Instant Hole™*

(Windows onto Reality)

Terry S. Yoo

T. Marc Olano

Department of Computer Science
University of North Carolina
Chapel Hill, NC 27599-3175, USA

{yoo | olano}@cs.unc.edu
Phone: yoo - (919) 962-1875
Phone: olano - (919) 962-1918
Fax: (919) 962-1799

Abstract

Inspired by the inventive uses of the fictional product of transportable holes demonstrated in several animated cartoons, we have endeavored to create a environment that allows us to create windows within our virtual worlds. These windows open the often closed environments of head-mounted displays to the richness of the physical world that surrounds us. We have created a video-based stereo see-through head-mounted display to enable the merging of 3D windows onto reality. We describe the many obstacles to merging stereo images generated by a computer graphics system and views of the physical environment. While this paper describes only a single implementation, the issues remain general to anyone intending to construct a similar environment and the motivation is not implementation specific. We point out obstacles, solutions, work-arounds, and other issues for others wishing to attain the same goal. We also discuss the advantages of adding familiar of physical input devices to virtual environments.

*Instant Hole™ is a trademark of the ACME Tool, Die and Gag company, Toontown, CA.. and is protected under international copyright laws.

Introduction

Instant Hole™ is a product that we all know well, and love for its remarkable properties. It is also a product that does not exist in reality. This fictional product is a liquid that can be painted, poured, peeled-up, and reused to make holes in solid surfaces, or to modify reality in mostly humorous and often violent situations. Like its counterpart, Vanishing Cream™, Instant Hole™ is used again and again to make the world change and seem different and to enrich the reality in which our favorite cartoon characters cavort, careen, cajole, and crash.

Instant Hole™ allows roadrunners to cut holes in bridges causing coyotes to plummet to the river below, enables coyotes to paint real highway tunnels onto solid rock, and frees Sergeant Pepper's Lonely Hearts club band from a trap laid by the Blue Meanies. Clearly, this useful product would be indispensable in the home or workplace.

But how does it work, and how do we go about creating these holes for ourselves? Seriously, the idea, while farfetched, does extend into the not-so-real world of virtual reality. We propose that the richness generated by these "holes" in our fictional, fantasy worlds of animated cartoon characters is not so fantastic that it cannot be realized in virtual environments.

We have created a world where physical objects can be brought into the virtual environment with the application of a little video magic. These "holes" allow the richness of the physical world to enhance our virtual environments. They enable us to use our keyboards and monitors while immersed in a virtual environment, to augment the range of virtual control devices at our disposal with physical devices such as dial boxes, trackballs, and joysticks, and they allow us to continue to perceive the world around us, giving us a frame of reference for the physical reality that surrounds us despite our disembodied perceptions.

Motivation

There are two primary ways of using instant hole to bring the real into the virtual world. The first method is to wrap an instant hole around a physical object, making it visible within the virtual environment. In this manner, a joystick, keyboard, telephone, or even a coffee cup can be made available visually.

The other technique is to use flat panels of instant hole to open portals, windows or doors, for viewing scenes outside the virtual space. Typical applications would be to open a window onto a workstation monitor, create a port through which the user converses with a coauthor, or perhaps to generate a panel that permits access to a whiteboard.

In both of these metaphors, three issues are essential to constructing and maintaining the illusion of merged realities. First, it is essential to have well matched physical and virtual environments. The dimensions of each should correspond as closely as possible; a measurement in one space should be reproducible in the other.

Second, the through the lens view of both worlds, one through the video camera, and the other through the virtual viewing plane, should display matched images. Small discrepancies between the virtual camera of the computer graphics system and the actual video camera manifest themselves as blatant perceptual errors. Anomalies in perspective arise from unmatched focal lengths; this corresponds to trying to match two pictures, one taken with a telephoto and the other with a wide angle lens. The result is that the image in the reality window looks less like a merging of realities, and more like a flat photograph or painting.

Finally, the third essential issue is to present the correct view to the user's eye within the head mounted display. This is separate from the through the lens view from the camera. The camera eye view requires knowing the specifications of the camera optics; the user's eye view requires a complete model of the view from within the head mounted display itself. This model is complex, and incorporates the study of visual perception, optics, and computer graphics.

Background

A head-mounted display (HMD) allows its wearer to integrate proprioceptive cues (spatial cues generated by muscles, tendons, and ligaments) into the computer to human interface. People naturally navigate three space by walking or gesturing, and they modify their viewpoint by moving their heads. An HMD allows the wearer to utilize these natural skills to command a computer to provide arbitrary views of a computer generated environment.

The idea of retaining a perception of reality while interacting with a virtual world through an HMD is not new. The original HMD by Ivan Sutherland [Sutherland 68], was constructed using monochrome monitors mounted above the forehead, where images were presented to the user using a set of mirrors. These mirrors were partially transparent, allowing the wearer to see past the virtual image to the room beyond.

However, one of the most important cues to perceiving depth in three dimensions is obscuration, where an nearer object occludes a distant object. Occlusion of physical objects with virtual objects (or the converse) is not possible using the optical partially silvered mirror approach (see Fig 1a). Current HMD products have tended to forego the real world in order to provide a more complete illusion of being immersed within a synthetic environment. Systems such as the Virtual Research flight helmet used at UNC leave the wearer effectively blind to the physical surroundings.

Recently there has been significant interest in matching camera views with synthetic image generation. Gleicher and Witkin published work on a system that performs "through the lens" control. While their work was primarily intended to address animation control systems, it is easily extended to include matching the view of a video camera [Gleicher 92]. In his presentation of high resolution virtual reality [Deering 92], Michael Deering carried this one step farther by demonstrating a monocular view of synthetic objects merged with the view from a spatially tracked camcorder.

Bajura, Ohbuchi, and Fuchs [Bajura 92] demonstrated compelling reasons to provide views of reality from within virtual worlds. Their application integrates a medical ultrasound image acquisition system with a 3D HMD-based display system. By merging the views of both the computer visualization of the medical images with input from a video camera, they superimpose pictures from within the body onto a monocular view of the real world. This analog of "X-ray vision" is inspiring enthusiasm among the medical community and reveals the promise of this technology in the area of computer human interfaces.

Creating the Illusion

The video mechanism behind instant holes is a common technique for overlaying two video signals. One can see this process in operation during ordinary weather reports, where live broadcasters are presented together with video maps and satellite images. This technique of masking or keying from color or intensity allows the actor to occlude portions of the active video display behind him. Through the use of computer graphics, we can create flat shaded objects that display the key intensity or color, enabling the video image from the live source to become visible. As with the ordinary television weather report, we acquire a sense of occlusion and a natural blending of the two signals into a single environment. (Fig 1b & 1c)

Figure 1a - a simulation of an optical see-through HMD

Figure 1b - a view from a video see-through HMD without Instant Hole

Figure 1c- the same view with Instant Hole.

However, the flexibility offered by creating key images through computer graphics raises some critical problems. We must create the illusion that objects in reality and objects in the synthetic worlds correspond. They must behave as if they exist in both worlds seamlessly, remaining invariant through changes of viewing direction and position.

For each physical object we bring into the virtual world, there is a virtual object that encompasses it, covering it with the particular color or intensity that makes it visible. Position, tracking, and dimensional accuracy become essential to preserving the illusion that the real object belongs in the virtual world. We incorporate the coordinate control system of transforms developed by Robinett and Holloway [Robinett 92] in order to aid in the control of the many dimensional variables. Through their model, the instantaneous relationship between two coordinate systems can be described with a single transform that converts a three dimensional point from one world space representation to another. Moreover, the notation represents the relationship of any point to the view position of each eye, to the user's hand, or to other objects that serve as a frame of reference.

For the visibility and occlusion to be convincing, it is not enough that dimensional correlation be correct, but also that the projections of the images onto the viewing plane match as close as possible. When the hole for a real object obscures a virtual object, the video image of that real object must land in that hole. A great deal of our effort was spent in overcoming the obstacles of correctly matching the world views. Photogrammetric techniques were used to measure the camera's center of projection, view orientation, and horizontal and vertical fields of view (FOVs). We then modified the computer image to exactly match these parameters.

However, even if all of the criteria for object correspondence are met, the effort is meaningless if the user does not have the illusion that he is immersed within a separate world. Effective virtual presence requires the perception of constant size, shape, and relative position of objects in a 3D scene. To achieve this, we require an in depth understanding of what the user sees within the head mounted display, what depth cues are essential, and what compromises are acceptable to make in order to achieve object constancy.

Beginning with the stereoscopic optics model formed by Robinett and Rolland [Robinett 92], we considered the parameters for generating merged virtual and physical stereo images for an HMD. Measurements taken of our HMD indicate that the optical axis of the eye does not correspond with the center of the screen (see Fig 2). This fact runs contrary to the nature of video cameras, whose lenses are coaxially mounted with the image plane. This contradiction leads to improper presentation of the images presented to either eye, impairing stereo fusion, and thus destroying the three dimensional illusion.

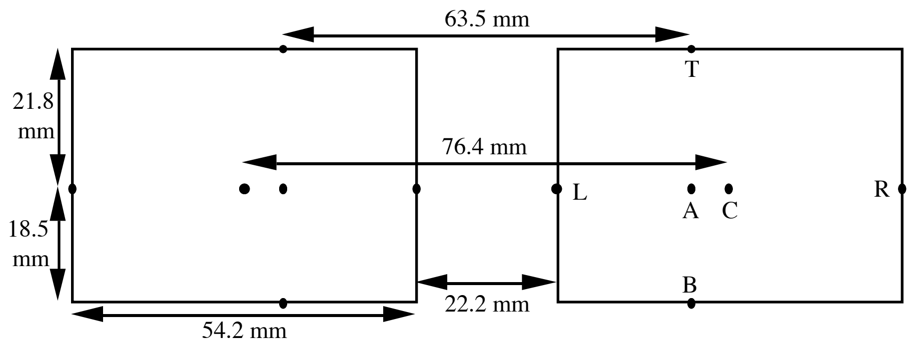

Figure 2 - diagram showing the dimensions of the screen offset

To compensate for the optical axis offset in the HMD, we altered both the placement of the stereo camera mounts on the helmet platform as well as modified the virtual view direction within the synthetic image. The projection axes for both the camera and the computer graphics system are canted outward from parallel, giving the user a comical froglike appearance. The resulting images are distorted in the periphery; however, stereo fusion is enabled, and result is quite satisfactory.

Other perceptual distortions arise from the improper positioning of the cameras. When the cameras are mounted away from the 1st nodal point of the eye, one of several effects become apparent. Errors arising from horizontal offset give a sense of exaggerated stereopsis. If the cameras are mounted above the eye horizon, an impression of added height is perceived. Careful placement of the cameras reduced some of the perceptual errors in our implementation. A system of mirrors is planned to remove other displacement effects that contribute to misleading perception of the world.

Implementation

We built our virtual environment using Pixel-Planes 5 [Fuchs 89], a real time 3-D graphics engine designed and built at UNC. Pixel-Planes 5 is capable of sustaining a polygon rate of over 400,000 Phong shaded triangles per second per eye at a frame rate of approximately 30 Hz. This provides sufficiently smooth image presentation to simulate natural movement in a virtual world. Pixel-Planes contains several frame buffers that, when configured properly, permit multiple users to interact within the same virtual world.

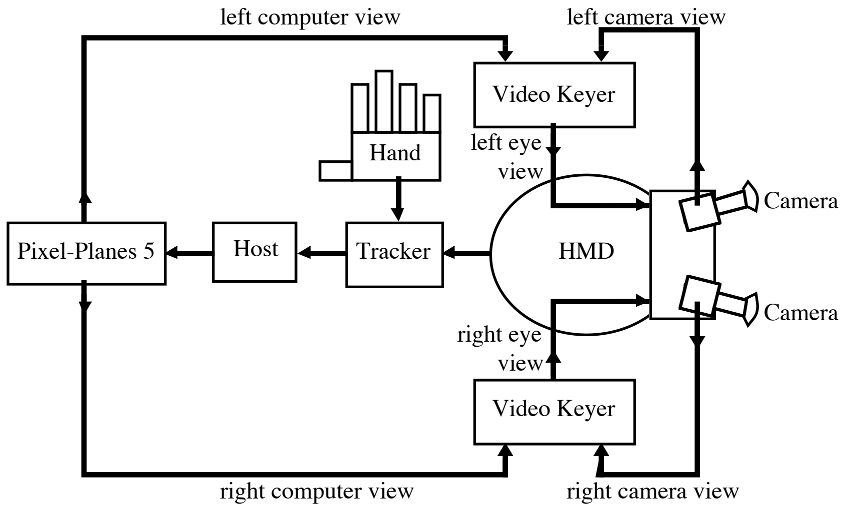

Figure 3 - a block diagram for the Instant Hole system

The software library for Pixel-Planes allows us to specify vertical field of view, pixel aspect ratio (and hence horizontal field of view), view point and direction for each eye, and the optical axis on the screen. We modified a version of 3DM [Butterworth 92], an HMD 3-D modeling program, to accommodate the viewing parameters determined by the camera optics, and the constraints of the video-see-through display.

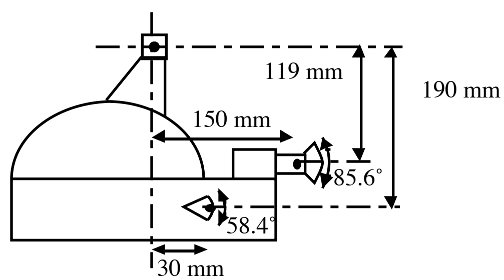

Fig 4 - side view of the Instant Hole HMD

The HMD helmet assembly used as the rigid platform for this work is made by Virtual Research. It is coupled with a Polhemus Isotrak electro-magnetic tracker for tracking head and hand position. To provide the stereo video, we used two lightweight Panasonic color CCD cameras. The cameras were mounted on the helmet about 7cm above the true eye position (see Fig 4). An intensity key system made by Grass Valley Group and a chroma key unit from Sony provide the video blending capability for instant hole.

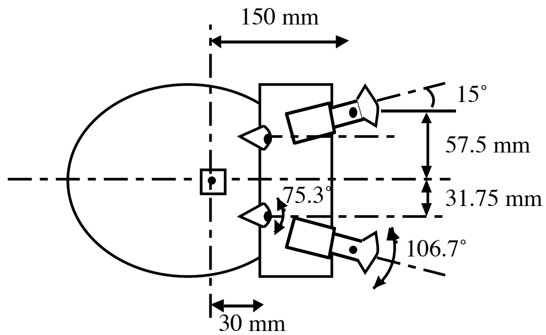

Fig 5 - Diagram showing a top view of the helmet assembly

Design and implementation of the Instant Hole system was largely a matter of accurate measurement, viewing parameter adjustment, and when the components proved inflexible, judgment in selecting a suitable compromise between a loss of fidelity and maintaining perception of 3-space. Most of the compromises made were due to limited capabilities of the cameras. In almost every facet of attempting to match the images with the expected view of the world, it was critical that the camera and computer views be aligned, in spite of the distorted appearance within the HMD view. The camera lenses have a 105˚ horizontal field of view while the HMD has a horizontal field of view of 79.3˚ per eye. In the resulting image, depth perception is slightly distorted so that everything seems farther away than it is. This condition is exacerbated since due to physical constraints, the cameras were about 3.75cm too far apart, which further gave the perception of an distorted sense of depth (Fig 5).

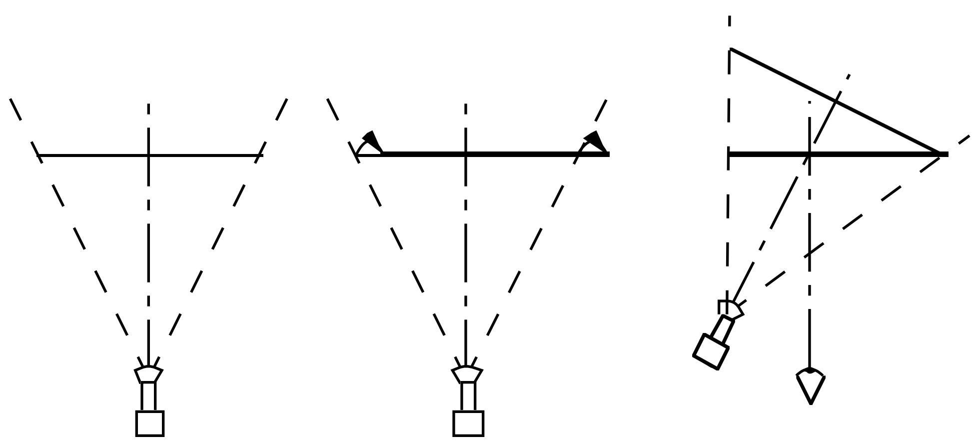

Figure 6 - 15 degree rotation compensation for screen offset

By far the most critical design issue was in determining a means to compensate for the offset of the optical axis within the HMD. The optics in the HMD are such that the view is perpendicular to the screen, but the axis directly in front of they eye is not centered (horizontal offset is approximately 5.9mm, vertical offset is approximately 2.4mm). The cameras, however, are rotationally symmetric and cannot be offset without building a custom lens assembly. But without compensating for this offset, it is impossible to fuse any stereo images. Our solution was to rotate the cameras outward. Rotation is a poor approximation for translation, but it allows stereo fusion for objects close to the center of view at distances further than about a 30 cm from the cameras (Fig. 6).

Results

The resulting system allows the user to model objects in 3 space using a six degree of freedom mouse. The modeler supports a virtual control panel with tools for drawing solid and polygonal figures. It also includes a color pallette for selecting hues. The a particular color (black in our case) can be selected to use as a transparent color, opening holes onto the real world whereever that color is applied.

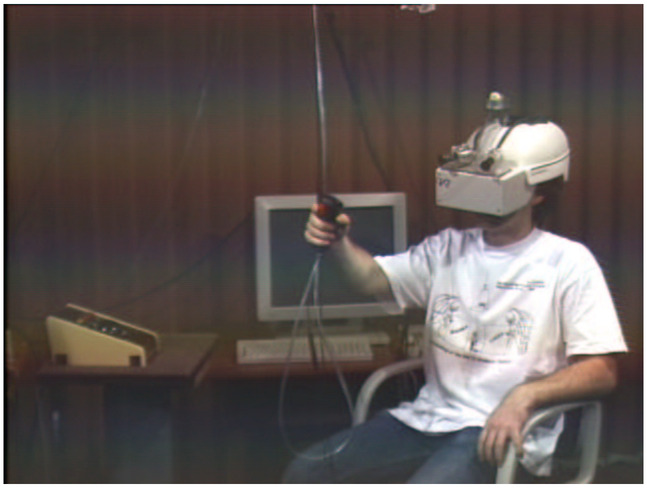

Figure 7 - An author using Instant Hole

Figure 7 shows one of the authors using the system. A keyboard, monitor, and trackball are all within easy reach; each device can be made visible to the user from inside the virtual world. Figure 8 is a view from inside a virtual office environment. This view shows holes onto the monitor screen and provides visual access to the keyboard. A reality window is also shown, allowing individuals outside the virtual environment to make themselves visible to the user.

Figure 8 - A view from inside a virtual office

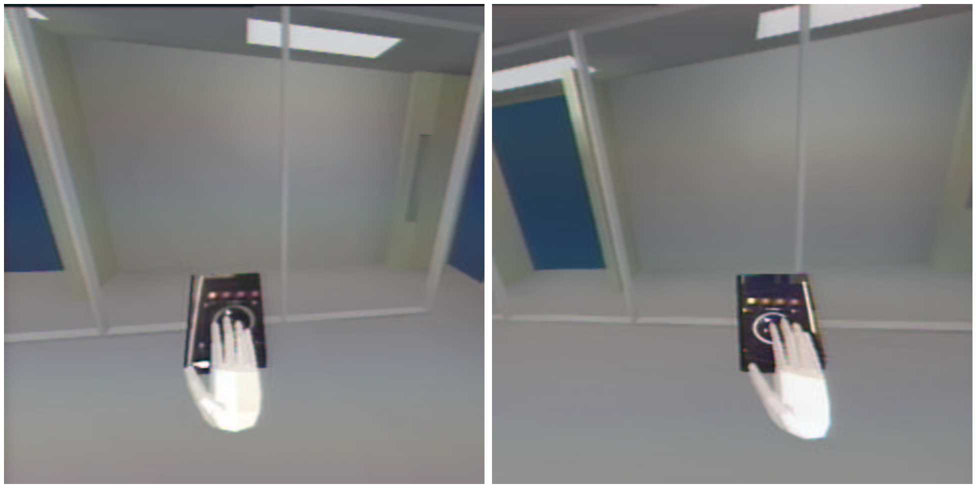

Figure 9 - stereo view of a physical device within a virtual world

Stereo imagery combined with head motion parallax help aid in spatial perception. Figure 9 shows a stereo pair of images where a virtual hand is seen above a physical input device (a trackball) within a virtual room.

Maintaining the Illusion

Several factors combine to counteract the illusion of objects crossing from reality to the virtual world. Largely these factors are limits of existing technology, that impact directly on some perceptual element. These errors often disrupt the spatial perception that we achieve through occlusion, perspective, head motion parallax, and stereo vision.

Latency

We must be concerned with latency when using HMDs because the delays between the actual head motion and the presentation of the updated view are directly apparent to the user. The images seem to swim around, lacking the firm steadiness of the real world. This is painfully apparent in see-through systems where the real world provides a contrast so the user can see just how far behind their view is lagging. The video image of a physical object swimming out from underneath the hole it is supposed to show through can destroy the illusion that the real and virtual environments have actually been merged.

It is important to note that latency is not equivalent to frame rate. We can present a new frame every 33ms. The tracker can provide a new sample every 16.7ms. But the position the tracker gives is 60ms old, and the frame being displayed is the one that was specified a little over two frames previously. The total latency in our HMD system is about 120ms — over a tenth of a second. The video, on the other hand has minimal delay between when the user moves her head and when the camera view changes.

There are three approaches to to reducing image latency. The first is in direct improvements in image generation and tracking technology. The second is to use predictive tracking. Methods using extrapolation methods such as Kalman filtering [Friedmann 92] can provide an estimate of where the head position will be at when the frame is finally displayed. The third is simply to match the latency of the video image to the latency of the virtual image. Video equipment can be used to delay the camera images by four frames, matching the slower HMD system latency. This approach, though it sounds like more of a work-around than a solution, can actually be combined with either of the other approaches to improve the fidelity of the illusion.

Tracking

As stated before, Instant Hole™ relies on an accurate map of both the virtual world and the physical environment. In order to maintain the perception of mixed reality, the system must remain registered at all times.

Electro-magnetic trackers are widely used in virtual environment research. But the nature of these trackers can present some serious problems when attempting to merge real and virtual images. Near the edge of the tracker's field (about a meter from the source for our tracker), the sensed position warps considerably from the true position. For see-through HMD, this translates to the computer generated world canting off sideways while the video image remains fixed. The tracker's field also distorts near metal objects, a problem brought swiftly home when we tried to use a metal stool as a stand for a test physical object. Without the view of the real world for self-calibration, it is possible to effectively work near the edge of the tracker range. With the real world for comparison, the warping quickly becomes obvious and distracting.

Other tracking technologies are emerging that avoid this problem. An experimental optical ceiling tracker [Ward 92] at UNC provides a much larger working environment. The larger working environment also allows placement of physical tools—keyboards, joysticks, etc.—in known places well within the environment without crowding the workspace.

Future Work

Instant hole has exposed many important problems in maintaining the illusion of merged realities and of HMDs in general. Researchers at UNC are working to find their solutions. Current obstacles being examined are exact field of view matching, the addition of projection axis offsets to the image plane of the CCD cameras, correction of the distortions introduced by the optics in the HMD, the improvement of tracking accuracy and calibration, and the development of higher resolution displays.

FOV matching: While the existing CCD cameras are light and compact enough to be used in this application, they do not have sufficiently flexible optical elements to match the FOV of the physical image with the FOV presented within the commercial HMD. An immediate solution to this problem is to add conventional lenses to the cameras at the cost of increased weight for the modified system. Adjustable zoom lenses will enable us to exactly match the FOV presented to the user.

Image plane offset matching: The difficulties with vertical and horizontal projection offsets in the commercial HMD products used in our laboratory can be overcome with an adjustment to the projection center of the cameras. Custom mechanical mounts for the camera optics can physically translate camera lens relative to the CCD plane instead of the axial alignment provided by the current lens screw mount.

HMD optical distortion: UNC researchers are currently addressing the issue of distortion from the HMD's LEEP optics through the design of custom optical elements. Moreover the custom optics will provide a fixed field of view exactly matched to the user and will incorporate the image plane offsets.

Tracking: Research is continuing at UNC to improve tracking technology with the UNC ceiling tracker. The ceiling already provides a larger working volume and reduced tracking latency. Plans are in place to create a tracker lab whose entire volume will be trackable.

Display resolution: The display resolution of the commercial HMD systems in use today at UNC is quite poor. At a distance of four meters, it was barely possible to tell that there might be writing on a standard eye chart. Efforts are currently underway to construct a custom HMD with a resolution of 512x640 color pixels. The displays will be built upon miniature Tektronics 180Hz monochrome monitors, with color supplied through RGB color shutters, to maintain a 512x640 60Hz non-interlaced image. Reality blending will be provided by a video capture system within Pixel-Planes 5. Field sequential frame buffers and video frame capture boards are current projects at UNC.

Improving Instant Hole™: Beyond the arduous task eliminating perceivable errors in the system, many new ideas for extending the utility of instant hole are being considered. The video elements used in our stereo system support the capability of external key images. Using additional frame buffers on Pixel-Planes 5, we propose to generate mask images separate from the existing views on the virtual environment. These mask images will enable psuedo-transparent objects to be merged from either reality, further enhancing the metaphor for instant hole.

Conclusions

In order to generate the Instant Hole™ metaphor for blended virtual and physical worlds, we needed to match the computer generated display with the through the camera lens view of physical reality, match the presented blended images with the view expected by the user dictated by the HMD, and correct for the perceptual infidelities introduced by the whole system.

To calculate the display parameters for these matched video/HMD correct images, it is necessary to consider the specifications of the HMD to be used, as well as the parameters of the cameras used to acquire images of reality. Well-matched cameras and displays are not easily acquired, so compromises must be made that do not destroy the persuasiveness of the illusion.

However, the motivations for pursuing blended physical reality and virtual environments are compelling. The result is a powerful blend of real and virtual, collecting the best of both; allowing the user to extend her perceptions through computer graphics without losing touch with terra firma. There remains no substitute for physical devices to provide accurate control over user input.

Acknowledgments

We'd like to thank Richard Holloway, Jannick Rolland, Ron Azuma and Mike Bajura for their inspiration and helpful suggestions. Special thanks to Gary Bishop for his enthusiasm for student research and his continued joy and wonder at old as well as new things in computer graphics. Our personal gratitude to David Harrison who enables video related work at UNC in computer graphics and to Penny Rheingans for her constant support. Thanks also to Henry Fuchs, Frederick P. Brooks, and Vern Chi whose leadership and continued funding make this and related work possible.

Behind all of this work are a host of people who continually work to broaden the frontiers in immersive display systems. We'd like to express our continuing admiration to members of the ultrasound, tracker, and head-mounted display projects at UNC.

References

Sutherland, Ivan. "A Head Mounted Three Dimensional Display." In Fall Joint Computer Conference, AFIPS Conference Proceedings 33 (1968), 757-764.

Gleicher, Michael, and Andrew Witkin "Through the Lens Camera Control," *Computer Graphics* (Proc. SIGGRAPH), Vol 26, No. 2, July 1992, pp. 331-340.

Deering, Michael "High Resolution Virtual Reality," *Computer Graphics* (Proc. SIGGRAPH), Vol 26, No. 2, July 1992, pp. 195-202.

Bajura, Michael, Henry Fuchs, and Ryutarou Ohbuchi. "Merging Virtual Objects with the Real World," *Computer Graphics* (Proc. SIGGRAPH), Vol 26, No. 2, July 1992, pp. 203-209.

Robinett, Warren and Richard Holloway. "Implementation of Flying, Scaling and Grabbing in Virtual Worlds." In *Proceedings of the ACM Symposium on Interactive 3D Graphics* (Cambridge, Massachusetts, March29-April 1, 1992). 189-192.

Robinett, Warren and Jannick Rolland. "A Computational Model for the Stereoscopic Optics of a Head Mounted Display." In *Presence* 1. 1 (Winter 1992), 45-62.

Fuchs, Henry, John Poulton, John Eyles, Trey Greer, Jack Goldfeather, David Ellsworth, Steven Molnar, Greg Turk, Brice Tebbs, Laura Israel, "Pixel-Planes 5: A Heterogenous Multiprocessor Graphics System Using Processor-Enhanced Memories," *Computer Graphics* (Proc. SIGGRAPH), Vol 23, No. 3, July 1989, pp. 79-88.

Butterworth, Jeff, Andrew Davidson, Stephen Hench, and Marc T. Olano. "3DM: A Three Dimensional Modeler Using a Head-Mounted Display." In *Proceedings of the ACM Symposium on Interactive 3D Graphics* (Cambridge, Massachusetts, March29-April 1, 1992). 135-138

Friedmann, Martin, Tad Starner and Alex Pentland. "Device Synchronization Using an Optimal Linear Filter." In *Proceedings of the ACM Symposium on Interactive 3D Graphics* (Cambridge, Massachusetts, March 29-April 1, 1992). 57-62.

Ward, Mark, Ronald Azuma, Robert Bennett, Stefan Gottschalk, and Henry Fuchs. "A Demonstrated Optical Tracker With Scalable Work Area for Head-Mounted Display Systems" In *Proceedings of the ACM Symposium on Interactive 3D Graphics* (Cambridge, Massachusetts, March29-April 1, 1992). 43-52.