VIRTUAL REALITY FOR HOME GAME MACHINES

Paper presented by
Allen Becker
President, Reflection Technology
Waltham, Massachusetts USA

at
"The International Symposium
on Artificial Reality and Tele-existence"
Tokyo, Japan

July 9, 1991

Virtual Reality for Home Game Machines

For the purpose of this presentation, Virtual Reality (VR) means a system with headmounted virtual displays, a head position sensor, and appropriate software, processing, and controls so that the user sees an artificial world around him as he turns his head.

For some time there has been a discussion of consumer products based on this type of technology. Most VR researchers believe that a high-quality low-cost system suitable for games, at a price point competitive with a 16-bit home game machine, will not be technically possible for several years.

They are wrong. We have built a demonstration prototype of a game system which can be price-competitive with the 16-bit game machines, including the game computer, display, head sensor, and controls. We have tested this system with game developers and with focus groups of teenaged boys. They prefer this game system to the Super Famicom by a ratio of 4 to 1.

Most workers in the field of VR will strongly disagree with our conclusions. Those workers should see our demonstration system, and they should bring their teenaged boys to use our demonstration system.

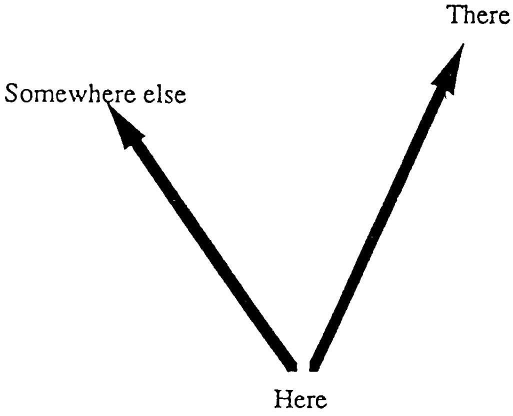
VR is primarily a visual experience, and most of the reasons why low-cost VR games are now possible are related to display development. Much of my talk will center on display parameters. However, other subsystems are important, and I will also discuss these systems to a lesser extent.

All of my presentation is in the context of what is important to make VR a good game experience, as opposed to what is important to make VR simulate the real world. Games and animated cartoons do not strive to duplicate the real world; they strive to provide entertainment at an attractive price by creating an artificial world which is internally consistent and enjoyable. Game designers have learned which features are cost-effective and perceptually important. They optimize these features and ignore parameters which are expensive or difficult and have little perceptual advantage.

VR system designers must learn the same skills for VR games to succeed as consumer products.

The most important principle to understand is that VR is not reality. For laboratory or research purposes, it is productive work to build systems that try to duplicate reality as closely as possible regardless of the expense. For consumer products it is more important to discover which parameters make the system fun to use at a low cost.

"Suspension of disbelief" is the English phrase used to describe the feeling people should have when they watch a play or a movie or a cartoon. They know it isn't the real world, but if the show is good, they accept it as a form of alternate reality which is a pleasant and consistent experience.



The VR system designer starts out **Here**, in the real world[1]. We want to make a system that will make the user think he is **There**, some different place in the real world. But that is not really possible; it will be many years before the user really thinks he is in some other part of the real world. So the user ends up **Somewhere else**.

Somewhere else might be fun or it might be dull. The most important point for a game designer to understand is that the amount of fun is not closely correlated with the amount of realism. Mario is a very popular game even though Mario does not look like a real person, and his motions don't look very realistic.

Like Mario, our demonstration system is not very realistic. In our demonstration system, the player moves in a monochromatic red wire frame world. Both eyes see the same image, and the system senses only head rotation, not translation.

But our system is fun to play. The player can drive through the city; as he turns his head sideways he sees the sides of the buildings; if he looks down he sees buildings below him. He can look up and see the helicopters flying; he can use his joystick to fly up and chase and shoot the helicopters. The player feels he is in a different world. Like watching a cartoon, it isn't a different "real" world, but it is a consistent and enjoyable world.

In our focus groups, players say that what they like is that they are "inside" the game and that the game is "realistic." They are not here, they are **Somewhere else**, and they think it is an enjoyable place to be.

Important parameters for VR game displays

I will discuss display parameters in four major, partially overlapping categories:

- Local pixel quality
- Global display appearance
- Global economic specifications
- Unit manufacturing cost

Local pixel quality refers to the detailed appearance of each pixel: smoothness, clarity, and contrast.

Global display appearance refers to the overall image quality: speed, persistence, need for text, and lack of a visible display background.

Global economic specifications refers to the display parameters that must be optimized in a price-performance envelope for best game-play at lowest cost: field-of-view (FOV) optimization, color, resolution, and gray scale.

Unit manufacturing cost is self-explanatory and extremely important.

I. Local pixel quality

A. No visible pixel structure--pixels should be locally smooth

With a full-sized display, the player sits far away and cannot resolve the pixel structure, and even for full-sized CRTs shrinking the shadow mask pitch is a needed technical improvement. With a clearly magnified virtual display, any pixel structure is much more obvious. Active elements of AMLCDs are visible, boundary lines between pixels are easily visible, and each of the separate color triads of a color display are easily distinguishable.

Human eyes have very good edge detection capability and bias, and visible pixel features make the scene less real. Using a diffuser to obscure the pixel structure, as is commonly done with LCD VR displays, is a poor solution. A better solution is to use a display technology which gives solid, contiguous pixels.

With Private Eye display technology, the pixel fill ratio can be 100%, and adjacent pixels can touch each other. This makes a substantial improvement in quality of image. See Figure 1.

Image clarity

Images should be in focus. Headmount displays are viewed through an optical system, and any defects in the optics detracts from the perceived quality of the image. The intentional use of diffusers to hide visible problems with the display would not be considered an acceptable technique in any serious discipline. Sharp images contribute to a feeling of fast game play and system quality.

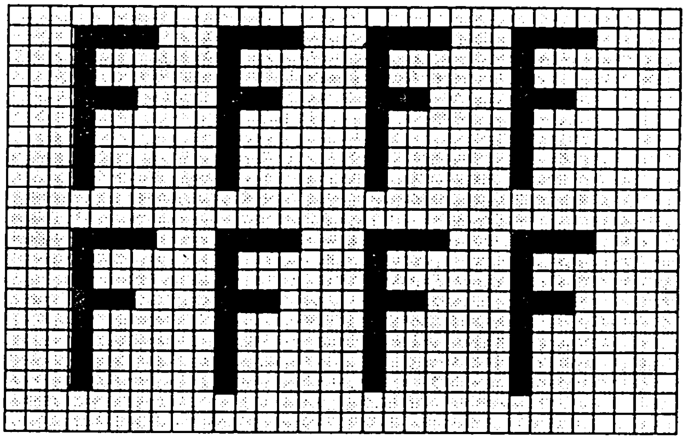
Although people can tolerate and adapt to defective optics, they don't normally choose to-- people who need reading glasses don't take them off to intentionally blur the image when using a low-resolution display.

Pixel structure and contrast

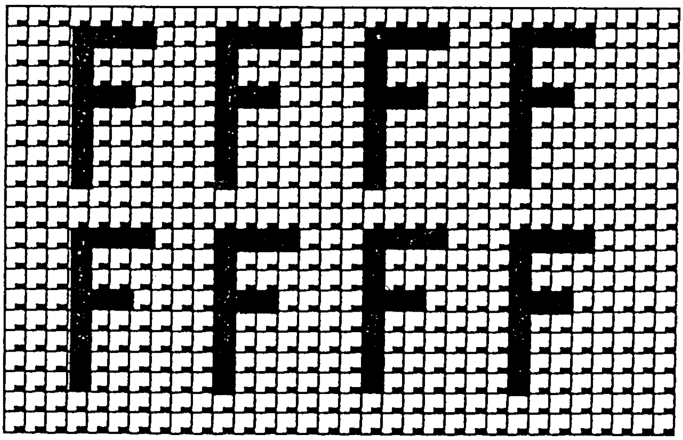
Figure 1

Private Eye

STN-LCD



AM-LCD



Most of the LCD-based VR systems use diffusers or more sophisticated optics[2] to blur the images because of the terrible problems with visible pixel structure. Sophisticated optical tricks to hide problems not such a good solution--a much better solution is to use a display technology which inherently has solid, contiguous pixels. With our display system, images are crisp and well-focused.

The unavoidable problem (at least with a raster system) is that of jagged lines due to limited resolution. Good anti-aliasing can reduce the perceived effects of this unavoidable problem. However, blurring is not a good way to anti-alias.

Crisp, well-focussed images are particularly important for text, and games frequently require text. See comments on Text below.

Contrast

High contrast is beneficial. This fact has nothing to do with the unique nature of a headmounted display. Just like any other display system, people appreciate high contrast and they rate contrast as an important perceptual parameter. We believe the extremely high contrast of the Private Eye display contributes significantly to the high perceived "presence" of the image.

Very high contrast also lets the display background be invisible, see below.

II. Global display appearance

Speed

Update speed and low display persistence is critical. Motion blur is tolerable in a handheld game; in a fast moving VR game with head tracking motion blur is terrible. Other parameters affect system speed, described below in the discussion of other system parameters, but the display should be as fast as possible.

Fast response is one of the most important parameters for making game-play fun! Walk into a video-game arcade and look at the systems.

No visible display background

LCD or CRT displays have a clearly visible rectangular display background. At any economic sized field of view, a visible raster gives the user the feeling of looking at a magnified TV. This perception detracts significantly from the sensation of telepresence. See Figure 2.

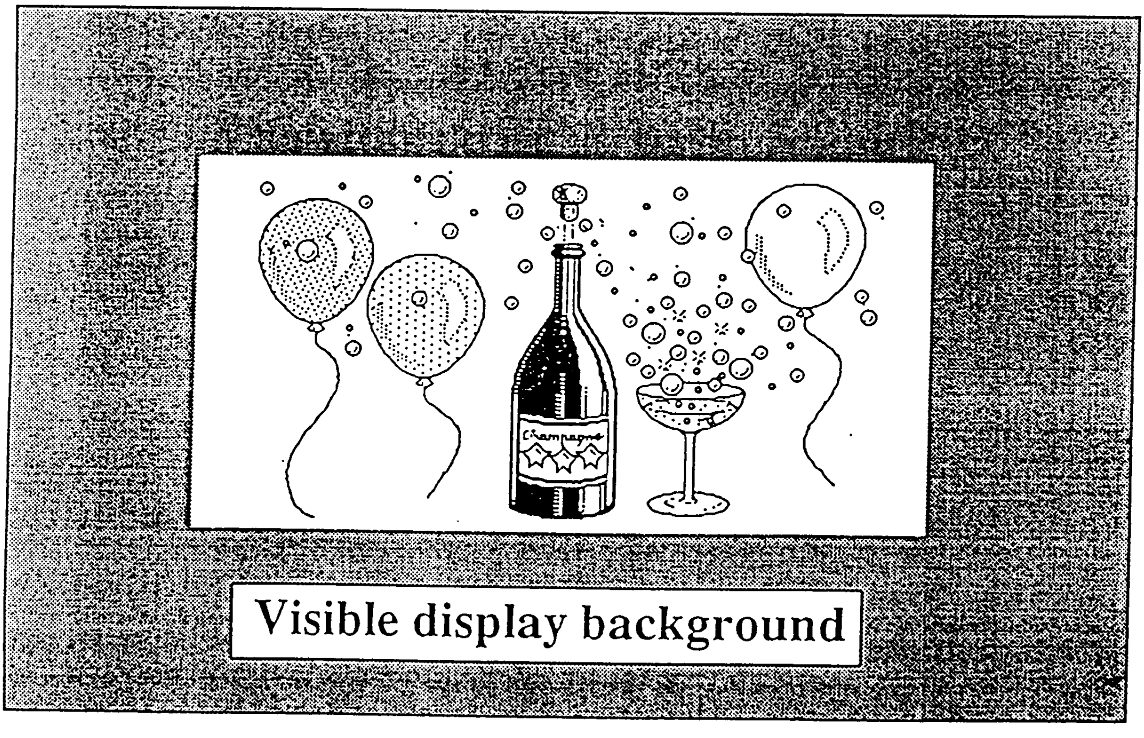
Our display is very high contrast red on black; it is an LED display and the LEDs are either ON or OFF. Because the contrast is so high, the player cannot see a rectangular display background; he sees only bright red objects in black space. The feeling with the solid, continuous 180° black background with clear red objects in the center of one's view is remarkable.

One VR-experienced user described it tersely:

> Looking at a Private Eye virtual display is like looking at red objects in space. Looking at a magnified LCD is like looking at a magnified LCD[3].

Figure 2

No visible display background



Visible display background

Perhaps more than any other single parameter (except cost), the lack of a visible display background is the most important advantage of using a Private Eye display for VR games.

Need for text

Most games have an occasional need for text. Car racing games should indicate speed, fuel remaining, lap time, etc., for good game play. Our simple flight simulator game demonstration has altitude, heading, and speed indicators which are important to experienced players.

Text in games usually is intended to be read quickly and unobtrusively, like the gauges in a car. Any display system which is seriously intended for low-cost VR games must deal with this fact.

Text in the game requires reasonably small, clear pixels. Blurred text is annoying and contributes strongly to a feeling of poor quality. Extremely large text blocks have an acceptable quality feel, but waste screen space. With a Private Eye, text looks like high-quality text--clear, sharp, easy to read, and small.

III. Global economic specifications

Most of these parameters are not related to the basic technology, but are design decisions involved with building a display system.

Field of view

Certainly a very wide field of view (on the order of 90°) contributes to the virtual reality experience. However, a very wide field of view is expensive in two ways: first, the optics are very expensive, and second, individual pixels become quite large and visible (unless resolution is increased to impractical levels).

A narrow field of view has two problems: first, the physical loss of peripheral vision, and second, the perceived loss of peripheral vision which limits one's sensation of telepresence.

Performance with varying fields of view in game-like situations has been an object of study[4,5]. Surprisingly, reasonable performance can be achieved at quite low FOVs--as low as 20°. Improvements in performance seem to taper off after about a 60° FOV is achieved, although anecdotal evidence indicates that there is continuing benefit all the way to 180°[6].

For game play, the difficulty of the game can be matched easily to the capabilities of a human with whatever limitations the system imposes. The real question is the trade-off between wider FOV and the <u>perception</u> of telepresence. This is not as easy to understand or measure.

Although our demonstration system has a FOV of about 10° vertical and 20° horizontal, with 100% overlap, there is a high perception of telepresence. Most users of LCD-based VR systems will find these numbers difficult to believe, especially after using the system. Nonetheless, there is a good sensation of being "inside" the game. Images are crisp and text is easy to read.

For game play, at approximately TV-type resolution, we believe an adequate FOV is no more than 20° vertical and 40° horizontal, about twice the FOV of our demonstration system. With this FOV, pixels will subtend about 5 arc-minutes. This field can be achieved with low-cost injection-molded plastic optics.

We believe that the exceptional contrast and black background of the Private Eye contributes to the reasonably good telepresence at relatively low FOVs.

Color

At a similar price point, and with a similar image quality, color is unquestionably preferred to monochrome. However, with existing technology, color has substantially worse image quality, due to the triad pixel structure required for existing color displays, and is substantially more expensive than monochrome, both in display technology and processing cost.

To our surprise, teenaged boys in focus groups have not been too concerned with color. In focus groups, only about 12% of players mention color as an important improvement, while at least 80% mention improved graphics (the demonstration unit is plain wire frame images). For a home computer game system, color is a desirable feature that can be traded away in return for better performance and an attractive price. As described earlier, the level of enjoyment is not closely correlated with level of similarity to the real world. The virtual world can have its own rules, as long as they are consistent.

There are some relevant comparisons for both image quality and cost. The CCSV[7] uses monochrome CRTs--the researchers found removal of the shadow-mask to be a substantial improvement in perceived image quality. Those who have used the CCSV have often mentioned that the higher quality monochrome screen was <u>more</u> realistic than color LCD VR images, mounted in an identical optical system, because of the improvement in image quality due to clear focus on pixels with no obvious pixel structure.

When cost is important, one should look to the handheld game business. The overwhelming market share leader is the monochrome Game Boy, competing with several more expensive color competitors.

Resolution and gray scales

Our demonstration system resolution is 200 (vertical) x 640 (horizontal) with no gray scales. Unequivocally, adding either higher resolution (we have built displays up to 900x1120 pixels) or gray scales (up to 32 levels are possible) is definitely an improvement.

Unfortunately, neither is free. Display cost increases with increasing resolution and (slightly) with more gray scale capability. Computation and memory cost increase with either parameter.

Our current belief is that 256x512 resolution with 8 gray levels will be optimal. However, additional focus groups are needed to verify the optimal cost/performance features.

Less than 100% overlap

Our current system has 100% binocular overlap; each eye sees the same image. It is simple to lower the overlap to increase the perceived horizontal FOV; some work has been done on the ergonomic results of this scheme in helmet-mounted displays[8].

A variety of problems can arise with lowering the overlap percentage; most are not likely to be serious performance problems in a black background headmounted display. The ergonomic results and perceptions are situation and task dependent. New research needs to be done in the game environment with prototype displays.

If the ergonomic effects are truly unobtrusive, the benefit of decreasing the area of overlap is a corresponding increase in horizontal FOV. The only cost is the increased computation cost for the wider FOV. This question needs further study.

IV. Unit cost

For consumer products, cost is of overwhelming importance. In high-volume manufacturing, a biocular display of 256x512 pixels with 8 gray scales can be built using Private Eye technology for about $35 in material cost.

Discussion of other system parameters

Fast response to head rotation

This is critical. There are three potential delays involved with screen update: position sensor lag, compute-and-draw lag, and display persistence. Our current demonstration system is limited by the position sensor lag, about 15 frames per second, and this is not fast enough.

For players to quickly hunt for objects by looking around and to track them by watching them, fast response is critical. Few things are more annoying than trying to track an enemy helicopter and having to slow your head rotation speed or overshoot your head rotation.

For a good low-cost game system, it is possible to bring the position sensor lag and display persistence times down into the millisecond range at very low cost. Minimizing the compute-and-draw lag costs money and we need to optimize the compute speed more carefully than we have done to date.

We are developing a system in our lab that will have the oldest pixel seen be no more than 100µ (yes, microseconds) out of date. We will use this system as an instrument to measure the optimal trade-off of improved telepresence versus processor cost.

There is a useful body of literature concerning head angle sensing because of the need for helmet mounted sight systems. (eg, [9]). For a game application, most of these systems have not made appropriate cost trade-offs.

Response to head translation

For good game play, there is no need to respond to head translation. Most game play involves distant objects, where head translation is not a significant change in view. Obviously, the system must allow linear motion in virtual space, but need not respond to head linear movement. In addition, for safety reasons it is not desirable for players to be leaving their chairs while wearing displays over each eye.

The most beneficial effect of ignoring head translation is the dramatic reduction in the cost of head position sensing. We are evaluating two different proprietary head angle measuring systems, each of which can be built for about $5.

Finally, there is also a useful reduction in computational burden if head translation is ignored. This reduces the cost of the processor required.

Real time shading

Realtime shading is not justified. Computational costs are very high and stipple patterns are almost as good. Make the response faster and save the processing money!!

Shading algorithms are of much interest to researchers, but there are too many other issues that are much more important to game play to justify the computational expense. It is fine for CAD/CAM, but even expensive arcade games don't generally compute shading today.

At least one worker found that line art (wire-frames) were <u>superior</u> to shaded images in the artificial environment and constrained resolution of VR[1]. This is a controversial finding; we do not believe that simple wire-frames are adequate. Improved graphics are necessary, but stipple patterns and sprites are easier to cost-justify.

This is a clear example of a situation where trying to duplicate the real world is very expensive for little benefit. The virtual world doesn't have to follow the same lighting rules as the real world as long as it is consistent.

Physical ergonomics

Comfort level is important. For extended play, the helmet must be light, not press too hard on the player's head, be easy to use with eyeglasses, and so on. The system must be easy to put on and take off.

Our prototype demonstration system is ergonomically better than Eyephones, but we expect to do substantial additional work in this area to build a really comfortable home game system.

Stereo vision

It is valuable to present an image to each eye, but presenting a different view to each eye is not cost-effective. In the real world, the human brain uses a variety of other cues to judge distance (head movement parallax, foreground/background, size cues, etc.). For objects more than about 20' from the eye, the two eyes actually see the same image.

Most game play involves simulating situations which would be distant from the player, so not much illusion is lost by presenting the same image to each eye. Also, with low or moderate resolution systems, it is not possible to accurately display the true differences between the two viewed scenes. Adding binocular vision adds substantially to system cost, mostly in computation cost, for very little improvement in most game play experience.

Some workers in the field believe that users presented with a different image for each eye in VR systems have significantly more problems with motion sickness

Game controls

We have found a standard computer game joystick to be an excellent controller. The human interface to a joystick is not entirely natural, but it is more natural and much easier to learn than a glove interface. It is much less expensive than most of the other proposed controllers. The players are generally familiar with the joystick and can learn to use it well quickly.

Processor power

Processor power is a function of the programmer's attitude. If the programmers are intending to build a world that is a close technical approximation of the real world, many MIPs are necessary. However, if the programmers come from the world of video or PC-compatible games, they are accustomed to using tools and tricks that optimize game-play at the expense of "realism". These are the programmers that can make great games with low-cost processors.

Because most VR projects have been aimed at low quantity laboratory applications, there has been, quite appropriately, a tendency in the field to use powerful hardware rather than code or perceptual optimization to achieve fast response. For the game application, it is more important to use inexpensive hardware, carefully optimized code, and good understanding of which characteristics give good game play rather than technical realism.

For example, our demonstration system runs on a 386 with no accelerator hardware, and its response is constrained by the response of the head sensor we used in the demonstration system. The software required only minor modifications to an existing PC-compatible game.

Most PC simulation games already have a mechanism to separate "where you look" from "where you go" and calculate the scene from that information. It is necessary only to uncouple the control from the keyboard or controller and connect it to the head angle sensor. Companies that write PC simulation games have excellent development tools for this environment.

Like any good simulation game, a game developed and optimized for low-cost VR should use a mixture of calculated polygons and sprites. For sports games, like ice hockey or basketball, the majority of the game objects should be sprites. For games which are primarily polygons, like our flight simulator, substantial enhancement in game play can be achieved with addition of stipple patterns to the polygons--calculating shading in real time is possible, but requires a much more expensive processor.

There is no reason for the game processor chip with minimal on-chip graphics processing to cost more than a few dollars. Nor is there any need for more than 128KB of RAM; ROM for the game should be built into the cartridge.

Display technology description

Most of the following technical description is excerpted from technical papers taht Reflection Technology personnel have published, eg^{10}.

The Private Eye was developed specifically for virtual displays. It is now in use in a variety of applications, primarily industrial applications where the user has a need to see a display while his hands are occupied. It is the only display technology which was developed specifically for the virtual display environment, and it has better performance for a virtual display than do adaptations of technologies which were developed for full-sized displays.

The basic display mechanism comprises a linear array of several hundred tiny LEDs, a magnifying lens, and a scan mirror (Figure 3). The lens magnifies the linear array of LEDs so that it appears to be a line of pixels of the appropriate height. This represents a single column of the display. The vibrating motion of the scan mirror moves the virtual image horizontally, sweeping out a rectangular raster display area. The LED's are turned on/off in snyc with the mirror motion to create any image desired.

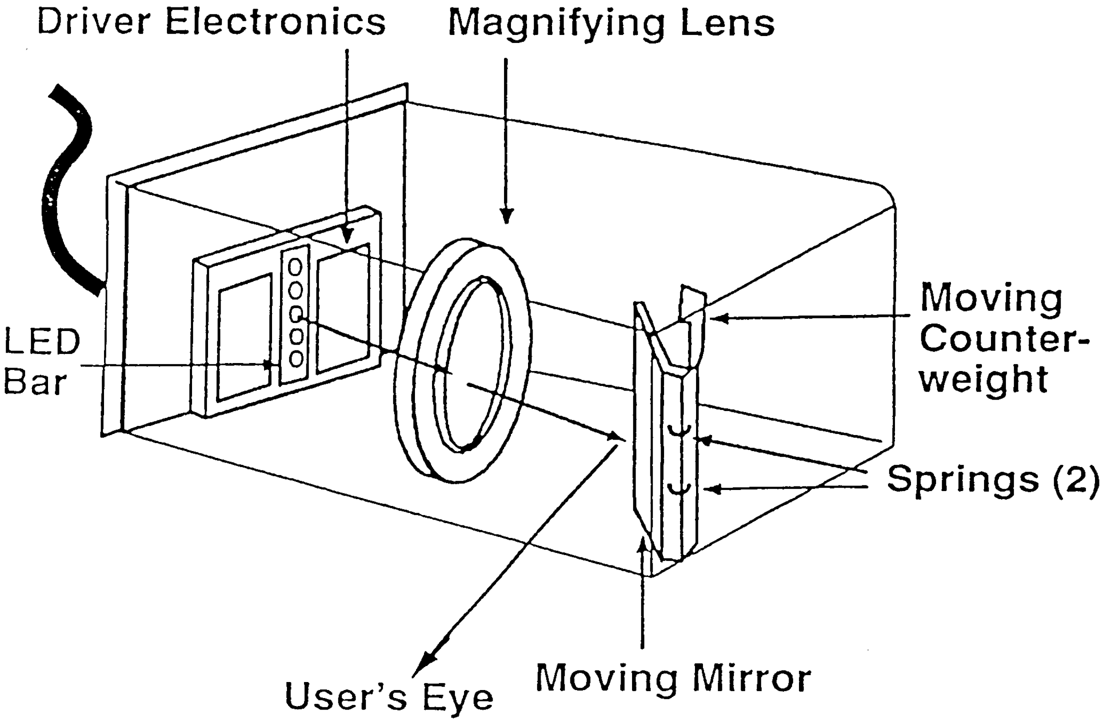


Figure 3

The technology is attractive because it requires only a one-dimensional array of LED's, rather than a two dimensional array as with an AMLCD. Only a few hundred active elements are needed, rather than a few hundred thousand elements. Yield is high, with no special processing. The mechanical scanning system is lightweight, low power, and low cost. A mechanical scanning system is impractical for full-sized displays, but virtual displays are very small physically and mechanically scanning is simple and stable.

There is a disassembled Private Eye display at the Tomen booth for those who wish to see the device in detail.

In addition to being low-cost, the display technology is easily designed to give superb image quality. Almost all first-time viewers of a Private Eye comment favorably on the image quality. There are several reasons for the image quality:

a. High fill ratio

The pixel fill ratio is close to 100%, and can in fact be made to be 100%. The scanning technology allows excellent control over the pixel geometry. Adjacent pixels touch with no visible boundary lines. See Figure 1.

b. Contrast ratio

The contrast ratio is very high. Although it is specified at 70:1 nominal, it is generally much higher; close to that of an LED in the "on" and "off" state. The extreme blackness of an "off" LED makes the raster invisible, enhancing the image of red objects existing in space, not on a screen. See Figure 2.

c. Fast response

The display is fast--pixels are displayed for only a few microseconds, so there is no blurring of pixels due to head rotation or object motion during play.

By comparison, image quality for a magnified LCD virtual display is a serious problem. LCDs are excellent displays for full-sized screens, where individual pixels are 100µ squares or larger (usually much larger). For a virtual display, however, 25µ is a good size for low-resolution displays.

AMLCDs suffer from very poor image quality at this small size. The active elements and the bus bars connecting the elements do not change transparency, and as pixel size shrinks the inactive portion of the pixel becomes objectionably large in comparison.

STN LCDs have no active elements in the display area, but even the gaps between pixels are noticeable and objectionable. The normal STN LCD problems of low contrast and slow response are still present, as well. The low contrast increases the feeling of seeing a magnified TV screen rather than real objects, and the slow response is a serious problem for fast game play and fast head rotation response.

Summary

Much VR work has been done on approximating reality as closely as possible. Little VR work has focused on understanding which approximations are most important to human perception or has studied the questions of which improved approximations add more user value than cost for a specific application.

We have considered these questions for a specific application, the home game machine application. We were able to study this question because we have been in the business of building virtual displays for other applications for several years.

With our low-cost high-quality display, it is possible to build consumer-priced VR games that are fun to play. There are other barriers to building such a game; none are insurmountable, because the important parameters are those of game enjoyment, not realism or technical sophistication.

People who work on cost-is-no-object VR systems don't like some of the techniques we have used to reach a cost goal. But for building a practical application, it is very important to reach a cost goal. It is important to build in only those features that add value as perceived by the end-user.

Our intended end-users are teenaged boys, and they greatly prefer our system over the 16-bit home game machines. A system like ours, built with today's technology, can be cost-competitive with 16-bit home game machines.

Cited references:

1. Bolas, M.T.: *Design for Virtual Worlds - Methods and Experiences.* Master's Thesis, Stanford University, 1989

2. Teitel, Michael A. *The Eyephone, a headmounted stereo display*
SPIE Vol. 1256 Stereoscopic Displays and Applications (1990)

3. Gold, Rich. Unpublished personal observation

4. Wells, M.J., 1, Venturino, M. 2 and Osgood, R.K.1, *The effect of field-of-view size on performance at a simple simulated air-to-air mission*
SPIE Vol. 1116 Helmet-Mounted Displays 1989

5. Wells, M.J., Venturino, M. and Osgood R.K. (1988a) *Using target replacement performance to measure spatial awareness in a helmet-mounted simulator.*
Proceedings of the 32nd Annual Meeting of the Human Factors Society, Anaheim, CA

6. Howland, Eric. Product brochure from LEEP Systems/Pop-optics Lab

7. I.E. McDowall, M. Bolas, S. Pieper, S.S. Fisher, J. Humphries *Implementation and integration of a counterbalanced CRT-based stereoscopic display for interactive viewpoint control in virtual environment applications*
SPIE Vol. 1256 Stereoscopic Displays and Applications (1990)

8. Melzer, James E. and Moffit, Kirk *Partial Binocular-Overlap In Helmet-Mounted Displays*
SPIE Vol. 1117 Display System Optics II (1989)

9. Ferrin, Frank J. *Survey of Helmet Tracking Technologies*
SPIE Vol. 1456 Large-Screen Projection, Avionic and Helmet Mounted Displays (1991)

10. Becker, A. *Design case study: Private Eye*
Information Display Vol. 6, No. 3, (March 1990)