

Collision Handling in Virtual Environments; Facilitating Natural User Motion

Jeffrey Jacobson & Michael Lewis
goshen@cmu.edu ml@sis.pitt.edu

Dept. of Information Science, University of Pittsburgh.
135, Bellefield Ave, Pittsburgh, PA, 15213 USA
+1 412 624 9468

ABSTRACT

As we move through the world, we get around most objects using low-level psycho-motor behaviors, which require little conscious thought. In most virtual reality (VR) applications, the user also needs to get around objects in a reasonable way. However, most interfaces provide neither the sensory input nor the body control for users to move about as they do in the real world. We studied three prototypical methods of collision handling in VR: Either the user goes through an object like a ghost, stops dead on contact, or slides around it.

In two experiments, subjects used a screen-and-mouse interface to navigate clutter virtual mazes. We found that the third method, sliding, enables superior navigation. While it does not simulate human avoidance behaviors, it does model their effect in terms of object avoidance. Significantly, screen-and-mouse is the both the most used and least studied VR interface.

INTRODUCTION

Collision handling strategies for existing virtual environment (VE) based applications, fall into three broad categories:

- **Ghost** No effect. User passes through object. Used in most architectural walkthroughs [1] and simpler VR implementations. [2]
- **Clunk** Complete stop on contact with object. Seen in the old tank games. [3]
- **Slip** User's movement is deflected making him slide around the object. Used in most high-end video games [4] and advanced graphics libraries. [5]

In this paper, we will:

- Discuss how we locate ourselves in the natural world [6- 8] and in virtual environments. [9-12]
- See how the notion of a personal space, or proxemic zone, [13, 14] can be used to model collision handling.

- Discuss the three collision handling schemes.
- Discuss two experiments which compare the usefulness of the three modes listed above. Subjects explores a succession of four mazes in the first study. fig (1-4) Only one in the second study. (fig 1)
- Analyze the results. In short, slip mode was found to be superior for navigating a cluttered space with only a screen-and-mouse interface.

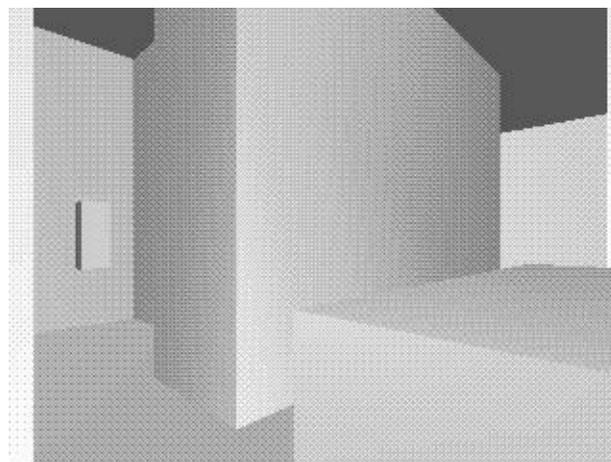


Fig 1: The Baffles maze. The floating box on the left is a treasure.

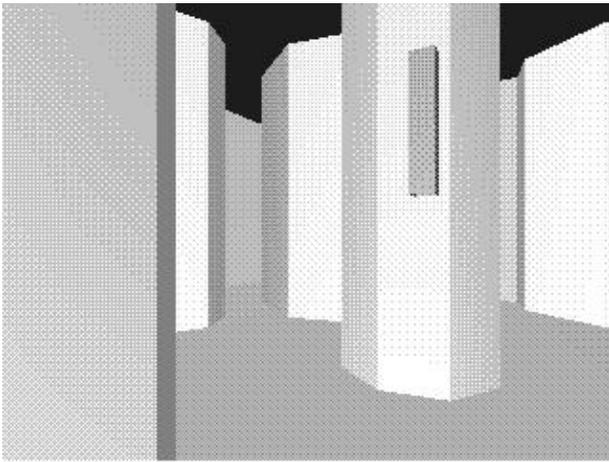


Fig 2: Cylinders maze.

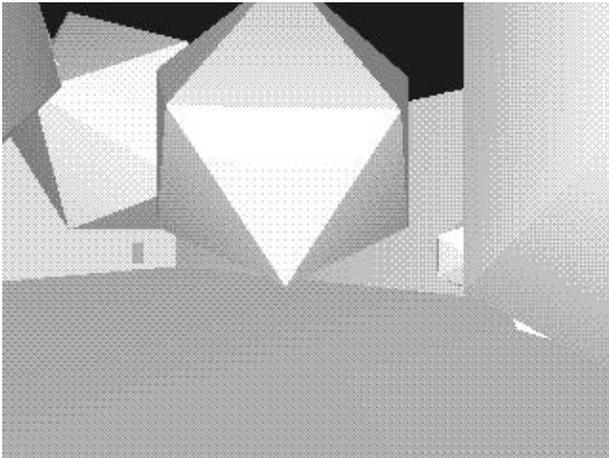


Fig 3: View of the third maze, spheres. A treasure is seen in the distance on the center left.

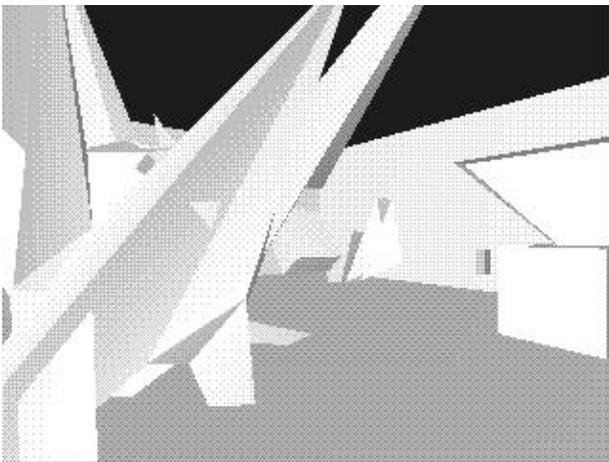


Fig 4: View of the junkyard maze. A treasure is shown in the distance, on the right.

SELF-LOCATION AND NAVIGATION

In the Natural World

We use a host of visual cues to locate ourselves in space-time, such as lighting, perspective, and motion effects to name only a few. [7, 8, 12] Gibson presents a framework for this, based on the concept of “flow perspective”. [6]

This refers to the way the entire scene around the observer changes as he moves through the environment. For example, when he moves forward, everything in the landscape appears to emerge from a point in front of him, (inflow) flow around him on all sides, barring collision, and disappear into the point behind him. (outflow) Objects may change, the landscape may change, but the point of inflow and point of outflow remain the same, relative to his motion. He uses these to locate himself in the environment.

Gibson’s discussion presupposed that the observer is free to move without constantly figuring out exactly where to place his feet to get past every obstacle in his path. We use semi-automatic behaviors for this, which in turn frees the mind to deal with the larger space. For example:

Someone is crossing a room to get to the door, but a chair is in the way. Rather than planning some ideal route, she will probably head directly for the door and step around the chair when she reaches it.

In Turvey’s terms, she is using navigational “atomisms.” These are behaviors one can think of as a single unit of action, such as step-around-medium-object

The Proxemic Zone

We live in a mixture of physical, psychological and social space. [13, 14] Anthropologists have long studied how people of different cultures divide the space between them depending on who they are and what they are doing. Gibson [Gibson] uses the broader notion of “affordances” to describe how people interact with inanimate objects. Relevant to our study, one will keep a certain distance from each object depending on what he is doing (or not doing) with it. For example, the average person will tend to put more distance between himself and a sharp table corner than a pillow.

For our study, we borrow Hall’s terminology to say that each person maintains a proxemic zone around her, within which she does not permit most objects while she is moving through space. In Gibson’s terms, the proxemic zone is the average minimal distance from most objects with the “going by safely” affordance. We use it as the basis for a simple model of collision handling.

In the Virtual Environment

As in the real world, the user locates his viewpoint in a VE with visual cues. For example, when the user rotates in the world, the entire scene pivots around the

viewpoint. That is one way she knows where her viewpoint is, or is supposed to be. [6]

In most VR applications, the user also needs to avoid obstacles. However, most interfaces provide neither the range of sensory inputs nor the body control for the user to move about as they do in the real world. [9] Some mechanism is needed to replace or obviate these behaviors in VR. Otherwise, the user will spend all his time trying to get around every obstacle.

Because of this, variations in the three collision strategies described in the introduction, *clunk*, *ghost* and *slip*, have been used in VE-based applications for some time. Which is best for screen-and-mouse?

METHOD

We conducted two major experiments, where users searched for goals or “treasures” in cluttered virtual mazes. In the first, users searched through four different mazes, under an implied time constraint. They are: Baffles (fig 1), Cylinders (fig 2), Spheres (fig 3), and the Junkyard (fig 4). In the second experiment, they had the whole thirty minutes to Baffles only. (fig 1) By comparing how efficiently users in each group got the treasures, we infer the relative desirability of the navigation modes.

Implementation of Collision Modes

For each of the collision modes, we used a simple, representative implementation. In these, the proxemic zone is modeled as an invisible sixteen-sided cylinder, extending from floor to ceiling, which forms a cage around the user’s viewpoint.

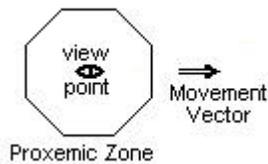


Fig 5: This is an overhead view of a proxemic zone. It had sixteen sides in the experiments, but shown with eight for clarity.

In *ghost* mode, (fig 6) the viewpoint is moved according to mouse input without regard to collisions. The user goes through walls often, which can be disorienting.

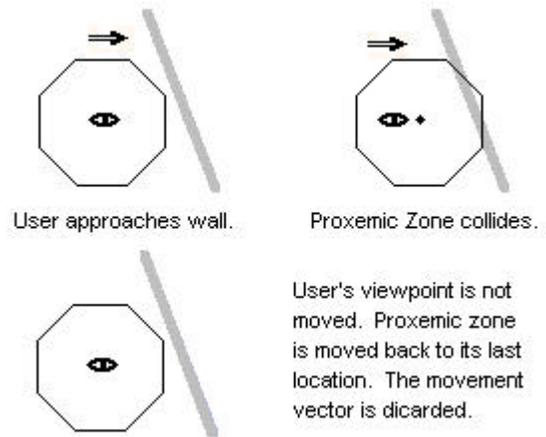


Fig 6: Implementation of clunk mode.

In *slip* mode, (fig 7) the proxemic zone also moves ahead of the viewpoint at the beginning of the frame as with clunk mode. If the zone intersects an object, his movement vector is recomputed so that it preserves as much of his original direction as possible, but is unlikely to result in collision in the next frame.

Testing Procedures

Subjects navigated the mazes (figs 1-4) using a walkthrough application of our own design, which ran on a Pentium class PC in a private lab. Paid subjects were

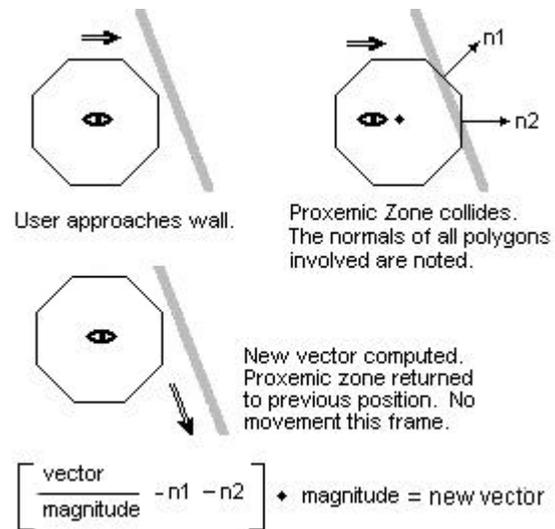


Fig 7: Implementation of slip mode.

recruited from the University of Pittsburgh community through fliers posted on campus. Each were assigned a collision handling mode, shown how to use the application, and told to find all of the treasures in each maze. He was not told anything about the layout of the maze nor the number of treasures in it. When he

thought he had all of the treasures, he was to hit the space bar.

In the first experiment, the tester also told the subject that there are four mazes to get through, and pressing the space bar would move him to the next one. Users appeared to try to get through all the mazes in the thirty paid minutes, usually giving more time to the first one. As a result, they were more rushed than in the second experiment, which had only one maze; there, they were told clearly to take all the time they wanted.

The analysis is based on the average time it took each subject to get from one treasure to another. This excludes time spend before getting the first treasure (orientation effects) and after getting the last one (diligence effects).

FIRST EXPERIMENT

Note how the slip mode users in the baffles maze were able to get more more quickly. (fig 8) Also, the fact that their data points cluster in the upper left of both graphs suggests less variation in their overall experience. The sample sizes are too small for statistical comparisons of significance, but, the results are they are suggestive and led us to perform the second experiment described later.

The results for the other three mazes (figs 9-11) show much convergence in test subject performance. This is probably because the latter mazes are more open than the first, and because of significant learning effects.

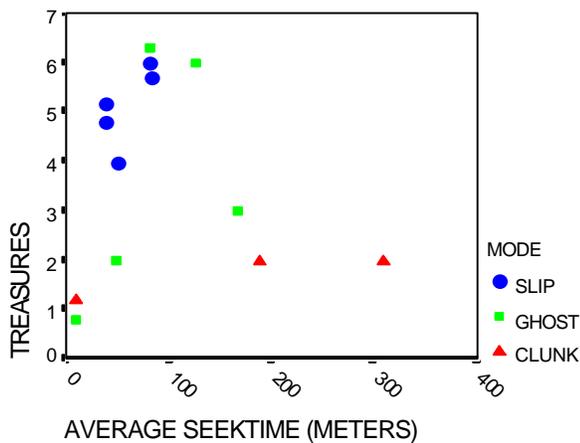


Fig 8: In the Baffles maze. Each data point shows the total number of treasures each subject found in the baffles maze by the average time he spent between treasures

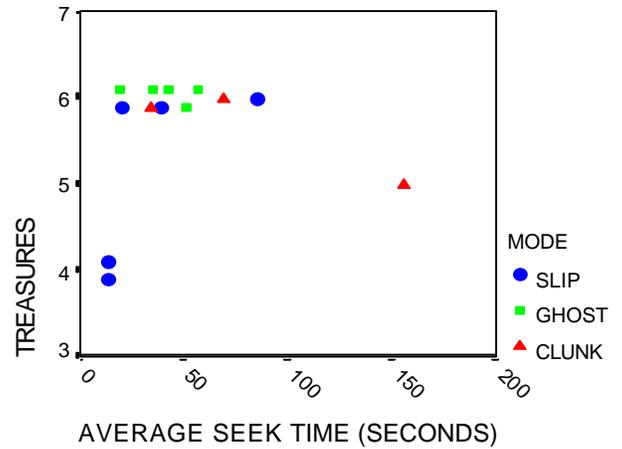


Fig 9: In the Cylinders maze.

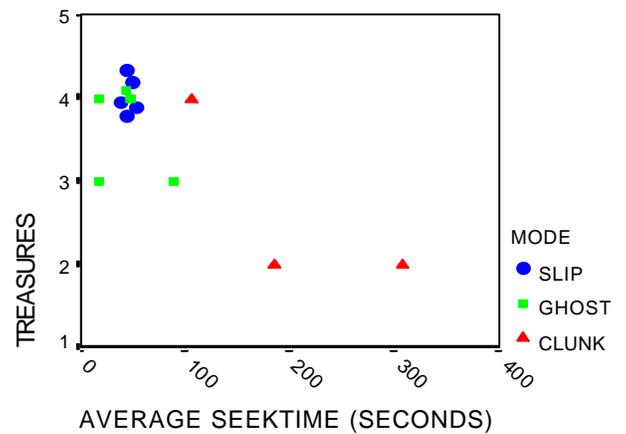


Fig 10: In the Spheres maze.

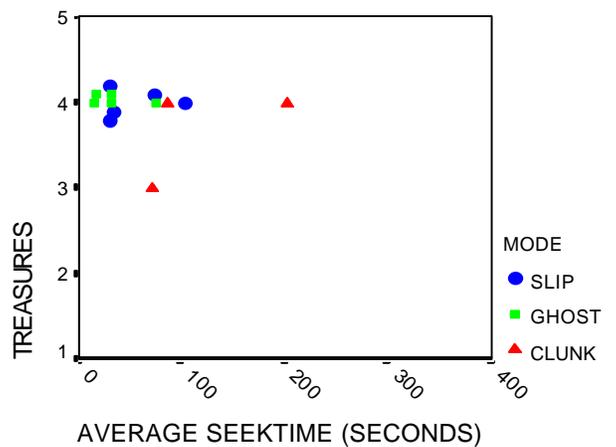


Fig 11: In the Junkyard maze.

We found another interesting measure by comparing the number of treasures found against the average distance the user had to go to get each one. In the case of the Baffles maze (fig 12), which is the only one shown, the resulting graph is remarkably similar to the time

comparisons. Of the three ghost mode users who did best, covered much greater distances than the slip mode users who found a comparable numbers of treasures.

of treasures found by each subject in the baffles maze with the total distance he had to cover during the test.

SECOND EXPERIMENT

In the second experiment, test subjects had the entire thirty minutes to traverse the baffles maze, and some

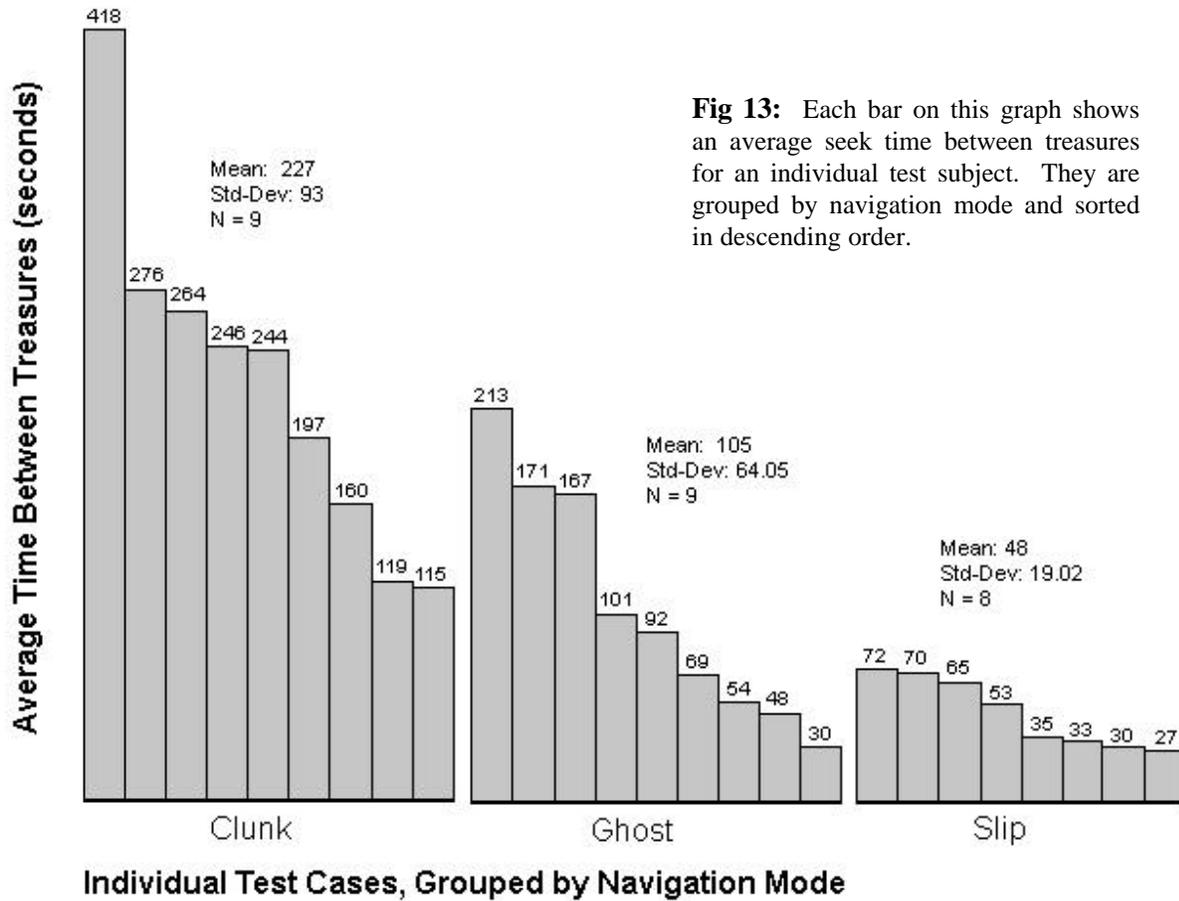


Fig 13: Each bar on this graph shows an average seek time between treasures for an individual test subject. They are grouped by navigation mode and sorted in descending order.

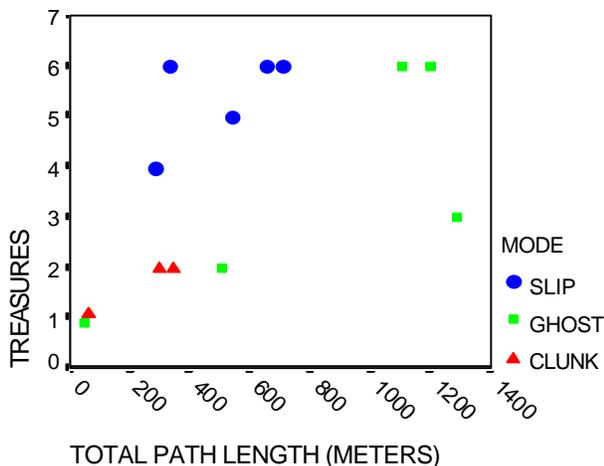


Fig 12: In the Baffles maze. Compares the number

took longer. Consequently, all subjects got all six treasures in the maze. The chart below (fig 13) shows the average time between treasures for all twenty-six subjects. The between-treasure-search-times of the collision resolution strategy group means were found to be significantly different $p < 0.01$.

ANALYSIS

These results indicate that slip mode is an efficient strategy. The Baffles maze (fig 1) was designed to be difficult to traverse, but its features are typical of what is found in many existing virtual environments. The independent development of similar strategies in video games and animation libraries supports the generalizability of our findings.

Why Clunk Mode is Difficult

The user simply loses too much time having to back up, reorient and start moving again after every collision. This can be especially difficult when the collision is with a low-lying obstacle, potentially out of sight, or it is just a glancing touch on something. In Turvey's terms, [15] such an unnatural interaction violates the atomisms we use to navigate around objects. In Gibson's terms, [6] the objects lack sensible affordances with respect to collision.

Why Ghost Mode is Difficult

When a user goes through a wall, it causes an immediate and total change the scene. This interrupts her flow perspective, eliminating the landmarks changing the visual cues she may have been using to locate herself. In Gibson's terms, the flow perspective is interrupted. The effect is more acute with the screen-and-mouse interface, because visual cues are the *only* ones the user has.

In this situation, the only way the user can maintain self-orientation is to estimate the distance travelled while going through the wall and use that to maintain his location in a mental map of the maze. However, ghost mode users have to sort through much more information to build this map, because they have to deal with both sides of every surface and extra garbage between them.

He would have an easier time if the virtual environment were open enough so that landmarks would always be visible. For example, moving through a virtual forest-in-winter (no leaves), he could float his virtual body through the tree branches. However, if he moved his *viewpoint* through a branch, there would be a vertiginous moment where he is seeing the inside of a branch and nothing else. The next moment, his viewpoint will emerge at the other side, and the landscape will look much the same. While effect may be unpleasant, the user is unlikely to become lost because of it.

Another problem is that when the user pushes his viewpoint too close to a wall, it fills his view and he cannot see anything else. Suddenly, the VE looks like nothing more than a sea of the wall's particular color or texture. [16, 17] While this may still happen with clunk and slip mode, it will *always* happen in the moments before the ghost mode user intersects a wall.

Slip Mode is Best

With all this in mind, it could be said that slip mode is most successful, because it allows for the greatest *continuity* in the optic flow and user motion. Rather than attempt to replicate natural interactions, it simply minimizes the delay imposed by obstacles.

CONCLUSION

The utility of collision handling schemes for VR warrants more detailed study. In situations such as desktop VR where interaction techniques are somewhat arbitrary, it is especially important to identify the good ones. Because its low cost and vast user base, most VE-based applications use screen-and-mouse. We envision the eventual development of a wide array of such pseudo-realistic interaction techniques for navigation and manipulation which preserve the intuitions and affordances offered by VR while simplifying and facilitating users' interactions .

REFERENCES

1. Walkthrough Project, UNC
<http://www.cs.unc.edu/~walk/>
2. C. Marrin & B. Campbell. *Teach Yourself VRML2 in 21 Days*, Sams Net, 1997. ISBN 1-57521-193-9
3. Battlezone. Original game by Atari Inc., 1980. PC version available at <http://www.microsoft.com/games/arcade/prodinfo.htm#ZONE>. Atari version: ATARI, INC., Consumer Division, 1312 Crossman Ave., P.O. Box 61657. (800) 538-8543 (in USA) Sunnyvale, CA 94086
4. Quake, ID Software
<http://www.kinglink.com/quake.html>
5. IRIS Performer.
<http://www.sgi.com/Technology/Performer/>
6. J. J. Gibson, *The Ecological Approach to Visual Perception*, Houghton Mifflin Company, 1986. ISBN: 0-395-27049-9
7. R. Warren & A. Wertheim, *Perception and Control of Self-Motion*, 1990, ISBN 0-8058-0517-6
8. I. P. Howard & P. Ian, "Spatial Vision Within Egocentric and Exocentric Frames of Reference," S. R. Ellis, *Pictorial Communication in Virtual and Real Environments*, 1990 ISBN 0-74840-008-7.
9. M. Slater & M. Usoh, "Perceptual systems, Perceptual Position, and Presence in Immersive Virtual Environments," *Presence*, Vol. 2, No. 3, Summer 1993, pp. 221-223
10. Spring, Michael B., "Being There" or Models for Virtual Reality. *ASIS Mid-Year Meeting, Santa Clara, CA*, 1991.
11. D. Zeltzer, "Autonomy, Interaction and Presence," *Presence*, Vol 1, No. 1
12. J. Psocka "Factors Affecting the Location of virtual Egocenters: From the Renaissance to Cyberspace." Draft paper from U.S. Army Research Institute,

ATTN: PERI-IIC, 5001 Eisenhower Ave.
Alexandria, VA 22333-5600

13. E T. Hall, *The Hidden Dimension*, 1966.
ISBN 0-385-08476-5
14. D. Stokols & I. Altman. *Handbook of Environmental Psychology* 1987. ISBN 0-471-63016-0 (v. I).
15. T. T. Michael, & P. N. Kugler, *Information, Natural Law, and the Self-assembly of Rhythmic Movement*, Lawrence EarlBaum Associates, 1987. ISBN 0-89859-640-8
16. M. Slater & M. Usoh. "Simulating Peripheral Vision in Immersive Virtual Environments," *Computers & Graphics*, Vol 17, No. 6, 1993, pp. 643-653.
17. G. Thomas, et. al., "Video Texture Cues Enhance Stereoscopic Depth Perception in a Virtual Reality-Based, Telerobotic Interface." 1995 IEEE 0-7803-2559-1/95 \$4.00