SEM. The dashed line represents ideal performance.

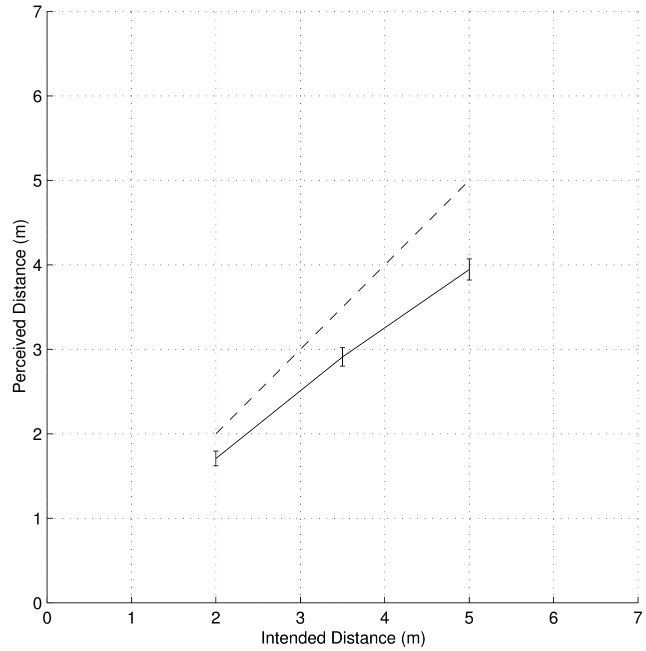

Figure 5: Average intended versus perceived distance in the computer graphics modeled hallway. Error bars represent ±1 SEM. The dashed line represents ideal performance.

due to missing information in the scene's rendering or due to artifacts of the display technology. In order to examine this effect, we conducted trials in a real world hallway, a virtual image-based panorama environment, and a traditional, polygonal, virtual environment.

Our data shows that subjects were quite accurate in the directed walking task in the real world hallway. We have also found a significant difference in the perceived distance in the real world versus both of the virtual environments. These results provide us with confidence that our experimental methodology is correct since our results are consistent with those found by both the spatial vision and computer graphics research communities.

It was expected, prior to the experiment, that the average perceived distance in the computer geometry condition would have been less accurate, resulting in a statistically significant difference in perceived distance between the image-based panorama environment and the computer geometry condition. This increased sense of space may be due, in part, to the geometric complexity found in our model of the hallway, in comparison with simpler geometries used in other research. Additionally, the high resolution of our display system may have contributed to more accurate correlation between intended and perceived distance.

This technical report presents some of the first data to evaluate egocentric distance perception using high-resolution photographs in an immersive display. Because the difference in perception between the computer geometry and photographic panorama conditions was not statistically significant, it suggests that the head-mounted display is one source of perceived spatial compression in virtual environments. While these conclusions are interesting, understanding the causes and magnitude of spatial compression in virtual environments requires substantial additional investigation and should be explored thoroughly.

6 Acknowledgments

The authors are grateful to Andrew Beall for suggesting the use of high resolution panoramic images for studies of this sort. We also thank William Thompson, our subjects, Bruce Gooch, John Regehr, and the members of the University of Utah computer graphics groups. Thanks to Nathan Riddle and Michael K. McCarter from Mining Engineering for helping us measure the target distances. Special acknowledgment goes to Sarah Creem-Regehr, John Rieser, Herb Pick, and Jack Loomis for attempting to teach us about perceptual psychology. Thanks to Alias|Wavefront for their generous donation of Maya Complete, which we use to create the geometric models for our experiments. This material is based upon work supported by the National Science Foundation under Grants CDA-9623614 and IIS-0080999. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the sponsors.

References

BILLINGHURST, M., AND WEGHORST, S. 1995. The use of sketch maps to measure cognitive maps of virtual environments. *Virtual Reality Annual International Symposium*, 40–47.

CHEN, S. E. 1995. Quicktime VR - an image-based approach to virtual environment navigation. In *SIGGRAPH 95 Conference Proceedings*, 29–38.

CUTTING, J. E., AND VISHTON, P. M. 1995. Perceiving layout and knowing distance: The integration, relative potency and contextual use of different information about depth. In *Perception of Space and Motion*, W. Epstein and S. Rogers, Eds. Academic Press, New York, 69–117.

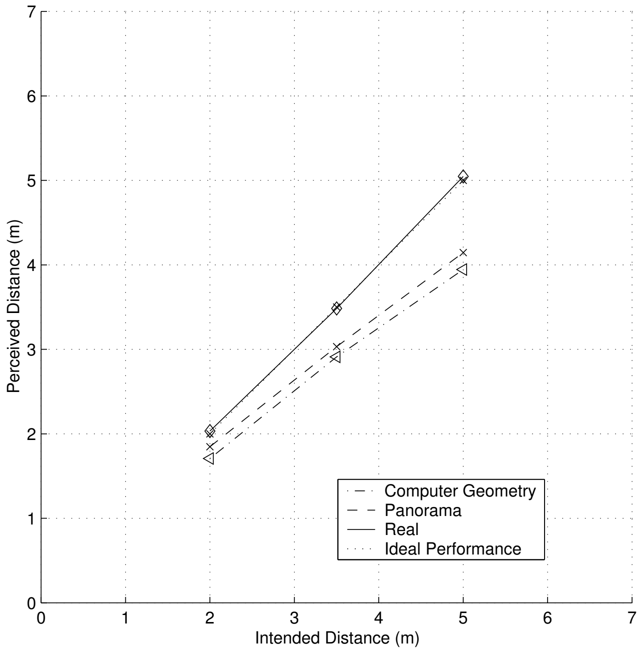

Figure 6: Intended versus perceived distance. Comparison of results from the three environments.

EGGLESTON, R. G., JANSON, W. P., AND ALDRICH, K. A. 1996. Virtual reality system effects on size-distance judgements in a virtual environment. *Proceedings of VRAIS '96*, 139–146.

ELLIS, S. R., AND MENGES, B. M. 1997. Judgments of the distance to nearby virtual objects: Interaction of viewing conditions and accommodative demand. *Presence: Teleoperators and Virtual Environments 6*, 452.

HENRY, D. 1992. *Spatial Perception in Virtual Envionrments: Evaluating an Architectural Application*. Master's thesis, M.S.E. Thesis, University of Washington, College of Engineering.

HU, H. H., GOOCH, A. A., THOMPSON, W. B., SMITS, B. E., SHIRLEY, P., AND RIESER, J. J. 2000. Visual cues for imminent object contact in realistic virtual environments. *IEEE Visualization*.

KNAPP, J. M. 1999. *The visual perception of egocentric distance in virtual environments.* PhD thesis, University of California Santa Barbara.

KOH, G., VON WIEGAND, T. E., GARNETT, R. L., DURLACH, N. I., AND SHINN-CUNNINGHAM, B. 1999. Use of virtual environments for acquiring configurational knowledge about specific real-world spaces: I. preliminary experiment. *PRESENCE 8*, 6 (dec), 632–656.

LAMPTON, D. R., MCDONALD, D. P., SINGER, M., AND BLISS, J. P. 1995. Distance estimation in virtual environments. In *Human Factor and Ergonomics Society*, 1268–1272.

LOOMIS, J. M., AND KNAPP, J. M. in press. Visual perception of egocentric distance in real and virtual environments. In *Virtual and Adaptive Environments*, L. J. Hettinger and M. W. Haas, Eds. Erlbaum, Hillsdale, NJ.

LOOMIS, J. M., DA SILVA, J. A., FUJITA, N., AND FUKUSIMA, S. S. 1992. Visual space perception and visually directed action. *Journal of Experimental Psychology: Human Perception and Performance 18*, 906–921.

NEVEU, C., BLACKMON, T., AND STARK, L. 1998. Evaluation of the effects of a head-mounted display on ocular accommodation. *PRESENCE 7*, 3, 278–289.

PALMER, S. E. 1999. *Vision Science: Photons to Phenomenology*. The MIT Press.

PHILBECK, J. W., AND LOOMIS, J. M. 1997. Comparison of two indicators of perceived egocentric distance under full-cue and reduced-cue conditions. *Journal of Experimental Psychology: Human Perception and Performance 23*, 1, 72–85.

PHILBECK, J. W., LOOMIS, J. M., AND BEALL, A. C. 1997. Visually perceived location is an invariant in the control of action. *Perception and Psychophysics 59*, 4, 601–612.

RADEMACHER, P., LENGYEL, J., CUTRELL, E., AND WHITTED, T. 2001. Measuring the perception of visual realism in images. *Proceedings of the 12th Eurographics Workshop on Rendering*.

RIESER, J. J., ASHMEAD, D. H., TAYLOR, C. R., AND YOUNGQUIST, G. A. 1990. Visual perception and the guidance of locomotion without vision to previously seen targets. *Perception 19*, 675–689.

STANNEY, K. 1995. Realizing the full potential of virtual reality: Human factors issues that could stand in the way. *IEEE VR 1995*, 28–34.

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

WANN, J. P., RUSHTON, S., AND MON-WILLIAMS, M. 1995. Natural problems for stereoscopic depth perception in virtual environments. *Vision Research 35*, 19, 2731–2736.

WITMER, B. G., AND KLINE, P. B. 1998. Judging perceived and traversed distance in virtual environments. *Presence: Teleoperators and Virtual Environments 7*, 2 (April), 144–167.

WITMER, B. G., AND SADOWSKI, JR., W. J. 1998. Nonvisually guided locomotion to a previously viewed target in real and virtual environments. *Human Factors 40*, 478–488.

ZELTZER, D., AND PIOCH, N. J. 1996. Validation and verification of virtual environment training systems. *Proceedings of VRAIS '96*, 123–130.