

Visualization and (Mis)Perceptions in Virtual Reality

J. Campos, H-G. Nusseck, C. Wallraven, B.J. Mohler, H.H. Bühlhoff

Max Planck Institute for Biological Cybernetics,
Spemannstrasse 38, 72076 Tübingen
E-mail: heinrich.buelthoff@tuebingen.mpg.de

Abstract: Virtual Reality (VR) technologies are now being widely adopted for use in areas as diverse as surgical and military training, architectural design, driving and flight simulation, psychotherapy, and gaming/entertainment. A large range of visual displays (from desktop monitors and head-mounted displays (HMDs) to large projection systems) are all currently being employed where each display technology offers unique advantages as well as disadvantages. In addition to technical considerations involved in choosing a VR interface, it is also critical to consider *perceptual* and *psychophysical* factors concerned with visual displays. It is now widely recognized that perceptual judgments of particular spatial properties are different in VR than in the real world. In this paper, we will provide a brief overview of what is currently known about the kinds of perceptual errors that can be observed in virtual environments (VEs). Subsequently we will outline the advantages and disadvantages of particular visual displays by focusing on the *perceptual* and *behavioral* constraints that are relevant for each. Overall, the main objective of this paper is to highlight the importance of understanding perceptual issues when evaluating different types of visual simulation in VEs.

1 Introduction

VR technologies are currently used in many diverse application areas including training, simulation, visualization, product design, therapy, tele-operations, as well as gaming. Although more recent VR technology also retains the potential to produce controlled multisensory experiences, most of the technologies developed for this field of research have focused primarily on the simulation of visual input. As will be discussed in greater detail below, visual displays ranging from simple desktop monitors, multi-panel tiled displays, immersive head-mounted displays (HMDs), and large, curved, projection screens are all currently being employed. Each of these displays offers unique advantages as well as disadvantages.

In addition to technical and logistical considerations involved in choosing a VR interface (such as issues related to space restrictions and cost), we feel, however, that it is also critical to consider *perceptual* and *psychophysical* factors concerned with those visual displays. A common finding in this context, for example, is that perceptual judgments of particular spatial properties are not veridical in VR. For instance, in-depth distance intervals have consistently been shown to be perceptually compressed in VEs (Witmer et al. 1998). Specifically, when real-world distances are accurately rendered in a VE, they will appear much shorter. There may also be differences in perceptual-motor coupling in VR as evidenced by differences in perceived speed in real and virtual environments (Banton et al. 2005). Understanding the cause of such misperceptions is essential when designing systems for particular applications and when interpreting the behavioral data of interest. Specifically, understanding the cues that are important for real-world absolute distance perception (e.g. field of view, angular declination from the horizon, and eye height) will provide insight into the cues that are necessary for veridical distance perception in Virtual Environments (VEs).

Here we will begin by describing what is currently known about the perceptual errors observed in VEs; in particular the above-mentioned compression of absolute distance perception and the underestimation of visual speed. Subsequently we will outline the advantages and disadvantages of particular visual displays by focusing on the perceptual and behavioral constraints that are relevant for each. Overall, the main objective is to highlight the importance of understanding perceptual issues when evaluating different types of visual simulation in VEs. Such issues become highly relevant when they impair an observers capacity to effectively perceive spatial features and to behave and interact accurately within that space. This is not only important for directly interpreting behaviors that are produced within the VE, but also when considering the consequences of transferring such knowledge to the real world (whether this is a desired or undesired result).

2 Non-Vertical Perception in Virtual Environments

2.1 Distance Compression

While much research has now demonstrated that humans are very good at estimating the distance between themselves and a stationary target in the real world, the same distances are consistently underestimated in immersive VEs by as much 50% (Witmer and Kline, 1998; Knapp and Loomis; 2004; Thompson, et al., 2004). Considering the importance of accurately representing distances in simulated spaces, it is important to understand the potential causes of such compressions. There are several aspects of visual inputs that are important for absolute distance perception and that may be differentially represented in VR compared to the real world. Such factors include, for example, the quality and/or fidelity of the visual graphics, restrictions to the field of view (FOV), and potentially the absence of near-space scaling information¹.

One hypothesis that was tested addressed whether absolute distances are compressed in VE's due to the absence of particular visual cues that enhance the quality of the rendered scene such as realistic textures and shadows (Thompson et al., 2004). Thompson et al. (2004) tested this hypothesis by comparing participants' target distance judgments across three versions of the same VE (an indoor hallway) of varying visual qualities. Results demonstrated that there were no differences in participants' distance estimates across the three environments, indicating that the quality of the visual graphics could not account for the distance compression effect.

A second hypothesis relates to the reduced FOV that often accompanies immersive visual displays (see also Figure 7 below). While some investigations have demonstrated a distance compression effect when the FOV is reduced and the point of view is stationary (Witmer and Kline, 1996), others have shown in the real world that when head movements are allowed under restricted FOV conditions, these effects disappear (Knapp and Loomis; 2004; Creem-Regehr et al., 2005). For instance, Knapp and Loomis (2004) demonstrated that when reducing the FOV in the real-world (through a viewing aperture) so that it approximated the FOV reductions when wearing HMD's, this did not lead to a compression of perceived distance. The FOVs reported above (Knapp and Loomis: 47° horizontal x 43° vertical; Creem-Regehr et al., 2005: 42° horizontal x 32° vertical) are still quite large, and it is known that when the horizontal FOV is systematically reduced to smaller FOVs, distances become increasingly compressed with greater reductions (Wu et al., 2004: <21° horizontal x 21° vertical).

Significant insights into the visual information necessary for distance perception comes from a series of work proposing a "sequential surface integration process" (Wu et al., 2004; He et al., 2004; Wu et al., 2007). Based on this theory, the near ground information is used to appropriately scale the far ground information so that when an individual scans the ground from

¹ Note that it is most likely the case that the distance compression effect observed in VEs could be a combined effect of several different factors.

the space around their feet to targets within action space (< 20m), they sequentially integrate distance information. Support for this hypothesis comes from the fact that if you disrupt the ground between yourself and a target (i.e., by a hole in the ground, by an obstacle, or by a change in texture from grass to concrete), distances are systematically underestimated (He et al. 2004). Further, if you reduce the vertical FOV to eliminate information about the ground in one's immediately surrounding personal space, this also causes distance underestimations, whereas if you reduce the horizontal FOV so that only the periphery is masked while the near ground information is maintained, no distance underestimation is observed (Wu et al., 2004). Considering that many important visual cues to absolute distance (i.e. binocular disparity, convergence, accommodation, etc.) can only be used at near distances, the accurately scaled information within this area can be used to draw inferences about farther distances in which visual cues to absolute distance become sparser. Further, it has also been shown that, with a known eye-height individuals can use the angular declination from the horizon to calculate absolute distances (Ooi et al., 2001). Overall, this elegant and comprehensive series of studies provides support for the importance of vertical FOV, rich texture information and accurate perception of eye-height when simulating visual environments.

2.2 Speed Underestimation in VR

Another, less studied perceptual difference between virtual and real environments that has also been reported, is the misperception of visual speed in VEs. Banton et al. (2005) required participants to match their visual speed (presented via an HMD) to their walking speed as they walked on a treadmill. When facing forwards during walking, visual speeds were increased by about 1.6x that of the walking speed in order to appear equal. When facing the walls or the floor, this underestimation of visual speed was no longer apparent, thus suggesting that the optic flow provided by peripheral/lamellar optic flow is important for speed perception. Considering that the magnitude of optic flow is greatest peripherally and lowest at the centre of expansion/centrally, having access to the faster inputs may improve performance.

As mentioned previously, perceived eye-height affects distance perception, but as a function of this, it also impacts speed perception. Specifically, as eye-height increases (i.e. farther removed from the ground), the magnitude of optic flow decreases (Gibson, 1979). To explicitly test the implications of this effect in an applied task, Rudin-Brown (2006) evaluated speed perception in a simulated driving task where drivers were positioned at varying eye-heights. Two eye-heights were chosen to represent the height experienced in a tall vehicle such as an SUV and the height experience in a low-slung vehicle such as a sports car. The results of this task demonstrated that speed perceptions were reported as faster when evaluated from a low eye-height (reflected in slower driving speeds) and slower when evaluated from a high eye-height (reflected in faster driving speeds). The variability in speed maintenance was also greater for high eye-heights and lower for low eye-heights. As such, the impact of eye-height on perceived speed and distance has important implications when developing VE's.

2.3 Summary

While this is a short summary of some of the many potential perceptual issues that may become relevant when developing simulated environments, much continues to remain unknown about the causes of such effects and how they might be overcome. It is however, important to consider many of these factors when designing and implementing scenarios and applications using VR. In recent years VR technologies have become more affordable, accessible and diverse. As such, the availability of different types of systems has increased. Therefore, evaluating the various options for one's own purposes and intentions becomes important. Again, there are practical considerations such as cost and space requirements, but there are also perceptual considerations, such as those discussed above. When considering what type of visual display is ideal for one's end goal, it is important to carefully evaluate the advantages and limitations of each.

3 Advantages and Limitations of Different Visual Displays

3.1 Desktop Displays

Traditional desktop displays consist of stationary computer monitors in which an external control device (i.e. a joystick or mouse) is used to interact within the VE. In recent years the quality and resolution of desktop displays has been steadily increasing. Current technologies are now providing dramatically superior visual displays than in the recent past. This trend encompasses both the available resolution as well as the contrast levels that displays can support. A recent development, for example, is the Brightside High-Dynamic-Range display that supports luminance up to 100,000 cd/m² and a contrast ratio of 200,000:1, which is much closer to the contrast range the human eye can actually perceive. Also the development of high resolution flat panels reaches the domain of the optical resolution of the human eye. The eyevis 56" DHD display for example, comes with a resolution of 3840x2160 pixels.



Figure 1. Brightside DR37-P HDR display. This display has a resolution of 1920x1080 pixels and a contrast ratio of 200,000:1. It is able to simultaneously show both dark and very bright pixels. www.brightsidetech.com



Figure 2. EYE-LCD5600DHD. This display has a resolution of 3840x2160 pixels and a panel size of 1224x729 mm (56" diagonal). www.eyevis.de

The trade-offs associated with high resolution displays are that they tend to be limited in size (reduced FOV, see Figure 7 for a comparison of the typical FOV of a desktop display in comparison to the human FOV) and do not provide peripheral visual information, which is known to be important for many perceptual and spatial tasks. Further, desktop displays are not immersive, but instead are viewed within the context of the surrounding visual environment. The stability of the surrounding environment can essentially “ground” the observer as they interact or navigate within a VE and as such may alter the perceived realism or reduce the impression of perceived self-motion. Further, no natural movement (other than minimal head movements) can be accommodated when viewing a desktop display. That is, all exploratory or interactive behaviors must be achieved through an artificial control device which may also impact the ecological validity of certain tasks.

3.2 Multi-Panel Tiled Displays

Multi-panel tiled displays consist of several, high resolution monitors arranged together to collectively form a much larger visual display. Such set-ups maintain the advantages of high resolution while also maintaining the capability of projecting to a much wider area, thus increasing the FOV. Further, such displays are also flexible in how they are arranged. They can be mounted along a flat surface, or can be curved inward at the sides, thus providing peripheral information.

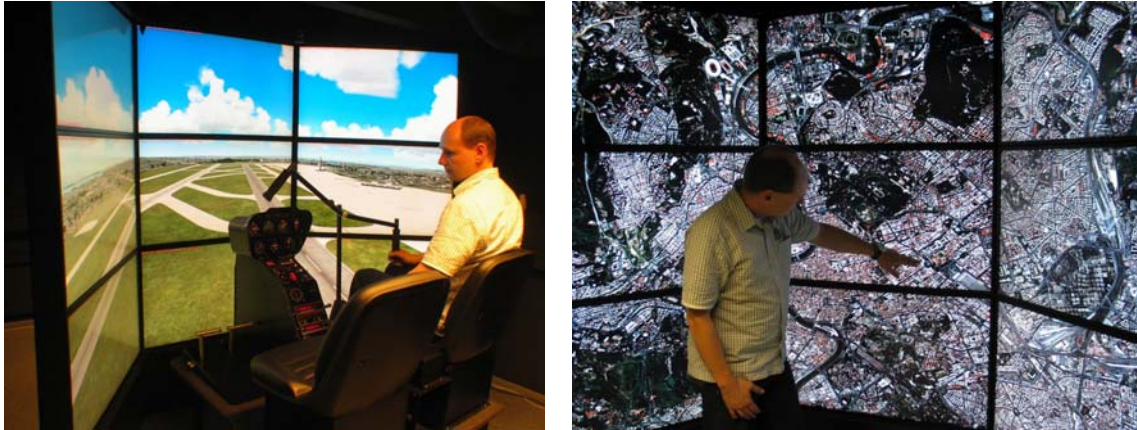


Figure 3. Multipanel tiled display installed at the Max Planck Institute for Biological Cybernetics. Each panel has a resolution of 1366x768 pixels and a panel size of 1018x573 mm (46" diagonal).

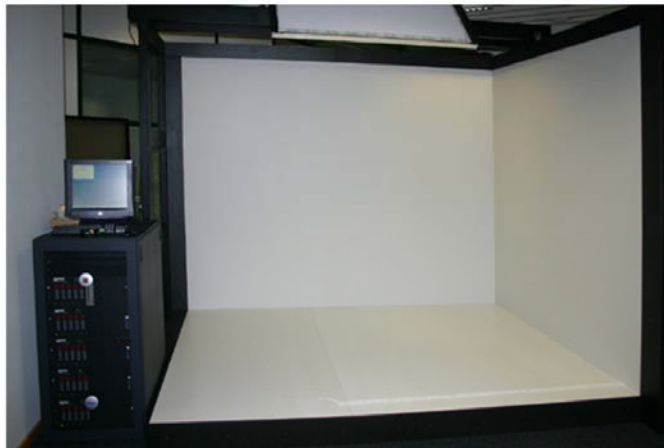
This configuration also has the advantage of not causing geometric distortions that can occur when projecting onto curved surfaces. One limitation of multi-panel tiled displays is the seams that remain visible at the intersections of the monitors. Such seams disrupt the fluidity of the visual scene and can be used to disambiguate perceived self-motion from object-motion. Specifically, the stability of the grid frame will cause a conflict with the movements represented by the visual display. If the seams could somehow be incorporated into the simulation (i.e. serve as the frame of the car in a driving simulator or remain masked by the cockpit of a flight simulator), their detrimental effects would be greatly reduced.



Figure 4. Multipanel tiled display based on 67" SXGAplus back-projection cubes (eyevis ImmersiveCUBE system).

3.3 Panoramic Projection Displays

Some of the first panoramic projection displays include the Cave Automatic Virtual Environments (CAVE™)² in which the four walls (and sometimes floor) of a small, square room are back-projected with an image of the visual scene (see Figure 5 for an example of an “open” CAVE™). Such displays are often projected with two-slightly different images (accounting for the inter-pupillary distance), which, when paired with stereo glasses can provide a 3-dimensional display of the environment. In addition to the full FOV and high level of immersion provided by these set-ups, they also provide the ability for active movements within the virtual space. The extent of these movements however, is restricted by the confines of the space (apart from other, non-natural forms of movement through the space such as via a joystick or walking in place). The other advantage of this system is that you maintain a visual representation of your own body within the virtual environment. This not only eliminates the conflict of not having your body represented in the virtual environment, but also provides scale information that may allow you to better interpret information at farther distances in the visual scene. One of the limitations of the CAVE™ is that, similar to the tiled displays, the points at which the walls intersect form an unnatural juncture in the visual scene.



Source: www.viscon.de (with permission)

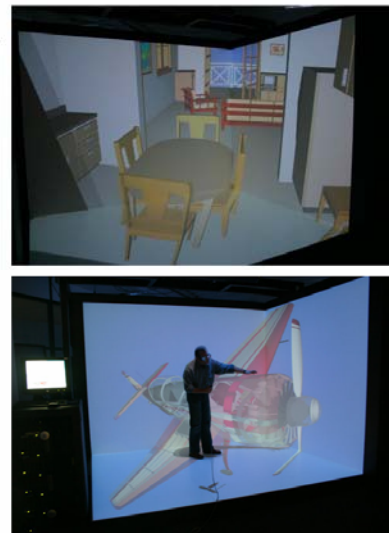


Figure 5. 3-sided CAVE with and without projection.

More sophisticated types of panoramic projection displays have more recently been developed which consist of large dome or curved projection screens. These curved projections wrap around the observer to provide an image that encompasses the entire visual field (see Figure 6 for an example, also see Figure 7 for a comparison of the typical FOV of two different curved screen types in comparison to the human FOV). The projection screen can be offset from the observer's own position (i.e. in front of the observer by approximately 3m) so that the distance information in near space is maintained. Based on the previously described sequential surface integration process, the near distance information is critical when evaluating more distal visual information. The limitation of these displays is the geometric distortions that are caused when projecting onto a curved surface.

² CAVE is a trademark of the university of Illinois

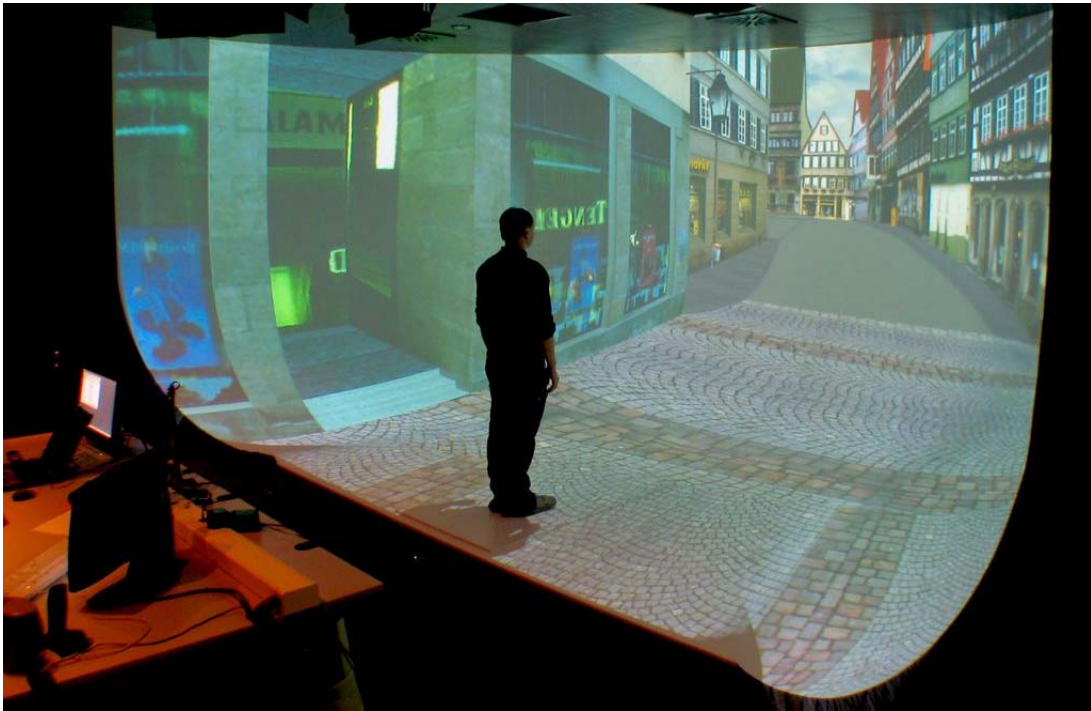


Figure 6. Spherical screen installed at the Max Planck Institute for Biological Cybernetics. It consists of four projectors which project onto a semi-spherical screen. Image warping for geometry correction as well as edge blending is done by openWARP® technology. www.openwarp.com

The main advantage of large field dome or cylindrically shaped projection screens is the decrease in conflicts between the distance of the physical screen and the scene that is projected onto it. Compared to CAVE-like systems, which also covers a huge area of the human FOV, no screen sections are visible. This results in a subjectively more immersive experience.

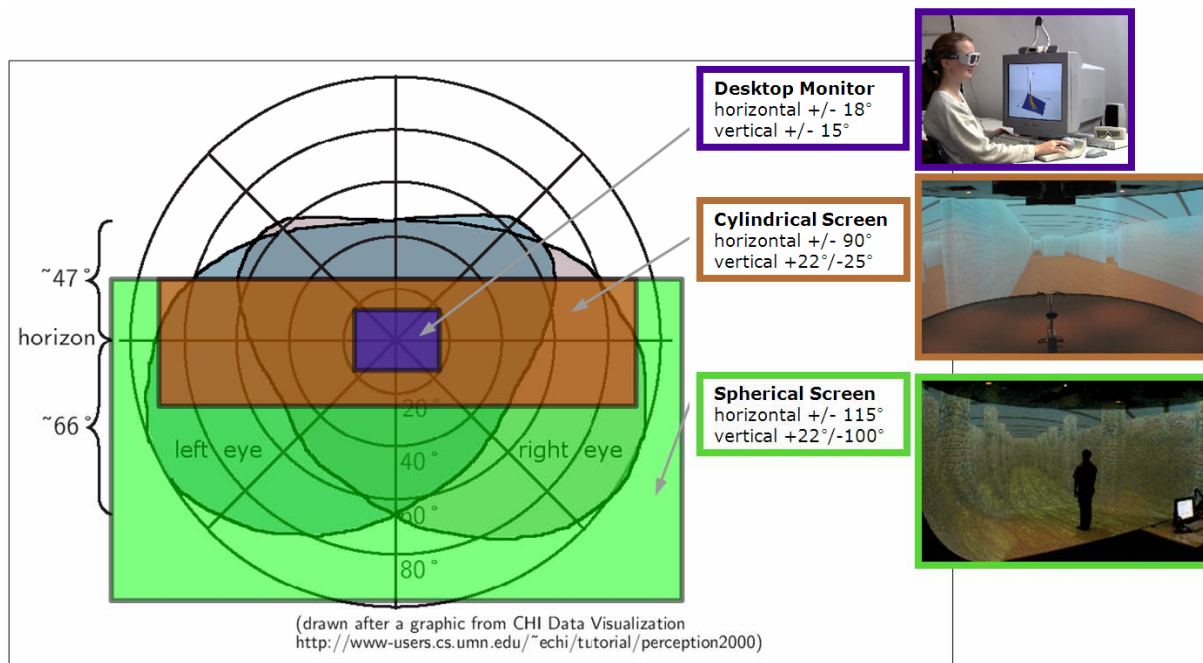


Figure 7. FOV comparison. The spherical plot shows the extent of the human FOV for both the right eye and the left eye. The overlaid boxes show the FOV for several display systems. As can be seen, the spherical projection screen covers a much larger extent of the human field of view.

3.4 Head-Mounted Displays

Apart from desktop displays, head-mounted displays (HMDs) are perhaps the most widely used visualization system. HMD's consist of two small displays that can be worn in a way that presents a visual image to each eye. The two images can be the same (bi-ocular) or can present two slightly different images of the same visual scene resulting in a stereoscopic display. HMDs range widely in size, resolution and FOV. The average FOV of HMDs fall around 40°-60° diagonal with FOVs as low as 24° and greater than 100°. The typically small FOV is one of the main limitations of using HMD's, especially considering that the vertical FOV is particularly reduced. This restriction can be partially ameliorated however by pairing the HMD with a motion tracking system which can update the visual image as a function of the observer's own head movements. This allows for a greater sampling of the environmental space and for the capacity to scan near distances that may not be accessible when facing forwards. Particularly with head-tracking, HMDs also provide a highly immersive experience as the visual information is typically restricted completely to that experienced through the display. Perhaps the greatest advantage of HMDs is the extent of mobility that is possible, thus allowing for natural, large-scale movements through space (see Figure 8). One of the limitations of wearing the visual display (and all associated hardware) is that the system can be heavy and/or cumbersome. In fact, it has been shown in the past that the weight of the HMD may actually account for some of the distance compression effects observed in immersive VEs (Willemson, et al., 2004).



Figure 8. User wearing an HMD and a tracking helmet for navigation in large VEs (setup at the Cyberneum (www.cyberneum.de) of the Max Planck Institute for Biological Cybernetics).

3.5 Summary

Overall it is clear that effectively and accurately representing virtual simulations of visual experiences is very important to the goals for both applied fields and in basic research. Understanding the characteristics of visual displays that achieve these goals is essential in the development and interpretation of the work conducted using VEs. While we have focused on the visual aspects of the virtual experience, it is important to note that several of these visual displays are now being paired with additional devices that provide a more multi-sensory experience (e.g. haptic devices, auditory information). Further, with full-body motion tracking, it is now possible to represent avatars of both the observer and others in the virtual environment, thus allowing for visual feedback of self-movements as well as multi-user interactions.

4 References

- Banton, T., Stefanucci, J., Durgin, F., Fass, A., Proffitt, D. The perception of walking speed in a virtual environment. *Presence-Teleoperators and Virtual Environments*, 14(4), 394-406, 2005.
- Creem-Regehr, S. H., Willemsen, P., Gooch, A. A., And Thompson, W. B. The influence of restricted viewing conditions on egocentric distance perception: Implications for real and virtual environments. *Perception* 34(2), 191–204, 2005.
- Gibson, J. J. *The ecological approach to visual perception*. Boston: Houghton Mifflin, 1979.
- He, Z., Wu, B., Ooi, T.-L., Yarbrough, G., and Wu, J. Judging egocentric distance on the ground: Occlusion and surface integration. *Perception* 33, 7, 789—806, 2004.
- Knapp, J. M., And Loomis, J. M. Limited field of view of head-mounted displays is not the cause of distance underestimation in virtual environments. *Presence: Teleoperators and Virtual Environments* 13(5), 572–577, 2004.
- Mohler, B. J. The effect of feedback within a virtual environment on human distance perception and adaptation Doctoral Dissertation, University of Utah, January, 2007.
- Ooi, T. L., Wu, B., And He, Z. J. Distance determination by the angular declination below the horizon. *Nature* 414, 197–200, 2001.
- Rudin-Brown C. M. The effect of driver eye height on speed choice, lane-keeping, and car-following behavior: results of two driving simulator studies. *Traffic Inj Prev*, 7(4), 365-72, 2006.
- Thompson, W. B., Willemsen, P., Gooch, A. A., Creem-Regehr, S. H., Loomis, J. M., And Beall, A. C. Does the quality of the computer graphics matter when judging distances in visually immersive environments? *Presence: Teleoperators and Virtual Environments* 13(5), 560–571, 2004.
- Thompson, W. B. “Visual perception” In *Fundamentals of Computer Graphics, Second Ed.* A. K. Peters, Ltd., Natick, MA, USA, 2005.
- Witmer B.G., Kline P.B. Judging perceived and traversed distance in virtual environments. *Presence: Teleoperators and Virtual Environments* 7, 144–167, 1998.
- Witmer B.G., Sadowski, W.J. Nonvisually Guided Locomotion to a previously viewed target in real and virtual environments. *Human Factors* (40)3, 478–488, 1998.
- Wu, B., Ooi, T.-L., and He, Z. Perceiving distance accurately by a directional process of integrating ground information. *Nature* 428, 7377, 2007.