Quantifying Spatial Presence

Cedar Riener and Dennis Proffitt
Department of Psychology,
University of Virginia

Keywords: spatial presence, illusions, visual perception

Summary

The human visual system uses a series of complex assumptions, defaults, and biases to reconstruct the depth relationships in the three-dimensional world from the two-dimensional image on the retina. Some of these biases are independent of the image on the retina itself, and dependent on whether the observer is viewing a picture of the real world or the real world itself. For investigators interested in the sense of presence, this characteristic of the visual system presents the possibility to quantitatively measure the spatial presence evoked by a virtual environment by determining the degree to which these biases are evident. Two case studies are presented, one in which a virtual environment resembled the real world, and one in which it resembled pictures. The evaluation metrics discussed are particularly important because the benchmark they provide is not necessarily how well the virtual environment replicates the real world, but whether it seems real to the human perceptual system.

1 Introduction

With the advent of new technologies for creating and displaying virtual environments, a research need has arisen for the evaluation and comparison of these environments. Guided by subjective accounts of the compelling nature of certain virtual environments, many investigators have focused on “the sense of being there” or a sense of presence as a way to evaluate virtual environments. Upon closer inspection, however, the sense of presence has appeared to be a complex construct, and not measurable by a single test. Some investigators have suggested that sub-factors such as “involvement,” “realness” and “spatial presence” ((Schubert, Friedman, & Regenbrecht, 1999) may contribute in non-additive ways to the total sense of presence. Presence in most virtual environments depends critically on using two dimensional visual display surfaces to convincingly evoke a three-dimensional world. An environment that contains depth relationships that mirror that of the real world has been called “spatial presence,” in recent presence research.

In seeking to provide a compelling illusion of reality, many creators of virtual environments have assumed that the problems of creating virtual environments are analogous to the problem of human visual perception. While a virtual environment uses two dimensional LCD-screens or projections, the human visual system uses an inverted two-dimensional projection of the world on the back of the retina. This analogy compares the visual system to a camera (or a video camera), and thus ignores certain fundamental characteristics of the human perceptual system. Several layers of complex processing take place after an image is projected on the retina, and the resulting representation (if it can be called a representation at all) is not a direct replication of reality, but a human reconstruction of reality, filled with systematic biases and distortions. This reconstruction is not based on the retinal projection alone, but also makes critical assumptions about the state of the world and the observer’s place in that world.

These biases and distortions are often seen as adaptive assumptions that the visual system makes in order to efficiently reconstruct a “best guess” of the spatial relationships of the three-dimensional world. Because these “errors” can be reflective of depth and size relationships that our visual system believes to exist in the world (independent of the particular retinal projection), the magnitude of certain biases and distortions varies as the retinal projection remains constant. For example, certain size scaling errors can be dependent on whether the observer is viewing the real world, or simply viewing photographs of the world, although the visual angles subtended by the stimuli are identical.

Virtual environments can be evaluated using these distortions as metrics. The magnitude of distortion observed in the virtual environment can be compared to the documented levels observed for viewers of the real world, and viewers of photographs. Such a test is potentially valuable because it is independent of subjective, conscious measures of presence, yet it provides a continuous scale of spatial presence anchored by two endpoints of comparison (photographs and the real world).

In the following paper, we will present two case studies of research investigating this litmus tests for quantifying spatial presence. The first set of studies identified differences in the magnitude of an illusion known as the vertical-horizontal illusion between photographs and the real world, then evaluated a virtual reality head-mounted display (HMD) using scenes that evoked this illusion. The second set of studies also identified differences between photographs and the real world, but evaluated a large panoramic projected environment, coupled with three-dimensional surround sound.

2 Case Study #1: The Vertical-Horizontal Illusion

The prototypical illustration of the vertical-horizontal illusion is pictured in Figure 1. Although the lines are of the same extent (AB=AC), the vertical line segment appears to be slightly longer. This illusion is relatively small in magnitude (around 3 to 6% vertical overestimation), but highly reliable, in the pictures and line drawings in which it has traditionally been studied. The few studies that have investigated this illusion in real-world settings have found much larger effects, from 20 to 40% (Chapanis & Mankin, 1967; Higashiyama

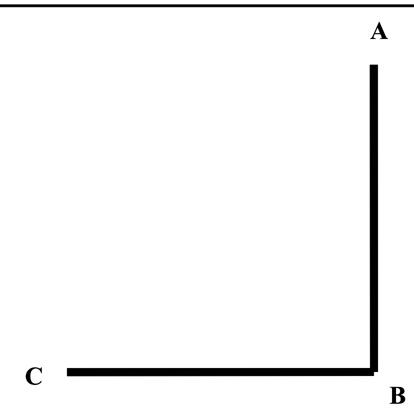

Figure 1

& Ueyama, 1988). Yang, Dixon, and Proffitt (1999) conducted a series of studies to determine how and why this discrepancy occurs. They administered the traditional geometric line-drawing version of the vertical-horizontal illusion on a desktop computer to establish a baseline. Another group of participants viewed an outdoor scene, and were asked to match vertical and horizontal extents. A third group viewed photographs of the same outdoor scene and were given the same task (matching vertical to horizontal extents). The visual angles subtended by the lines in the vertical and horizontal extents in all three conditions were held constant. The groups who viewed the photographs and line drawings overestimated approximately 3%, and the group that viewed the real outdoor scene overestimated the vertical extent by about 12%. To ensure that the increased overestimation of the outdoor scene was not simply a result of the increased field of view, another group of participants viewed the outdoor scene through a window that cropped their view in the same degree that the desktop monitor constrained the photographs. This manipulation did not affect the level of vertical overestimation of those viewing the outdoor scene.

Utilizing the data from the above studies as a comparison, Yang, Dixon and Proffitt (1999) then presented the same outdoor scene in a head-mounted display. Again, visual angles were held constant. The magnitude of overestimation of these participants matched those who viewed the real scene. One final experiment was conducted to demonstrate that the comparison between the real scene and the HMD was not a result of the optical characteristics of the HMD, but of the interpretation of the VR world by the participants. Participants in this "virtual desktop" condition viewed the outdoor scene through the HMD as before, but this time were led into a virtual office, and viewed the scene as if it were presented on a virtual desktop computer. The scene was exactly the same one presented to participants in the earlier HMD condition, but these participants were led to believe that they were looking at a desktop. Participants in this condition overestimated to the same degree as those who viewed pictures on a real desktop terminal.

The above experiments using the vertical-horizontal illusion provide several key insights into the nature and measurement of spatial presence. Although the image in virtual reality may subjectively seem to be less realistic (low field of view, low resolution, low color accuracy), the HMD is effective in conveying a sense of depth, as measured by the magnitude of the vertical-horizontal illusion. This suggests that while high-quality photographs or films may provide a higher resolution of detail, for applications such as an architectural walkthrough, an HMD will actually convey a more realistic sense of the depth and size relationships in a virtual environment. The importance of the assumptions and processing of the visual system (independent of the image on the retina) is emphasized by reiterating that in each condition, the visual angle of the lines judged was held constant, and in the final two conditions (the virtual outdoor scene and the virtual desktop condition) the exact same images were displayed to the observer.

These experiments using the vertical-horizontal illusion accomplish several goals. First, they demonstrate the viability of an objective litmus test based on perceptual biases, not on the geometric properties of retinal projection. Second, a particular virtual environment technology, virtual reality presented in a head-mounted display, was found to pass this litmus test, and have the potential to resemble reality rather than pictures.

3 Case Study #2: The Ponzo Illusion

A second possible application of this litmus test for the spatial presence of virtual environments is the Ponzo illusion, also known as the railroad tracks illusion. Depicted in Figure 2, the Ponzo illusion refers to the overestimation of line A in comparison to line B (the lines are actually the same size) when viewed in the context of lines slanted towards line A. One possible explanation, advanced by Gregory (1965), is that the slanted lines are taken as parallel lines converging at the horizon, as they are in the case of railroad tracks, or a

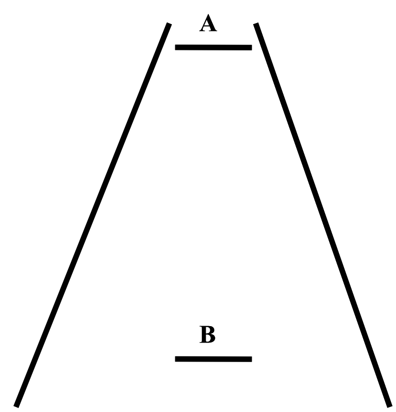

Figure 2

road. The illusion is thus a result of the induced depth relationship that line B is closer to the viewer than line A, and a misapplied size constancy scaling (objects that are further away, but project the same visual angle as closer objects, are larger). Following this line of reasoning, one might suspect that the more depth cues in a scene, the larger the magnitude of the illusion.

Liebowitz, Brislin, Perlmutter and Hennessy (1969) investigated this possibility in a study comparing line drawings, different types of photographs, and real world viewing (both monocular and binocular). There were 6 conditions (control, geometric figure, photograph texture, photograph perspective, actual scene monocular and actual scene binocular). They found the magnitude of the illusion in line drawings to be 10%, standard for geometric line instances of this illusion. There were two types of photographs, those with a texture (a field of mown grass), and those with strong perspective cues (railroad tracks). The average overestimation of the bottom line for the texture photographs was around 20%, while this rose to near 30% for the perspective photograph. Those viewing the actual scene with only one eye did not overestimate more than the perspective photograph, but the binocular viewing of the real scene resulted in an average overestimation of 45%.

Seeking to use the Ponzo illusion to evaluate spatial presence in a particular virtual environment, we presented instances of the Ponzo illusion on a large panoramic projected display screen and an ambient three dimensional surround sound environment. Informal observations strongly suggest that the use of high-fidelity three-dimensional sound environments has potential to greatly enhance the sense of presence. Using Liebowitz et al's (1969) levels of overestimation as metrics for spatial presence, we sought to determine if the large size and field of view of the panoramic display evoked spatial presence, or if spatial presence was enhanced by an ambient surround sound environment, which matched the outdoor scene viewed.

The participants viewed an outdoor scene, centered on two sets of railroad tracks, projected onto a large curved projection screen. The image subtended approximately 110 degrees of visual angle, and was projected using 3 Sharp Notevision 6 Projectors (2200 lumens). The physical dimensions of the projected image were approximately 1.2 meters high by 5 meters wide. The subjects stood behind a small pedestal and viewed the scene. On the pedestal (1 meter high) was placed the trackball for adjustment of the lines. Participants adjusted the lower line (Figure 3; Line B) in the scene to match the absolute length of the higher one (Figure 3; Line A). When finished adjusting one set of lines, the participants pressed a button, and viewed the next set of lines. There were 42 sets of lines.

Figure 3: Ponzo Illusion Stimulus Scene

One half of the participants heard ambient environmental sounds (such as crickets chirping, leaves rustling, a dog barking in the distance) created with a ProTools system, and played on a 6 speaker surround sound system (3 front, 2 back, and one high front). The other half of participants did not have any sounds played. Both groups (N=52) overestimated on average of 29%. While these results fit in with the framework that Liebowitz provided, they provide evidence that the current virtual environment did not produce spatial presence commensurate with the real world, but rather that the magnitude of the illusion was similar to that of photographs having strong perspective cues, even with a surround sound environment. Thus, while a subjective sense of presence may be enhanced by a virtual sound environment (Hendrix & Barfield, 1996), the perceptual bias responsible for the Ponzo illusion is present (at least for this virtual environment) only in the magnitude that it would be in a photograph.

4 Conclusion

While a "sense of presence" is often a powerful and compelling subjective phenomenon, there is a paucity of tools to quantitatively evaluate virtual environments along different scales of presence. Some aspects of presence can be measured by subjective questionnaires, but others call for careful comparison to the perceptual distortions that we see in the real world. A litmus test for spatial presence, such as the one discussed above, offers the possibility of presenting quantitative endpoints of presence (photographs and the real world) and evaluating different virtual environments for where they fall on this continuum. This litmus test also reinforces the importance of perceptual processing that is independent of the projected image on the retina. The existence of this litmus test should serve as a reminder to designers of virtual environments that replicating the real world's projected image on the retina will not necessarily result in the same spatial perception as is evoked in the real world.

References

Chapanis, A., & Mankin, D. A. (1967). The Vertical-Horizontal Illusion in a Visually-Rich Environment. *Perception & Psychophysics, 2*(6), 249-255.

Gregory, R. L. (1965). Inappropriate constancy explanation of spatial distortions. *Nature, 207*(4999), 891-893.

Hendrix, C., & Barfield, W. (1996). The sense of presence within auditory virtual environments. *Presence: Teleoperators and Virtual Environments, 5*, 290-301.

Higashiyama, A., & Ueyama, E. (1988). The perception of vertical and horizontal distances in outdoor settings. *Perception & Psychophysics, 44*(2), 151-156.

Leibowitz, H., Brislin, R., Perlmutter, L., & Hennessy, R. (1969). Ponzo perspective illusion as a manifestation of space perception. *Science, 166*(3909), 1174-1176.

Schubert, T., Friedman, F., & Regenbrecht, H. (1999, April 6-7, 1999). *Decomposing the sense of presence: Factor analytic insights.* Paper presented at the Paper presented at the 2nd International Workshop on Presence, University of Essex.

Yang, T. L., Dixon, M. W., & Proffitt, D. R. (1999). Seeing big things: Overestimation of heights is greater for real objects than for objects in pictures. *Perception, 28(4) 1999*.