2.3 Current Research and Development

The following research laboratories and small companies are performing research related to visual displays for VE systems. These can generally be divided into two different categories: (1) the development of new types of 3-D displays, and (2) research into how to use these displays, including investigation of human factors issues. By far the most

Specification

Resolution	680 x 480 pixels
Weight	Not applicable
Brightness	72 fL
Contrast Ratio	100:1

Photo courtesy of Dimension Technologies, Inc.

Figure 1. Virtual Window

active research activity at the present time is the attempt to develop autostereoscopic displays. This is because the basic technology for constructing HMDs and shutter glasses is already well known and available as commercial products. The autostereoscopic systems described below typically have multiple applications, including both VE systems and telepresence systems, and more ordinary uses such as 3-D computer displays or television.

The extent to which HMDs and shutter glasses are likely to be replaced by autostereoscopic systems is unclear, and depends upon the particular application and how much of a problem head-mounted gear is for that application, as well as the performance and cost of those autostereoscopic systems that are developed into commercial products. However, it is clear that autostereoscopic systems can be used not only for non-immersive applications, but also immersive applications by using large-screen projection displays, similar to those described in the CAVE project, discussed in Section 2.3.12 below, that is being undertaken by the University of Illinois at Chicago.

2.3.1 ATR Communications System Research Laboratories, Japan

The ATR Communications Systems Research Laboratories are developing virtual space teleconferencing systems in which participants experience a sense of telepresence—in this context, the sense that they are physically in the same environment as other participants they are communicating with, even though they are geographically separate. The researchers are led by Katsuyuki Omura.

One specific objective toward their long-range goal is the development of autostereoscopic display systems. This type of display was chosen in preference to holography and volumetric approaches because of the requirement for full color, a large display size, and imaging in real time. The researchers have chosen an approach using lenticular screens in which a display screen is overlaid with a sheet of tiny lenses so that a given pixel can be seen by only one eye. The primary problem with the basic lenticular approach is that viewers are not free to move their heads laterally beyond a trivial amount. Movement beyond the immediate area results in either no view (because the viewer is seeing the dead space between pixels) or a reversed 3-D effect if the viewer moves further laterally. The ATR group is resolving this problem by tracking the viewer's eye position so as to change the display appropriately when his eye moves into another area. Previous work demonstrated the feasibility of this for a single viewer and a high-density LCD projector.

Recent work has focused on extending the approach to multiple viewers. In this more recent work, the position of each of several viewers is tracked using magnetic sensor or other techniques, and the image projection modified according to the movement. The basic approach is that for every viewer, there is a projector that projects both a left and right image on a screen. The viewers are on one side of a large (100 in) screen, while the projectors are on the other, creating a rear-projection image on the screen. Each projector is mounted on guide rails that allow it to move (driven by motors) a maximum of 0.9 m laterally and from 1.2 to 2.5 m in the front and rear directions: the position of each projector typically mirrors the position of the viewer it projects images for. If the viewer moves to the left (from the viewpoint of the viewer), the projector will also move to the left (from this same viewpoint). If the viewer moves towards the screen, the projector also moves towards the screen. This mechanical arrangement does substantially restrict the mobility of viewers. In the prototype system, only two viewers can be accommodated, because the size of the projectors allows only two projectors, although the researchers envision more viewers in an eventual product. Each projector uses a specially developed projector lens with a wide 65 mm exit pupil, and uses a CRT rather than an LCD because of the CRT superiority in wide angle projection. The projection screen consists of two layers of lenticular lenses, with a diffusion layer between the layers. Tests of the system confirmed that two viewers could in fact see different stereoscopic images. The researchers expect future work to involve developing a system for determining viewer eye positions by the use of video cameras.

As part of developing systems that have high fidelity telepresence, the researchers are studying certain errors in the perception of stereoscopically presented objects. This is motivated not only by the goal of allowing participants to perceive themselves as being present at the same location as other conference participants, but also by the goal of being able to manipulate objects in that environment with the same dexterity that they could if they were physically present. The approach taken is to use a large (709 in) stereoscopic projection system, using the autostereoscopic techniques discussed above, rather than HMDs, which the researchers consider too cumbersome. The focus of the study concerns the visual perception of objects in this space, in which both virtual and real objects exist, and in particular the need for accurate perception of the location of objects to allow a human to reach out with his hand and "grab" a virtual object. This requires the perception of an object location to be consistent with the representation held by the computer controlling the VE, and also requires the perception of the object to be stable when the user's hand approaches it. This can be difficult because in environments in which virtual and real objects are mixed, a variety of phenomena occur, including the following: (1) conflict between accommodation of a viewer's eye at the location of the real or virtual object, as opposed to accommodation at the location of the screen; (2) the tendency for real objects to become transparent when they pass behind a virtual object; (3) the differences in hue, brightness, contrast, and similar characteristics between real and virtual objects; and (4) errors in tracking the user's head and hand positions.

To date two main issues have been focused on. The first concerns the depth error resulting from a mismatch between the assumed inter-pupil distance (IPD) and the actual IPD of a viewer. It is well known that a mismatch between the assumed ("setup") and real IPDs of an observer can result in significant depth error. What has not been previously investigated is the extent to which this mismatch and resulting depth error can be caused by convergence, in which the individual eyes rotate to converge on an object at a given distance. This rotation results in changing of the IPD. The researchers measure the error in depth perception as a function of the position of the (virtual) distance between the displayed object and the screen it is displayed on. A mathematical model predicted that if in fact there was error due to a change in IPD resulting from convergence, there would be a systematic, linearly increasing error of a certain magnitude. In fact, the actual error was more than twice that of the predicted value, suggesting that while changes in the IPD resulting from convergence may be part of the source of error, there are other sources of error. The magnitude of these errors is viewed as unimportant for large-screen applications such as teleconferencing, but potentially quite significant in the case of see-through HMDs.

The second issues that has been examined is the effect of fuzziness of a displayed image on depth perception. In an experiment, the researchers had subjects make judgements about the depth of sharp and blurry images. Results showed that all subjects perceived the blurry image as farther away than the sharp image, though there were large individual differences among subjects, with the effect very small for some subjects.

Current and future research is focusing on the following: (1) determining the other source of depth error (other than convergence); (2) determining what factors are involved in depth error that may interact with blur; and (3) investigating the effect of fuzziness in mixed virtual and real environments.

2.3.2 British Aerospace plc, United Kingdom

Researchers at the Sowerby Research Centre, British Aerospace, plc, are investigating the extent to which viewers can adapt to the unusual accommodation that is typically necessary when using HMDs and heads-up displays. Accommodation is important, in that improper accommodation can result in blurred objects and a failure to detect objects. A variety of factors can influence accommodation. In darkness, accommodation moves to a resting position. If a stimulus is blurred, a reflex drives a change in accommodation in an attempt to resolve the blur. Accommodation also tends to change if there is a change in vergence, the lateral movement of each eye that causes the two eyes to track together. The cognitive knowledge that an object exists that is close to the eye may cause accommodation, and specific conscious mental effort can result in a lapse of accommodation. The experiments being performed by the British Aerospace researchers are intended to resolve how these factors work together when HMDs and HUDs are used.

In an initial series of experiments, the researchers measured resting accommodation in darkness, finding that subjects focused at about 1.4 Diopters, or about 0.7 m. When an optical combiner was placed directly in front of the subject's eyes, the accommodation did not change, even though subjects were aware of the combiner being in front of their eyes. Then an array of hash symbols that formed a sharp, high-contrast pattern was projected as a virtual image on the combiner, collimated such that subjects should accommodate at infinity. This particular image was chosen since it should serve as a high-quality stimulus that triggered an accommodation reflex. Only three of the subjects were able to maintain accommodation at or near infinity, with the other five accommodating at various levels, including as short as 1 m. When the virtual image projected on the combiner was changed to a word, all subjects showed an accommodation substantially closer to the subject than with the previous image, even in the cases where the earlier accommodation was at or near infinity. These experiments suggest that subjects are substantially misaccommodating when virtual imagery is present, particularly when they are mentally processing information in the imagery.

In another experiment, subjects were tested for their ability to accommodate to virtual imagery when it was superimposed on the real world. In this experiment, subjects looked out an open window to see a brick wall and bushes about 28 m away. Three conditions were run. In one, subjects were asked merely to view the scene and to keep the wall and a light fitting on the wall in focus. In the second condition, an array of hashes was superimposed on the outside world with a beam splitter, with the virtual image collimated so that it appeared to be at the same optical distance as the wall. In the third condition, reversed words were presented in the array of hash marks, and subjects were asked to read the words aloud. In the first two conditions, most subjects could maintain accommodation at or near infinity. However, when subjects were required to read the reversed words aloud, thereby mentally processing the information in the virtual image, every subject showed a lapse of accommodation inward, in most cases, one of quite substantial magnitude. There was little difference between accommodation by subjects reading words aloud when they saw a mixed real plus virtual scene than when they saw the virtual simulation. The implications of this experiment are substantial: it suggests that subjects using a HMD will not accommodate to infinity, if intended by designers, even if they have a mixed scene and real-world stimuli to focus on. Rather, their accommodation will lapse, resulting in blurred objects and the potential of failure to detect objects. Such misaccommodation can also result in misperception of the size and distance of an object.

In a follow-up experiment, subjects were provided with an information processing task presented either visually or aurally while viewing a simulated scene, and the shift in accommodation measured. Subjects shifted their accommodation in both cases, but less so when information was presented aurally. This suggests that it might be better to provide a mixture of information to persons viewing HMDs, with a substantial part of it aural, to reduce misaccommodation. However, it is possible that the experimental results could be due to the visual task being more difficult than the auditory analogue.

Other work at British Aerospace includes a study of why blurred images appear sharper when in motion, and studies of the conditions under which the movement of an image on a display is perceived as smooth movement by a viewer rather than jerky movements or multiple images. Still other work includes investigation of voluntary head movements during visual tracking and the resulting slippage of a helmet, and the use of eye pointing as an input media.

2.3.3 BT Laboratories, United Kingdom

Researchers at BT (formerly British Telecom) Laboratories are looking at how useful 3-D might be in video-teleconferencing systems. They see using glasses in teleconferencing applications as very undesirable, “which dramatically reduce eye contract” and which can “make the wearer look doubtful or sinister. Since eye contact with the person at the remote location is one of the key advantages offered by video-telephony and video-conferencing, spectacles based 3-D imaging approaches are not appropriate for these applications” (Jewell et al, 1995). The work of these researchers with 3-D displays also is applicable to telepresence systems and VE applications in which it is undesirable to have the user wear a HMD or shutter glasses.

The researchers are using a system in which they have positioned a lenticular sheet in front of an LCD display, where the sheet is viewed at a distance of 600 mm. (The lenticular display was chosen in preference to a parallax barrier because of the higher luminance throughput.) In such a system, the lenticular sheet consists of a set of columns of tiny lenses, with each column mapping to a pair of columns of pixels, one column of pixels displaying the image for the right eye and the other displaying to the left eye.

Early work looked at necessary properties of the LCD display used with the lenticular sheet. The geometry of the early display resulted in dark areas between pixels that were perceived by the viewer as dark stripes, particularly when the viewer turned his head. An LCD with a very limited color palette (3 bits per color) showed good quality 3-D but rough transitions between shades were particularly distracting in the case of skin tone. A newer display with 200,000 colors has provided good 3-D and much better color, and is currently in use. Preliminary tests have also looked at the transmission of display data over communication lines. Two separate channels were used with independent compression hardware and, though there is a potential problem if the two transmitted views become unsynchronized, overall results were good.

More recent work has concentrated on the development of a head tracking system to allow viewers to move around and still see an appropriate 3-D effect. The general problem with lenticular systems is that when a viewer moves laterally, movement beyond the immediate area results in either no view (because the viewer sees the dead space between pixels) or a reversed 3-D effect with further lateral movement. The BT researchers used a commercially available infrared head tracking system that tracked the lateral position of the viewer's head. They compared moving the lenticular sheet with respect to the LCD display, to correspond to head movements, against rotating the entire assembly of LCD display and lenticular sheet. Of these two approaches, the second proved superior. Since viewers seated in front of a video telephone rarely moved more than 250 mm on either side of a central resting position, and these movements were usually slower than 1.35 m/sec, the necessary rotations can be supplied by a standard stepper motor.

In the long term, these researchers expect that using video cameras with image processing hardware and software that is capable of locating and tracking individuals, together with displays with less dead space between pixels, will generally solve the problem of limited viewing areas for lenticular systems. They would also like to see greater resolution for displays, and image processing systems that can interpolate between source camera views to produce a greater number of intermediate images for viewers to see as they move their head. The goal of present work is extending their basic approach to a higher resolution display. However, they see the most practical future approaches to 3-D video telephony generally as not involving head tracking, but as using high enough resolution displays, together with some additional bandwidth to provide additional perspective views. The combination of additional resolution (which allows more movement laterally by viewers) and additional perspective views would eliminate the need for head tracking.

2.3.4 Canon, Inc., Japan

Researchers at Canon, Inc. are developing techniques for creating a large number of different viewpoint stereoscopic images from a smaller number of 3-D images, or creating 3-D images from 2-D ones by interpolation. The ability to present different perspectives of an image is needed for any type of display that is viewpoint dependent, for example, binocular displays and autostereoscopic displays using lenticular and parallax barrier approaches.

The Canon researchers have developed algorithms for creating interpolations of scenes from a unique perspective between two given perspectives. These interpolations have been experimentally tested by presenting them to viewers using a CRT monitor and shutter glasses. The interpolation method starts by constructing a data structure known as an epipolar-plane image (EPI), in which the separate images from a line of cameras are matched up together into a volume such that the position of the camera is the third dimension. If the cameras are aligned on a straight line, a point in 3-D space is seen on the EPI as a straight line, known as a trace line. All of the possible trace lines that pass a point are identified by searching for trace lines that have similar color values and a slope within a certain expected range. Every pixel that is not already part of a trace line is processed to locate a trace line it is part of. Then the view from a virtual camera is created by determining the value a trace line would have if it passed through the space of the virtual camera position. This step imposes a number of difficulties. One such difficulty occurs when two or more trace lines can intersect at a given point on an EPI, posing the problem of which line to use for an interpolation. In such a case, the line with the least slope is selected to use for interpolation. Another problem is posed by a background region with uniform color that can produce many trace lines and prevent selection of a true trace line, resulting in the background hiding objects in front of it. The algorithms developed can handle these and other problems.

The algorithms have been applied to a number of real scenes, including complex scenes that resulted in considerable motion parallax. In addition to the algorithm for creating a virtual camera position by interpolating between camera positions along a straight line, the researchers have developed an algorithm for reconstructing back-and-forth multiviewpoint images from a set of right-and-left view point images. These are not simply images that one might see by zooming in and out, but perspectives as might be seen by actual physical backward and forward movement.

Future work is expected to focus in two areas: (1) speeding up the interpolation process and making it more robust when operating in a range of different conditions; and (2) extending the algorithms to create VEs that consist of both real and computer-generated imagery.

2.3.5 Dimension Technologies, Inc.

Researchers at Dimension Technologies, Inc., are developing, with funding from the National Aeronautics and Space Administration (NASA), a prototype autostereoscopic display that produces multiple perspective views with full resolution. The overall goal of the work is to develop a device that is not only autostereoscopic, but has a form known as "look around" in which viewers see different perspectives as a result of moving their heads.

Initial work developed proprietary technology that uses the parallax barrier technique for autostereoscopic display. This technique involves placing a barrier between the viewer's eyes and the screen, such as a sheet of opaque material with narrow vertical slits. Such a barrier prevents both eyes from seeing the same column of pixels in the image, and allows presentation of different stereoscopic images to the left and right eyes. Dimension Technologies, Inc. has developed a form of this technology that doesn't actually use a barrier, but narrow vertical light lines that are positioned behind a LCD. The light backlighting the LCD consists of a substantial number of such vertical lines, spaced at equal horizontal intervals. The geometry of the situation is such that there is a vertical line of light for each two columns of pixels, allowing each eye to see only one column. The effect is similar to that of an opaque barrier with a vertical slit, but allows greater brightness and the ability to turn the vertical lines on and off as required.

This basic technology is used in the Virtual Window commercial product described in Section 2.2.18. Its primary limitation is that the reliance on the geometry of parallax creates specific viewing regions for the left and right eyes. If the user moves his head laterally a sufficient distance, he will leave the correct viewing regions and see either no 3-D effect (if at the edges of the regions for a particular eye) or, perhaps worse, an inverted 3-D effect (if the left eye is in a region intended for a right eye, and vice-versa). Therefore, the viewer must remain stationary or, more typically, a system must be provided for tracking the lateral position of the head and modifying the display accordingly. While this solution can work well for a single viewer, the fact that the display can only be adjusted to movement of a single viewer means that it works poorly for multiple viewers.

In the ongoing research, a number of enhancements are being added to the system. First, a very fast LCD was developed allowing individual pixels to be addressed and pixel values modified at a rate several times faster than current off-the-shelf LCD products. A prototype LCD, built by the David Sarnoff Research Center, has an approximate 180 Hz address rate and pixel response times of 0.5 ms off and 3.5 ms on. Second, an illumination system was developed that consisted of 24 fluorescent lamps and an opaque sheet with narrow vertical lines to allow transmission of light to the LCD. Third, a computer control system is used for the lamps that allows each lamp to be turned on and off selectively. Fourth, a sheet of lenticular lenses, consisting of 265 parallel, vertical cylindrical lenslets molded of plastic, and bonded to glass, is placed behind the LCD displaying the image. Fifth, a memory system provides for storing images received at 60 frames per second, so that they can be displayed on the LCD at 180 frames per second.

The resulting prototype display system has so far been able to produce six perspective views at 800 x 400 pixel resolution, with a 32-level grey scale, where each perspective view is visible within a viewing zone 6 cm wide. A viewer thus sees a single view and, if he moves laterally, will see the same perspective until moving into the next viewing zone, in which case a different perspective view is seen. Thus, 3-D images can be seen across a 36 cm viewing area. Brightness is 23 fL and a 15:1 contrast ratio was measured. Ghosting, a form of crosstalk where images intended for zones other than the current one are faintly visible, is relatively low, about 5% as bright as the proper image. The change in perspective as the viewer moves from one zone to another also is relatively low, implying a smooth “look around.” The measured critical flicker frequency is about 36 Hz high, and may be reduced with better equalization of the brightness of the different lamps.

Most of the technologies used in the prototype for electronics, lighting, and optics are proven variations on standard technologies, and Dimension Technologies, Inc. believes that the primary barrier to commercializing the prototype is the availability of higher resolution, higher speed, color LCDs that can be mass produced. They are engaged in discussions with manufacturers towards this end. They also expect that very small LCDs that have recently appeared on the market for use in projection systems can be adapted to autostereoscopic displays using their approach.

2.3.6 Dimensional Media Associates

Dimensional Media Associates is developing autostereoscopic 3-D volumetric displays that could be used in VE systems. Images acquired from objects, CRTs, LCDs, or other media, are projected as full color, solid objects floating in midair. This technology is being sold as a product, the High Definition Volumetric Display (HDVD), for specialized applications such as displays in retail stores where the effect is attention-getting.

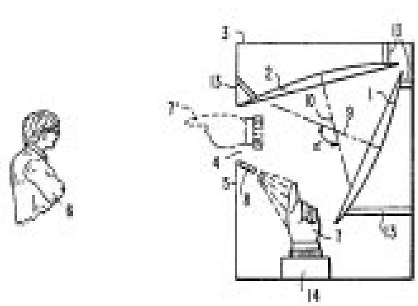

Figure 2. HDVD Overview

The HDVD technology is based on a technique described in a US patent (Summer & Katz, 1994) that uses a pair of large concave mirrors packaged with an image source (an object itself or a 2-D video screen), such that the image source is reflected by first one mirror, then the other, so that the object appears to be floating in space (see Figure 2*).* The image source in the current HDVD system is provided by 2-D display signals, such as those available from video sources or computer simulations. It is, therefore, only capable of projecting monoscopic images. However, the approach is capable of projecting a 3-D image if a volumetric display is used as the image source. Dimensional Media Associates is adapting an acoustic-optical scanner, a volumetric display that uses audio frequency to change the index of diffraction of a lens to diffract the light beam, for use with the concave mirror HDVD technology. The company describes the technology as having potential for application in information kiosks, medical diagnostics, air traffic control, video games, large-scale theme park rides, and data visualization. The company has a Small Business Innovation Research contract from ARPA to develop applications of their display technology to the simulation of surgery.

2.3.7 IBM Thomas J. Watson Research Center and Georgia Institute of Technology

James Lipscomb of the IBM T. J. Watson Research Center, in collaboration with Wayne Wooten of the Graphics, Visualization, and Usability Center at the Georgia Institute of Technology in Atlanta, is researching image processing techniques for reducing crosstalk in shutter glasses. In display systems using shutter glasses, crosstalk occurs when the image presented to one eye is unintentionally seen by the other resulting from leakage through afterglow of the phosphors on the CRT and leakage through the LCD shutter. The first case arises because while the entire LCD shutter turns on or off at one time, the same is not true of the image on the CRT viewed through the shutter. The luminance is turned on by the trace of a scanning electron beam that moves horizontally from left to right and then retraces additional horizontal lines until it reaches the bottom of the screen. It is turned off, effectively, when the light from phosphors on the screen decay sufficiently, in the same time sequence as the tracing beam. In the second case, leakage through the LCD shutter happens when a shutter begins to go black during the vertical blanking interval of the video signal but, when the next field begins to be traced by the electron beam, the shutter has not completely gone black and is still settling down. These characteristics mean that crosstalk varies considerably as a function of the vertical position on the screen. As measured by Lipscomb and Wooten (1994), the leakage through the shutter at the top of the screen (about 7% leakage), continues about 15% of the way down the screen (4.5% leakage). From then on, crosstalk actually increases due to phosphor persistence showing an afterglow from the electron beam tracing, and this crosstalk rises exponentially, reaching a maximum at the bottom of the screen.

Lipscomb and Wooten have developed an algorithm for minimizing crosstalk arising from these causes. First, the image is processed so that the brightness of each pixel ranges from 0.3 to 1, rather than 0 (black) to 1 (white). By reducing the darkest normal intensity level to a dark grey, a certain level of crosstalk can be eliminated since it will not show up against the lighter background. (This strategy incurs the penalties of reducing overall contrast and making images light pastel colors). Crosstalk is further reduced by predicting the amount expected for each eye and subtracting that level from the image. Since crosstalk varies greatly by color, because phosphors that produce some colors have greater persistence than phosphors that produce other colors, this subtraction is color specific. For example, blue and green phosphors have zinc sulphide, resulting in a longer persistence than red, which does not. Consequently, image subtraction is done at one level of magnitude for green and blue, but at a much lower level for red. The algorithm also caters for crosstalk differences based on the vertical dimension of the screen. In this case, the screen is divided into sixteen horizontal bands, with anti-crosstalk measures greatest in the top two and bottom four bands, and weakest in other bands.

2.3.8 Infinity Multimedia

Infinity Multimedia is developing an autostereoscopic display system based on research conducted at the Computation Laboratory and Engineering Department of the University of Cambridge in England. Instead of the more common parallax barrier or lenticular approaches, this display makes use of a patented beam steering system in which images are made visible to only one eye at a time.

The beam steering system allows viewer head movement by using multiple images that are presented via time-division multiplexing. Thus, rather than displaying 60 fields per second, the system displays six different perspective images, using a total of 360 fields per second. This significantly broadens the viewing area that viewers can move around in, allowing the viewer to move forward and backward with respect to the screen as well as laterally. The basic approach uses a CRT, a pair of lenses, and a ferro-electric LCD. The

CRT display produces an image, with the light produced by the display passing through the first lens and then through a narrow horizontal slit displayed on the LCD. The light then continues through a second lens to an eye of the viewer. The slit is in the focal plane of the second lens and this arrangement results in all of the light that forms the image on the CRT display being transmitted from the lens in a single direction, and thus viewable only from a single direction. In particular, the image is viewable by one eye but not the other. Changing the position of the slit (by an appropriate display on the LCD) changes the direction (and thus the eye) from which the image can be viewed. A stereoscopic image can thus be presented by displaying the appropriate image in synchronization with the appropriate slit position for each eye.

Infinity Multimedia presently has a 10 inch diagonal CRT-based proof-of-concept prototype built out of standard components. By spring 1996, the developers expect to have completed an engineering prototype, with a 25 inch diagonal display, using specially designed components. Full-scale production of a commercial product is expected by the end of 1996. Both the engineering prototype and production system are expected to replace the CRT display with LCDs and rear-projection methods. These systems require special high-speed LCDs, and one of their corporate partners, Litton Industries, is contributing the necessary technology, including both lens design and very fast LCDs based on Cadmium Selenide active matrix transistors.

2.3.9 NASA Ames Research Center

Researchers at the NASA Ames Research Center have continued the human factors research in VEs that was initiated in the 1980s with their pioneering efforts in creating the first HMDs. The recent work has had a variety of goals.

One goal of the research is that of developing techniques for calibrating displays for VE systems. Calibration is desired for two reasons. First, it is desirable to ensure that the viewing optics have a field of view that matches the field of view expected by the graphic display system. Second, the optics of HMDs, particularly those with low-cost lenses, introduce distortion. While algorithms exist that can predict the amount of distortion, these algorithms assume an idealized viewing situation and their accuracy in an actual system is unclear. In see-through displays, calibration is not a problem because computer-generated imagery can be superposed on actual physical objects, and the imagery aligned to the physical objects by vernier adjustments. However, in closed systems, some alternative technique is required that can calibrate the imagery and verify that the calibration and registration is correct. Work by Stephen Ellis and Kenneth Nemire has used psychophysical techniques that involve subjective judgement of visual direction. Thus, in an experiment, subjects were asked to point in the direction of either a real object or, in a different condition of the experiment, a virtual object, with both objects appearing as fence posts. Subjects were highly accurate in pointing to the physical object but much less accurate in pointing to the virtual object, a difference the researchers attributed to an incorrect scale factor for the viewing angle in the virtual condition. After modifying the scale factor, accuracy improved considerably. The researchers also attempted to align the virtual and physical environments together with what the subjects perceived as being straight ahead, but found that this did not improve the rotational errors, and concluded that some other factor must be responsible for this error. They are engaged in further experiments in an attempt to isolate the cause of this error.

Another line of research at NASA Ames is a test of whether the addition of a third dimension in a head-slaved telepresence situation (in which an operator is provided with a HMD and a remote camera that moves in a manner slaved to the operator's head movements) can enhance the awareness of a spatial situation. Traditional telepresence systems use two degrees of freedom: pan (azimuth, yaw) and tilt (elevation, pitch). Recently, there has been a trend toward adding "roll," so as to mimic the full capabilities of the human head. The issue here is whether the addition of roll is worth the additional cost and complexity required for its use. Bernard Adelstein and Stephen Ellis carried out experiments that compared the performance of subjects both with and without roll capability. They found that the ability of subjects to accurately determine the azimuth or elevation of an objection was not improved by the roll capability, but that the ability to determine the orientation of an object was improved by a roll capability.

In another line of research, the effect of the display of an object as pitched—that is, rotated up or down from the horizontal plane—on an observer's perception of gravity referenced eye level was investigated. Previous work had shown that when the display of a virtual box was pitched up or down, subjects' judgments of eye level were biased in the direction that the box was pitched. Recent work has shown that the magnitude of this effect depends upon the structure of the object being displayed. In addition, the effect was not as strong with a virtual box as with a real box. The researchers also found that longitudinal structure biased the perception of eye level more than did traverse structure. While the perceptual effect was not as strong with a virtual box as with a real box, only minor additions to the display of the virtual box were necessary to obtain the effect comparable to that of a real box. They found that observers adapted with experience by tending to increase their perceptual bias. Future experimentation may better reveal how observers adapt to the VE.

Other research at NASA Ames focuses on perceptual phenomena that can degrade perceptual performance. In particular, work is ongoing on the source of errors in the perceived distance to virtual objects. Here, the researchers found that superposing a virtual object on a physical backdrop changes its position as judged by an observer. Specifically, if a physical surface is introduced at the depth of the stereoscopic virtual image of an observer, the virtual object is judged to be closer to the observer.

2.3.10 Purdue University

Researchers at the School of Electrical Engineering, Department of Psychological Sciences, and the Biomedical Engineering Center at Purdue University are engaged in experimental studies to determine the extent to which viewing images in stereo can support better visual task performance than viewing images monoscopically. In particular, they are investigating whether the presentation of X-ray images as fused stereo pairs can provide radiologists with depth information that is similar to that obtained with techniques such as computed tomography, but with considerably less radiation dosage and less cost.

The researchers have experimentally tested the ability of subjects to detect a potential tissue abnormality, that may signify a small early cancer, from breast X-rays. In the two experimental conditions, images were displayed on a CRT display while viewers wore LCD shutter glasses. In the stereoscopic condition, the views for the left and right eyes were displayed in alternate video fields. In the monoscopic condition, images were presented side-by-side, so that depth was not perceived. In half of the trials, the task was to decide whether an object of “higher density” was present, while in the other half the subjects were asked to decide whether a particular target arrangement of objects was present. The results of this experiment showed that the subject performance in detecting high density objects was comparable under stereoscopic and monoscopic conditions, but the stereoscopic presentation did increase subject performance in detecting a specific arrangement of objects.

2.3.11 Terumo Corporation, Japan

The Terumo R&D Center, Terumo Corporation, in collaboration with the Department of Radiological Technology, Nagoya University of Medical Technology, Japan, is developing an autostereoscopic display using LCDs. The display system uses an unusual method of head tracking and providing for a parallax barrier. While many variations have been reported, the basic idea is that of having an infrared-sensitive television camera capture the image of viewers illuminated by an infrared light. This image then is displayed on a black-and-white display screen that serves as a backlight to a color LCD that displays the actual image to be viewed. This scheme serves the same purpose as a head tracking system and a parallax barrier system—allowing a viewer to see a particular perspective only from a certain location.

In the first version of the system, a single monochrome display is used with a single TV camera but a pair of infrared lamps, one illuminating viewers from the left side, the other from the right. The stereoscopic signal is time sequential, with alternate fields viewed by the left and right eyes. This is accomplished not by shutter glasses, as is conventional, but by turning on each infrared light according to whether the field is for the left or right eye. If it is for the left eye, the lamp illuminating the left side is turned on. This results in the image of a half-face on the monochrome display, which backlights a color LCD with the actual image to be displayed.

An alternative system is a time-parallel system. This has a pair of displays, each consisting of a black-and-white LCD, a large format convex lens, and a color LCD, with the lens set up so that the black-and-white LCD backlights the color LCD, upon which is displayed the image to be seen. Each of the pair of displays is arranged at right angles to each other, with a half-silvered mirror at a 45° angle such that the two images are combined as seen by the viewer. The viewer is illuminated by a pair of infrared lamps, one from the left and one from the right, as before, but in this case each lamp has an infrared filter in front of it, either 830 - 870 nm or 930 - 970 nm. A pair of television cameras are used, with one of the indicated infrared filters in front of each camera, and the output of each presented to the black-and-white LCD of each display. The image displayed on the black-and-white LCD is a half face image of each viewer. (In the case of the right eye, the right half face).

Still another system uses a similar approach but allows multiple viewers, is thin, and can be mounted on a wall. It is expected that some version of the Terumo Corporation system soon will be introduced in the United States as a commercial product.

2.3.12 University of Illinois at Chicago

For several years, researchers in the Electronic Visualization Laboratory at the University of Illinois at Chicago have been developing a VE and scientific visualization environment in which a person, or group, is surrounded by screens that provide visual displays. The environment is known as the CAVE, a recursive acronym that stands for CAVE Automatic Virtual Environment. Displays are projected on screens positioned at the front, two sides, and floor of a 10 x 10 x 10 ft room. (A sphere would actually be better, so the VE would be seamless, but this is computationally very expensive and beyond the capability of present graphics hardware.) Rear projection is used for the three walls and down-projection for the floor. Each screen uses a separate Silicon Graphics high-end workstation to create the graphics, providing a color display with a resolution of 1280 x 512 pixels. The visual display is at 120 Hz, with alternating fields for different eyes and users wear StereoGraphics CrystalEyes shutter glasses. Multiple speakers provide 3-D sound, and users wear electromagnetic sensors that track head and hand movements.

There are a number to advantages of the CAVE over HMDs, including the ability of users to easily see others in the same room, some physiological vision effects such as the ability to see objects appropriately in focus or out of focus, and significantly less occurrence of motion sickness: the researchers report that only two of the 9000 people who have visited the CAVE complained of nausea. The CAVE display also is claimed to provide a more accurate display than HMDs, which have optics that create geometrical distortion. In designing the CAVE the researchers had to confront a number of problems, including the difficulty in displaying green stereoscopically by projection (caused by especially long persistence of phosphors in projection equipment, and solved by a specialized tube), minimizing user shadows when projecting downward onto the floor (shadows cannot be eliminated unless the floor has projection from below, but projection from the top offset to the front minimizes shadows), the use of a shared memory arrangement among workstations to synchronize frames on different screens so as to avoid a problem in which “images in the corners crease and start to look sucked in like sofa cushions.”

The researchers analyzed the HMD, monitor, and CAVE situations with respect to the results of tracking errors that are errors of displacement or errors of rotation. Neither monitors nor the CAVE are sensitive to rotation error, since the image display plane does not move with the position and angle of the viewer. In the case of HMDs, rotational errors can be serious. Displacement errors for the CAVE and monitor suggest that for small distances between the viewer and the display screen, there is little difference in effect between the CAVE and monitor. For large distances, the angular error is less for the CAVE because of the typically larger distance between viewer and display. For small distances (for example, 20 cm) the monitor has the best performance, and the CAVE and HMD and also BOOM display have slightly worse performance. For large distances (for example, 500 cm), the HMD and BOOM have very good performance, while the CAVE has 2.5 times the error of the HMD and BOOM, and the monitor is worst at 9 times the error.

The CAVE does have some shortcomings. One is cost: the CAVE is large and expensive, although in some applications the cost per person might be more reasonable because it can be shared by a group of users.

2.3.13 University of New Brunswick, Canada

Researchers at the University of New Brunswick are studying the relationship between stereoscopic vision and visual environments responsive to head movements. As previously mentioned, they have experimentally tested the effect of stereo viewing and head coupling on viewers (Ware, Arthur, & Booth, 1993). The display system developed for this work uses an ordinary CRT display and StereoGraphics CrystalEyes shutter glasses. Head tracking was provided by a mechanically linked head tracker, which allowed coupling of the display image to user head movements. There are several advantages to this approach claimed over HMDs. First, as a result of both increased resolution of the monitor display over typical HMD resolution and a decision to reduce the field of view to 30° laterally, the resolution is 2 minutes of arc per pixel rather than the 12 minutes of arc per pixel typical for a HMD. Second, the monitor approach allows presentation of the 3-D images at a depth of field that the viewer can focus on, with images behind or in front of this depth out of focus, which is normal. This is in contrast to HMDs, in which the optics are typically arranged so as to force the viewer to focus at infinity. Third, an error introduced in HMDs when the eyes move is greatly reduced with the head-coupled shutter glass system because the eyes are further from the display. Finally, the monitor is part of the ordinary workspace, allowing (if desired) the user to see the normal desks, tables, and chairs rather than a completely synthetic environment.

The researchers performed two experiments. The first experiment looked at the subjective impression of 3-D images. Here two simple scenes were used that had strong depth cues even when seen without stereopsis. Subjects viewed the scenes in a number of conditions, including with and without stereo vision and with and without head tracking. Subjects were then asked which condition showed the strongest 3-D effect. Perhaps surprisingly, a non-stereo image with head coupling responsive to movement was typically perceived as having the strongest 3-D effect. A second experiment measured performance on a task in which visual depth was necessary and which used the same conditions as the first experiment. Two complex trees were presented in 3-D, with overlapping branches and one leaf tagged. The task required the subject to determine which root the leaf could be traced to. Results indicated that subjects made a much large number of errors for the stereo only condition (14.7%), compared with the head-coupled monocular condition (3.7%). The condition with both stereo and head-coupled display resulted in the lowest number of errors (1.3%).

In more recent work, the researchers have developed an algorithm that dynamically adjusts the apparent distance between the eyes for a computer-generated, stereoscopic image displayed on shutter glasses (Ware, Gobrecht, & Paton, 1995). The primary motivation for this work was the fact that different scenes appear to be best displayed by use of an eye separation that is different from the "correct" or expected one, given the apparent distance between the viewer and the object viewed. This assertion was tested experimentally by having viewers set their preferred separation for different moving scenes.

2.3.14 University of Washington

An intriguing alternative to the conventional CRT or LCD display is the use of a laser to "directly write" onto the retina. The motivation for attempting this is to increase the resolution and field of view of displays, and to eliminate display screens and heavy imaging optics and thus create a low profile display. The Human Interface Technology Laboratory (HITL) at the University of Washington is developing a prototype retinal display as a long-term project funded by Micro Vision, Inc., of Seattle, Washington, which hopes to manufacture and distribute the device as a product. The project has a series of long-range project goals that are very ambitious. The goals are to create a device that (1) is small and lightweight enough to mount on eyeglasses; (2) has resolution high enough to approach that of human vision; (3) has a large field of view, greater than 100° per eye; (4) has color resolution superior to standard displays; (5) is capable of displaying either in a dedicated or see-through mode; (6) is bright enough for outdoor use; (7) has very low power consumption; and (8) provides a true stereoscopic display.

To meet the acuity of the eye (about 1 minute, that is 0.016° of arc) while also surrounding the user to the maximum extent (based on a single-eye field of view of 135° horizontally and 150° vertically) would require a display with a resolution of about 8000 by 8000 pixels, far beyond the capabilities of today's displays (Holmgren & Robinett, 1993).

In principle, a direct retinal write, or scanned laser, display uses a laser source along with accompanying mirrors or other deflectors to direct a laser beam through the pupil of the eye and onto the retina with a scanning pattern similar to the scanning used in conventional television. Scanning a laser onto a retina has been previously used in scanning laser ophthalmoscopes. The feasibility of building a practical retinal scanner has yet to be demonstrated, and early prototype efforts by the HITL were criticized because of their use of acousto-optic scanners, which would be very cumbersome in a color version because of the need for six different sets of deflectors and lenses (one for the x axis and one for the y axis for each primary color). The primary problem posed is the difficulty of building a scanner that can deflect the beam horizontally.

The HITL/Micro Vision researchers now believe that a practical retinal scanner can be developed using a different approach than that used in their early prototypes. They have developed (and applied for a patent on) a mechanical resonant scanner that has only a single moving part and has a large scan angle and a high scanning frequency. The device is also quite small (0.9 x 1.3 x 2.8 cm), has a uniform and repeatable scan, reflects all colors at the same angle, and is made from common materials at a volume manufacturing cost estimated to be under $3. A bench-mounted prototype, using the mechanical resonant scanner to scan the beam horizontally, has been developed that provides VGA resolution images (640 points horizontal, 525 lines vertical) in either monochrome or full (RGB) color. The monochrome system uses a red laser diode, while the color system additionally uses green (helium neon) and blue (argon) gas lasers.

One problem with the present scanner is that the beam moves faster at the center of the scan than at the edges, which results in pixels that are wider and brighter in the center than near the edges. This can be corrected by varying the pixel display time and intensity as the beam scans across the image. A second scanner problem is a change in resonant frequency with temperature, which researchers expect to cure with a feedback mechanism that compensates appropriately for temperature variation. The primary disadvantages of the color display are its size and cost, primarily because of the blue and green gas laser sources, since diodes for these colors are not currently available. Work is being done on frequency doubling methods and the use of non-laser sources, including the development of blue and green LEDs.

Future work will concentrate on further development of the mechanical resonant scanner, resolving a problem of the exit pupil size (which is currently quite small), developing methods of generating color with hardware that can fit in a small package, increasing the resolution of the display, and testing for safety. The researchers expect that safety will not be an issue due to the low power of the light sources, but are planning to begin an extensive series of safety tests by an ophthalmologist.

Micro Vision, Inc. intends to enter the market with monochrome displays for applications that need a compact display with minimal power draw and a bright, medium-to-high resolution image. They expect to begin with a small hand-held prototype system that can be used to develop and test applications.

2.3.15 Xenotech, Australia

Researchers at Xenotech in Australia also are working on techniques for 3-D autostereoscopic display. The goal of the effort is to develop systems that have high image quality and that do not require viewing aids such as shutter glasses. They have built a number of prototype systems, with the design of most of them still confidential. However, they demonstrated an advanced prototype in October, 1995 at the Korean Electronics Show in Seoul, Korea.

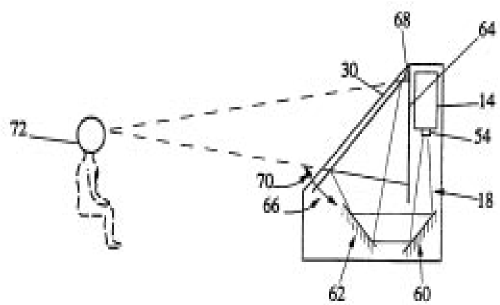

Figure 3. Xenotech Autostereoscopic Display Overview

The prototype system uses a 30 inch diagonal television monitor and a single viewer sits at a distance of about 1 m in front of the monitor. Using a pair of projectors and a series of mirrors, and a screen with a specialized material, the device presents a different image to each eye. The basic technique is shown in Figure 3, which is from a patent application filed by Xenotech (Richards, 1994). A pair of video projectors are contained in a housing, one for each eye, with the projector for the right eye (14) shown and the projector for the left eye not visible but behind the other projector. The image for the right eye (18) is projected onto two successive mirrors (60 and 62), and then onto a partially silvered mirror (not labeled but that swings on a hinge 68 between the arrows 66 and 70). The image is then reflected off of the partially silvered mirror to a “retroreflective” surface (64) that acts as a mirror, reflecting the image back toward the viewer (72), but passing first through the partially silvered mirror. This image is focused by optics consisting of the partially silvered mirror and the retroreflecting mirror specifically to the right eye of the viewer. A second image is projected in the same way by the second projector, through the same series of mirrors and optics. The retroreflective mirror consists of a special surface with a zigzag pattern at the pixel level, which has the property that an incident beam of light is reflected back at an angle 180° from the incident angle. For example, a beam of light that comes in at a 90° angle to the surface of the retroreflective mirror will be reflected back at a 90° angle, and is thus seen only by one eye and not the other.

A pair of television cameras at the left and right sides of the monitor track the head position and pupil locations of the single viewer. When a viewer moves, the image projected to each eye moves to compensate, either by adjusting the angle of the partially silvered mirror (for movement in the vertical direction) or by lateral movement of a carriage upon which both projectors are mounted (for movements in the horizontal direction).

Xenotech sees the primary advantage of their systems as their use of field sequential video at 50 (or 60 in NTSC) frames per second for each eye to eliminate the flicker that is a problem in competing systems that present video at 25 (or 30) frames per second. Presentation of images at this frame rate is achieved by holding each frame in a buffer memory and presenting it twice. The prototype system operates at standard NTSC or PAL television resolution, but the approach is applicable to much higher resolution. Other claimed advantages are very large screen sizes (the researchers have built a prototype as large as 50 inches diagonal), and very high image brightness. Xenotech has two specific markets in mind for development, these being military applications and the video games market. Current work includes the development of a prototype autostereoscopic display that will allow multiple viewers of the monitor.