Making Virtual Reality More Real:
Experience with the Virtual Portal

Michael F. Deering
Sun Microsystems Computer Corporation
2550 Garcia Avenue, Mountain View, CA 94043
e-mail address: michael.deering@Eng.Sun.COM

ABSTRACT

The technical limitations of early Virtual Reality systems made them mere teasers for showing the potential of the technology. Since these early days, many researchers have focused on understanding the display factors affecting quality and realism in Virtual Reality display systems. This paper surveys such work, and presents some new data based on experience with the Virtual Portal: a new high-resolution, low-distortion, inclusive Virtual Reality display system, built with three rear screen projectors covering three sides of a small room with head-tracked stereo display. Successes and limitations of this new technology are discussed.

KEYWORDS: Stereoscopic Display, Virtual Reality, Head-Tracking.

INTRODUCTION

Early Virtual Reality hardware made it hard to judge the potential of the field. Ultra-low resolution head-mounted displays, slow head-trackers, and even slower 3D rendering systems required a great leap of faith to believe that systems descendent from these would replace traditional displays for mechanical CAD, medical, simulation, architectural, and entertainment applications.

But advances in technology have improved the quality and impact of the virtual experience to the point where few would argue with the assertion that Virtual Reality is a very powerful new display technology. The more interesting issues revolve around the detailed techniques required for effective Virtual Reality display, and cost trade-offs.

This paper will discuss some of the main work aimed at improving the quality and usefulness of Virtual Reality display systems. Following this, some recent results from use of the Virtual Portal, a new high-resolution, low-distortion, inclusive Virtual Reality display system, will be described. Observations stemming from examining this system should be useful in the construction of other Virtual Reality display systems, as well as in building Virtual Reality software applications.

HISTORY

Ivan Sutherland's pioneering work in building a Virtual Reality system was described in his 1968 paper [29]. This system included most of the key components still present in today's Virtual Reality systems: six-axis head-tracking, stereo head-mounted displays, 3D graphics acceleration hardware, and software to tie the components together into presenting a stereoscopic virtual environment. The paper noted limitations that are only just now beginning to be addressed, such as the fact that the distance between the optical centers between a viewer's eyes varies as their convergence changes. The system also supported what is today being termed Computer Augmented Reality, where a wire-frame computer model of the world is optically superimposed onto the view of the real world.

The next large-scale effort to build head-mounted Virtual Reality systems was Tom Furness' work at Wright-Patterson Air Force Base [5][18]. The effort produced a number of systems, and extended the technology to the use of shaded graphics rendering.

In contrast to this head-mounted display approach, a number of systems were built utilizing external CRT or projection stereo displays. These include [19][27][14][16][28][23][10].

The availability in the mid-80s of inexpensive, lightweight, liquid crystal displays led a number of researchers, including this author, to build a new batch of head-mounted displays and associated systems. Parallel efforts at many sites started a new generation of researchers experimenting with Virtual Reality concepts. This technology still had many limitations. The displays were of very low resolution, many times on the order of 208x138 or less. Wide field of view optics introduced severe image distortions. Most systems employed magnetic position tracking hardware, which had extensive lag and position distortions due to metal interference. Graphics workstations were typically used as image generators, with attendant slow update rates, and limited realism. Nevertheless, such systems reignited interest in the field, and have led to most of today's commercial systems.

More recent research has moved on to address the technological limitations, and to work with real applications.

As will be discussed below, to mimic the interaction of real-world light with an observer under free motion, the real-time position and orientation of the user's head must be known. A high latency or position distorting head-tracker can severely limit the realism of an otherwise good VR system. This has spurred research into understanding the characteristics of ex-
isting tracking technologies [21][1], and methods of improving the quality of the data through post-processing [20][17]. Also, some VR applications need a longer tracking range than a few feet; this has motivated work on some new tracking technologies [30].

Virtual Reality's insatiable demand for more complex, more realistic, higher-frame-rate image generation has become a driving force in the architecture of many high-end 3D graphics accelerators. This was the case in the design of the triangle processor system [9]. The pixel planes systems [22] have also had much application as image generators for Virtual Reality systems. This author has also designed specific support features for Virtual Reality into 3D graphics accelerators for desktop systems [11].

Computer Augmented Reality is a form of Virtual Reality in which the virtual world is superimposed with the real world [15][12]. This technology has likely applications on the manufacturing line, maintenance, and in medical areas [4].

While this paper mainly address the visual aspects of Virtual Reality, work is also ongoing in many other areas, including virtual audio, tactile feedback [6], network- and world-building software, and applications [7], [3] and [24] are two very recent general audience books, with good coverage of the entire field of Virtual Reality.

As Virtual Reality starts to be applied in real applications and become a commercially viable technology, it is inevitable that interesting work will be done by commercial entities for profit rather than publication. All in all, this is to the good, but it means that some results and technical details will be proprietary.

LESSONS FROM PRIOR SUN VR WORK

Several different forms of Virtual Reality display systems have been built at Sun Microsystems over the last several years [10][12][25]. In this process, a number of lessons have been learned about what makes for effective Virtual Reality displays. While formal perceptual experiments with test subjects are only just starting, several empirical observations can be stated (many are discussed in detail in [10]):

Proper Generation of Head-Projected Stereo Images

The laws of physics define how rays of light fill the three-dimensional space we inhabit. As we move through space, our light-receptive organs (eyes) sample different portions of this field of photons by means of localized image planes, closely approximated by the mathematics of projective geometry. The human visual system evolved under such constraints, hard-wiring in many of them. Our visual sense of "reality" is fundamentally based on mono and stereo images changing "just so" as we freely move our heads and bodies through space. The practical question for Virtual Reality is how precisely does this physics have to be simulated to satisfy our visual system, and how does this satisfaction fall off as the simulation is simplified.

The evolving answer is that even small distortions in simulated images, caused by Virtual Reality display systems, are quite perceptible by all persons with near normal (corrected) vision. Such distortions rapidly become an impediment to accepting the virtual imagery as real, and rapidly diminishing the sense of "presence". Because most of our perception of space is processed by low-level brain mechanisms, unlike some other visual perceptual modifications, these distortions cannot be "learned" away by repeated experience.

Recent experiments [2] are showing that monocular motion cues caused by viewer directed head movement are an even stronger source of 3D information than static stereo imagery.

The Need for Accurate Physical Calibration

The physical geometry of the display must be accurately calibrated: what is the precise size or field of view of the display raster in physical units? Where is the viewer's head in relation to this? What is the viewer's individual intracocular distance? (With our systems, each viewer's individual intracocular spacing is first measured with an interpupilometer.)

Low-Latency Accurate Head-Tracking Information

Head-tracking data must be accurate, low latency, and predicted into the future: Frame rates must be in excess of 12 Hz, preferably closer to 20. At any lower rates the virtual images get out of sync with human head movement.

Correct for Sources of Optical Distortion

Optical distortions of the display must be corrected to very high accuracy: nearly all existing wide-field-of-view, head-mounted display optical designs have too far much distortion. [26] gives a good description of the distortion function and the negative effects on stereopsis caused by it. Even if stereo workstation CRTs are used, the magnification and curvature distortion of the CRT glass front plate must be corrected for [10]. Optical distortions make it impossible for human stereo vision to perceive a stabilized location in space for objects. As a result, objects appear to swim and shift in position with viewer movement.

Pixel Resolution

The resolution of the display must be better than "legal blindness": at 10 minutes of arc per pixel or better. There is a good description of resolution definitions and human perception in [8]. Our work at Sun is biased; we want virtual space to be a place in which our users can perform productive work; not being able to read normal 80-column text is too severe a resolution limitation.

Image Realism

While in general, we have found that the ultimate in photorealistic image rendering quality is not necessary for effective virtual displays (especially when traded off against minimal frame update rates), higher quality base primitives help. Thus anti-aliased dots and lines, and smooth shading can greatly increase the perceived resolution of the display.

Trading off Limits

If one accepts lower accuracy and higher latency in a Virtual Reality system, what is the associated trade off in the quality
of the Virtual Reality experience? Our results show that the virtual images lose their stability in space; viewers cannot accurately localize them in space relative to their bodies or other physical objects. The images seem less real, more artificial. This has greater impact on some applications than others. An MCAD virtual machining or assembly task may have very high accuracy requirements in contrast to what is necessary for an architectural walk-through application.

THE VIRTUAL PORTAL

Initial work on the Virtual Holographic Workstation [10] showed that a high-resolution head-tracked stereo display could produce strikingly effective three-dimensional imagery. The system offered several advantages over most head-mounted displays: much higher resolution, nearly no image distortion due to optics, and a very lightweight user interface. But with a field of view of less than 45°, it did not support fully immersive Virtual Reality. The Virtual Portal was the result of our attempt to build an inclusive display interface while retaining many of the advantages of the Virtual Holographic Workstation. Our approach was to use stereo rear screen projection CRTs to cover the user’s field of view with pixels. The Virtual Portal (see figure 1) is a small 6-foot by 6-foot room, three walls of which are actually floor-to-ceiling (8 foot) rear-projection screens. Behind each screen is a dedicated projection CRT and a controlling graphics workstation. The user dons a lightweight pair of stereo head-tracked glasses, and is free to move about the room, as well

Figure 1. Diagram of the Virtual Portal. The viewer’s position in the room is dynamically tracked by the computer. Each projection screen is displaying real-time animated stereo images, as determined by the viewers current position.
as interact with the virtual environment with a 6-axis 3D mouse. The CAVE [8] is a similar system.

All three of the video projectors are stereo genlocked together, and in sync with the users field-sequential stereo shutter glasses. The Electrohome ECP4100 projectors have a special StereoGraphics Corp. fast decay green phosphor CRT tube to minimize “ghosting” one eye’s image into the other. Currently the Virtual Portal uses three SPARCstation 2GTs for its image generators, connected via Ethernet. The system master broadcasts head-track and simulated physics information to the two other workstations. The display resolution is 960x680 square pixels per eye, per screen. The screen is refreshed at 108 Hz, 54 Hz for each eye. The motion update is at a lower multiple, typically 13 to 18 frame per second.

The Virtual World Displayed

A viewer in the Virtual Portal experiences a series of animations, each segment emphasizing a different facet of the display possibilities of the portal. Written text is a poor substitute for the experience itself, but some of the flavor of the display can be gleaned from the description of the sequences below:

Calibration Grid. To start out, the walls are made visible by displaying a simple calibration grid. This also allows the system to be visually inspected to ensure that no projection parameters have been electronically or physically changed.

Sea Cliff. As the sequences start, the calibration grid fades away, and the viewer finds herself floating above a sea cliff overlooking a bay with a boat at anchor. Two inquisitive seagulls fly down to investigate (see figure 2). Meanwhile, the viewer starts floating down to sea level, stopping with only her neck above the surface of the water.

Under Sea. The viewer is instructed to duck down under the water (by physically squatting), and now is viewing the cliffs under the waterline. Now two large fish swim up to see what’s going on.

Large Mirrors. The viewer’s head is mirrored on all three sides by synthetic heads, using the six-axis head-tracking information. The heads then expand to 27 times their normal volume, still mimicking the viewer’s every head motion.

Hatchet & Arrow. Several hatchet blades chop through the walls, and then a five-foot-long arrow flies through from the right side (see figure 3).

Object Potpourri. Twelve different MCAD, BioCAD, and other 3D objects are shown in rapid succession.

Floating Cubes. Several hundred multi-color cubes, about 2 inches in size, form a lattice in space, into which the user can walk.

Radiosity Lit Studio Apartment. The viewer is standing on the ground floor of the interior of a complete, two-level small apartment, all pre-lit by radiosity techniques.

Space War. This is a full three-dimensional implementation of the original computer graphics game, as two space ships fight while orbiting about a central body. In interactive mode, the 3D mouse is used to control one of the ships.

Large Virtual Lathe. A virtual lathe demonstration with a six-foot lathe stock. The viewer can cut into the stock with the 3D interactive mouse, causing sparks to fly, and appropriate audio grinding noises.

Toothpaste. A thin tube of material is extruded wherever the user waves the interactive control, allowing the creation of complex, hanging rope shapes, including knots.

Night Swamp. The viewer is traveling through a swamp of tall plants at night in a very dark environment, lit by occasional lightning flashes.

3D Programming Environment. Here, three poster-sized pieces of parchment are actually three VT100 text terminal emulations, with a 3D cuckoo clock to keep the time.

In running subjects through the Virtual Portal experience described above, a number of visual effects were noticed.

Perceptual Resolution

To see an object in more detail in the real world, one moves closer to it (or it closer to you). Closer viewing distance translates into a larger field of view of the object, imaging the object onto a greater number of rods and cones. In the Virtual Portal, resolution is many times limited by pixel density on the projection screens. So, as one moves closer to a virtual object, if the object's position is on the side of the screen opposite the viewer, the size of the object in pixels actually decreases, even though its retinal image grows. The perceptual effect is that somehow the object is less detailed, the exact opposite of our real-world-based expectations. Similarly, leaning back can unexpectedly increase the resolution of a distant object. Objects at the position of the screen remain constant in pixel size as the viewer moves, but do not grow in resolution when closely approached. Objects inside the screen with the viewer act in the opposite way to those outside: they change resolution, more like our natural expectations dictate. Even these exhibit surprising behavior. When one approaches a small object within the room, it will appear to gain an incredible amount of resolution. It is only when the screens are viewed without the shutter glasses that one realizes that even a physically small object's projection might occupy most of the screen area.

Depth of Field

The major visual cue not properly simulated by most headtracked stereo display systems is eye focus, or depth of field. With the Virtual Holographic Workstation, some viewers sitting 16 inches from the screen cannot converge images of virtual objects more than a few inches in front of or behind the screen. In the Virtual Portal, viewers are typically further from the screen than that, and fewer depth of field problems have been reported. Another reason is the lack of contrasting cues: in the Virtual Holographic Workstation, the viewer can still see and focus on other parts of the room and workstation, while in the portal, there are no such reference cues.
Practically everything the viewer can see has been generated by the computer.

**Realism Achieved**

While there are as yet no objective measurements of the "realism" of virtual displays, the Virtual Portal scores high on the subjective measurement of nearly all viewers who have been in it. This can be illustrated by the Hatchet & Arrow segment of the presentation. After distracting the viewer with a few virtual axe blades chopping through the walls, a five-foot virtual arrow is thrown from right to left in front of their eyes. Since the system has real time knowledge of the viewer's head placement, the arrow is always thrown at current eye level, six inches in front of their face. The arrow passes into the left-hand wall, with the shaft still hanging in front of the viewer's eyes. Viewers seem to instantly figure out what has happened without much cognitive processing. Many viewers then duck a few inches so that their head passes under the arrow shaft and move to the other side of the arrow to look at it (see figure 3).

Another indication of the degree of presence achieved is how rapidly and completely the viewer forgets the location or presence of the projection wall screens. We found that we had to put up a confining railing to keep viewers from walking or reaching right through the screen itself. Indeed, even when actively attempting to perceive the location of the screen, unless one focused on a slight texture in the screen material itself, its location is impossible to accurately gauge.

**Effect of Wide Field of View**

The viewer’s field of view is limited much of the time only by the edges of the stereo shutter glasses, approximately 95° horizontally by 77° vertically, with 74° binocular overlap; about the same as normal eye glasses horizontally, slightly worse vertically. The considerable amount of peripheral vision greatly adds to the sense of presence. One example of this is illustrated by what happens in the Floating Cubes segment as the viewer walks through the lattice of cubes. Even after a nearby cube has left her field of view to the side or below, she has a very strong sensation that the "cube is still there", and that she knows exactly where it is, and could even bite it.

**Contrast Ratio**

As a display system, the Virtual Portal can achieve very high dynamic contrast ratios. This is because the projection CRTs are the only source of light in the Virtual Portal. The walls and ceilings have been painted matte black, and the floor is covered with black carpet to minimize any internal reflection of light behind the screens. Thus the viewer can be plunged from bright illumination into near pitch blackness.

An effect found years ago by the flight simulation community is that at low light levels, different portions of the human visual system are utilized, and low light images can actually feel much more realistic. This effect worked out well for training for night landings. The effect also holds in the Virtual Portal. We built the Night Grass segment with no illumination. The only image of the plant stalks perceived was when the stalks occluded a dim star field background, or a small colored sunset ramp at the horizon. Despite this impoverished background, many viewers felt that this was the most realistic environment in the system. To give some additional information, we added occasional “lighting flashes”, where for one frame (1/13 or 1/18 second) the plants are fully illuminated and the sky goes from black to yellow. Synchronized with digital thunder sounds, the overall effect is quite striking.

**Use of 3D Mouse**

For user interaction, we employ a variant of Logitech’s 3D mouse, comprised of a pistol grip with three buttons. This is used to direct the cutting tip of the Virtual Lathe, as the drawing tip for Toothpaste, and as an orientation control for the Space War game.

**Technical Limitations**

**View Angle Dependent Screen Brightness.** Our three rear projection screens are actually made up of one piece of material, wrapped at 90° around a thin cable under tension at two corners. While the cables are thin enough to be almost transparent, the screen joint is easily noticeable by the abrupt change in intensity of the image. The reason for this is the off-axis attenuation of the image on the screen material itself. Typically, one screen is being viewed at a low incidence angle, the next one over at a high angle. Newer screen materials may greatly reduce this effect. Another alternative is to correct the intensities digitally. Since the eye position of the viewer is known, the amount of attenuation can be calculated. Then the brighter of the two screens can be dynamically reduced in intensity to match the other. Since the angle of incidence actually varies across each screen, the intensity could be ramped down according to this function.

**Screen Warping.** Even though the screen edge cables are under high tension, the tension on the screen material is greater, and the cables bow by nearly half an inch in the middle of the span. We correct for this using the pin cushion distortion adjustment of the projection CRTs to match the slight curve of the screen edge. It is not mathematically correct, but reduces visual discontinuities at the seam. To put the resulting distortion in perspective, over the six-foot-wide screen area, the roughly 1% distortion is less than that found on desktop CRTs, and roughly the same as the present absolute positional accuracy of the tracking technology.

**One Person at a Time.** The Virtual Portal, like nearly every other head-tracked stereo Virtual Reality display system, is inherently a single user system, in that only one person at a time can properly experience its effect. Additional people can wear stereo shutter glasses, but the stereo view is being computed for such a different point of view that the additional people often cannot converge the image, or will see a very distorted image.

**6-foot by 6-foot Room.** The small size of the room limits the area in which the viewer can walk about. A larger room would require enormous rear areas for the projector light paths, and would further limit brightness. An alternative might be a form of 2D treadmill. So far, we have limited ourselves to objects displayable inside the room, or to putting
the viewer on a virtual platform the size of the room that travels about the virtual world.

Not 360° View. The three projectors in the current Virtual Portal only cover 50% of the possible view: the floor, ceiling, and wall behind the viewer are black. For applications that need more coverage, this can be achieved with additional projectors at further expense ([8] covered the floor with a front projector at SIGGRAPH92). For our experimental purposes, we have found the three-screen approach to be an acceptable compromise, fewer viewers ever note seeing the non-projected areas.

Projector Tweaking. While overall we are very pleased with the quality of the video projectors, the convergence of the projectors drifts over time and must be touched up almost daily for accurate calibration.

Graphics Technology Limitations

Front Clipping Plane Problem. In the conventional graphics pipeline, the front and rear clipping planes are parallel to the image plane. But for the Virtual Portal, this is not optimal. The reason is that although the image plane is always parallel to the screen (it is the screen), the viewer is not always facing a screen head on, but can be facing it at a high angle. The problem is that the front clipping plane, which we usually place a few inches in front of the viewer’s nose, is not parallel to the plane of the viewer’s face, but always to the screen. This means that objects coming at the viewer get clipped at an apparently high angle, and is not as natural as parallel clipping. This is only soluble by going to a non-standard front clipping plane equation.

Z-buffer Range Restrictions. As is known by the simulation industry, the conventional Z-buffer formulation has numerical round-off problems when both very near and very far geometry must be shown. The problem is characterized by the ratio of the distance to the front clipping plane F to the distance to the back plane B. If objects are allowed to come right into the room close to the viewer before being clipped, then F must be set to on the order of 10 cm. Given a 24-bit Z-buffer, the entire far half of the display space, distances B/2 to B, can only use 24 - log2(B/10) bits to represent distances. This will be further reduced by numerical round off by log2(n/2) bits for a n-pixel wide display polygon. This is not too bad for virtual worlds out to a few meters, but a 100 pixel polygon at 1 km will have to live in a Z-buffer with only 8 bits for the entire last half km of virtual space (2 meter quantization). This problem can be avoided by a non-standard formulation of Z.

Scene Complexity. The requirement to keep the image rendering complexity to that which can be rendered in 1/26th of a second (13 Hz for both eyes) or less severely restricts the polygons budget for virtual worlds. Even though the current image generators are rated at 100K triangles per second, 1/26th of this leaves only about 3,000 triangles per scene. For many application areas, much higher complexity is needed. To understand these future requirements, several hundred industrial objects have been analyzed for typical rendering performance, and to determine what is driving the triangle counts [13]. Initial results show the expected: many industrial Virtual Reality applications need one to two orders of magnitude of improvement in display performance.

FUTURE WORK

Our initial set of experiments with the Virtual Portal were not targeted at applications per se, but at understanding the range and quality of visual experiences that the new technology could produce. Now that the display potential is better understood, appropriate application areas can be investigated. This next stage of work will also include more formal experiments measuring user performance in visual application tasks.

CONCLUSIONS

As technology allows us to build display systems to more and more completely match the visual cues expected by low level human perception, Virtual Reality displays will continue to increase in realism. This trend is confirmed by experience with the Virtual Portal: low-latency, low-distortion, high-resolution, high-frame rate, wide field-of-view shaded stereo images can increase the degree of presence.

ACKNOWLEDGEMENTS

The author would like to thank Will Shelton, Michael Neilly, and Keith Hargrove for their help in making the Virtual Portal a reality.

REFERENCES


Figure 2. Entrance to the Virtual Portal, Sea Cliff segment.

Figure 3. User ducking virtual arrow, Hatchet & Arrow segment.