PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

Real-Depth imaging: a new 3D imaging technology with inexpensive direct-view (no glasses) video and other applications

Eugene Dolgoff

Real-Depth™ imaging: a new 3-D imaging technology with inexpensive direct-view (no glasses) video and other applications

Gene Dolgoff

Floating Images, Inc.
95 Post Avenue, Westbury, NY 11590

ABSTRACT

Floating Images, Inc. has developed the software and hardware for a new, patent pending, "floating 3-D, off-the-screen-experience" display technology. This technology has the potential to become the next standard for home and arcade video games, computers, corporate presentations, Internet/Intranet viewing, and television.

Current "3-D Graphics" technologies are actually flat on screen. Floating Images™ technology actually produce images at different depths from any display, such as CRT and LCD, for television, computer, projection, and other formats. In addition, unlike stereoscopic 3-D imaging, no glasses, headgear, or other viewing aids are used. And, unlike current autostereoscopic imaging technologies, there is virtually no restriction on where viewers can sit to view the images, with no "bad" or "dead" zones, flipping, or pseudoscopy.

In addition to providing traditional depth cues such as perspective and background image occlusion, the new technology also provides both horizontal and vertical binocular parallax (the ability to look around foreground objects to see previously hidden background objects, with each eye seeing a different view at all times) and accommodation (the need to re-focus one's eyes when shifting attention from a near object to a distant object) which coincides with convergence (the need to re-aim one's eyes when shifting attention from a near object to a distant object). Since accommodation coincides with convergence, viewing these images doesn't produce headaches, fatigue, or eye-strain, regardless of how long they are viewed (unlike stereoscopic and autostereoscopic displays).

The imagery (video or computer generated) must either be formatted for the Floating Images™ platform when written, or existing software can be re-formatted without much difficulty.

Keywords: 3-D imaging, 3-D video, 3-D projection

1. BACKGROUND

Three-Dimensional Imaging Techniques[1] by Takanori Okoshi (1976) analyzes virtually every 3-D imaging method ever devised to that date. Since that time, most 3-D imaging technologies that have been developed use variations of the technologies disclosed in Okoshi's book, and no radically new approaches have been proposed. A good summary of the history of 3-D imaging was recently published in the *SPIE Proceedings.*[2] When discussing the prospects for 3-D television, Okoshi calculated that the bandwidth required to transmit a Lenticular-Sheet 3-D Image, which he asserts is the lowest bandwidth required for an autostereoscopic 3-D picture, would be 750 MHz. This corresponds to about 125 of today's conventional TV channels.

Okoshi's analysis shows that neither he nor anyone else knew how the bandwidth requirement could be reduced enough to transmit 3-D video, nor did anyone imagine what kind of display device could be devised to show it. Okoshi predicted that the next generation of television broadcasting would feature highresolution, wide-screen display that gave only an "illusion" of depth sensation. To support his prediction, Okoshi cited the beginning of an "epoch" in movies when Cinerama was introduced. Cinerama was a two-dimensional (2-D), curved, wide-screen technique. According to Okoshi, the popularity of Cinerama resulted in a dramatic decrease in efforts to develop other forms of 3-D.

The popularity of Cinerama, as mentioned above, is especially interesting because it was only a 2-D image projected on a very wide curved screen; yet it produced a realistic depth-containing experience. The effect indicates that something about the display was providing the brain with depth information.

The depth cues present in a Cinerama display were very important and compelling. First of all, a variety of 2-D depth cues were present. As an object gets farther away, it gets smaller, higher up in the frame, less distinct, less color saturated, less contrasty and less bright. Object points get closer together as they recede, such as train tracks appearing to get closer as they get further away, and foreground objects obscure background objects. Second, the objects depicted on screen were often small compared to the huge screen size. Third, the extremely wide screen necessitated that the viewers focus their eyes differently, the process known as accommodation, when viewing objects at the center of the screen as compared to when they viewed objects on the sides of the screen. This effect was more pronounced the closer one sat to the screen and the smaller each on-screen object was. If an object was nearly as big as the screen, however, the brain would reduce the viewer's perception of apparent depth. This is because different parts of such an object appeared at different depths, depending on where each part was on the wide screen. The brain automatically lost depth perception with regard to that object and just saw it as a curved flat object.

With Cinerama, although depth appeared limited, the appearance of depth was striking because accommodation and convergence seemed to match when looking at different parts of the big screen. The eyes had to both converge differently and refocus.

Further analysis of this observation led to the development of the technology outlined in this paper. This new technology is better understood when considering the following analysis of 3-D perception.

The experience of 3-D is created in the presence of four conditions. The first condition consists of what are collectively called **2-D cues**. These cues mainly consist of objects getting smaller, higher, closer together, dimmer, less distinct, less contrasty and less colored as they get further away, as well as foreground objects blocking the view of background objects. These cues are recorded and reproduced in the course of standard 2-D image recording.

The second condition is **parallax.** Parallax occurs when a change of position of the viewer produces a different view in which background objects previously hidden by foreground objects become visible. Conventional 3-D stereo techniques lack the ability to convey parallax. Although parallax is not absolutely essential for a viewer to experience 3-D, its presence adds a great deal of realism to 3-D imaging.

The third condition is **lateral binocular disparity.** This means that the lateral (horizontal) relationship between at least two objects in the scene is different for each eye. This results in different amounts of convergence of the eyes to form a single perceived image of each different object in the scene. It can be reproduced by recording images of a scene from two (or more) different points of view.

The fourth condition is **depth disparity,** in which at least two objects in the scene are not in focus to the eye at the same time. Thus, **accommodation** or re-focusing of the eye, is required when shifting attention from one such object to another. This phenomenon, present in real life, is not reproduced with conventional 3-D imaging techniques. It requires focusing of at least two different components of a scene into at least two different depth locations in space.

The author has found that this condition is very important because, as the brain acts to refocus the eyes from one depth to the other, it experiences the perception of true depth. Lack of the depth disparity phenomenon leads to eye strain and headaches after an extended period of viewing in current 3-D imaging systems, because of the conflict caused between accommodation and convergence. Preferably these four conditions correlate to one another when viewing a 3-D image so as to provide the same relationships found when viewing a real life scene.

In real life, accommodation is an extremely important, but almost completely ignored, part of what causes depth perception. The eyes constantly refocus on nearer and farther objects in a scene. When focusing on objects at one distance, objects at another distance are seen as being out of focus, and the brain adjusts lens and corneal muscles to bring into sharpest focus whatever one concentrates on. Due to the eyes' limited depth of field, one can never get all objects (or even all parts of any single three-dimensional object) into best focus at any one time. As attention is shifted to blurrier objects in an attempt to sharpen their image for clearer recognition, the eyes keep refocusing. Objects keep shifting from clear to blurry, creating a scene in constant flux made of a mix of sharper and blurrier images which keep changing their focus. This effect is even observable with one eye, creating the basis for limited "monocular depth perception."

While observing a scene, the brain also constantly shifts attention from nearer objects to farther objects in the attempt to merge all perceived double images. When viewing an object at a selected depth, the two eyes are aimed at that object so that the two views of that object overlap precisely, creating a single image in the brain. This is called **convergence**. At the same time, other objects at other depths do not line up and therefore appear to the brain as double images. As the brain constantly shifts attention among nearer and farther objects, it experiences a continuing flurry of single and doubled images.

Through life experience, the brain forms a correlation between each degree of accommodation and each degree of convergence in response to viewing objects at different depths.

Almost all 3-D imaging systems in use today utilize stereoscopy or autostereoscopy. In stereoscopy, two images are recorded, one corresponding to the left-eye view of the scene and the other corresponding to the right-eye view. These two images are different, providing what is called "binocular disparity." This difference forces the viewer's eyes to aim at objects at each selected depth to see them properly. Each eye is made to view only its corresponding image through the use of an optical device such as red and green glasses, polarized glasses, or lenses which focus one image into each eye, such as in a stereoscope.

Autostereoscopy directs the corresponding images to the eye without the use of any optical device near the person. Instead, optics are located near the images, restricting the angle of view of each image so that each eye still sees only one of the images. This has been done with lenses, prisms, and light-blocking barriers, for instance. Since the angle of view for each image can be made very narrow, many images can be taken from many angles and viewed one at a time as one moves their head. This provides an aspect of 3-D perception, called **motion parallax**, not available from stereoscopy. With motion parallax, one can look around foreground objects to see previously hidden background objects. The images displayed using stereoscopy and autostereoscopy, however, are all in one plane, so the constant refocusing and perception of a mix of sharp and blurry images - which the author has found to be so important to the real 3-D experience - is absent. Due to this lack of variable accommodation cues, stereoscopy and autostereoscopy present a difference from reality that is significant.

In stereoscopy and autostereoscopy, as the object gets farther from the plane of the image, in front or behind, the convergence of the eyes increases, but unlike reality, the accommodation stays the same (since all image information is in one plane). The farther away from the image plane an object is, the larger is the discrepancy between accommodation and convergence. The discrepancy causes the brain to change the convergence and accommodation of the eyes back and forth in an attempt to create a match between them based on past experience.

In such an experience, fatigue, eyestrain, and headaches result since the objects are, at least in part, not really in focus in the same plane as convergence makes them appear. Also, the further out of the plane of best focus the image appears due to convergence, the harder it is for the viewer to see a 3-D image instead of a double image.

2. REAL-DEPTH™ 3-D IMAGING

The new technology outlined in this paper relates to 3-D imaging for television, computer monitors, and projection for numerous applications including home and arcade video games, animated features for the VCR, CAD, military simulation and display, corporate presentations, Internet/Intranet imaging, and consumer television, both standard and high-definition.. Numerous attempts have been made over several decades to devise a practical system for 3-D display, including full color moving electronic video in particular. To date, no satisfactory methodology has surfaced that produces affordable 3-D imaging of acceptable quality. Additionally, no method has been devised for transmission of 3-D images over conventional bandwidths.

The new technology provides what the author calls "Real-Depth"™ imaging, since there are real differences in depth between images of objects in a scene. Different objects at different depths in a scene are presented in different planes in space, one behind the other. The object areas of each plane are perceived as opaque, and the non-object areas are clear to allow observation of images in other planes.

Experimentation to determine the minimum number of different planes needed to create a satisfying "Real-Depth"™ experience yielded the surprising answer that only two planes are needed. This simplifies the necessary system, and reduces the information content to a level which can be transmitted on a conventional TV channel or over the Internet.

Consequently, with Real-Depth™ imaging, a scene is divided into two sections; a foreground and a background, one placed on each plane, utilizing 2-D cues, including perspective and background object occlusion in each plane. The eyes continually shift attention from plane to plane, changing accommodation and convergence. As in real life, the scene always appears to be a mix of sharper and blurrier images and single and doubled images which keep shifting. If one changes one's viewing position, one perceives both horizontal and vertical parallax. At all times, in all viewing positions, the two eyes see different perspective views, creating binocular disparity to further stimulate the brain into experiencing a real-life depth experience. Most important, accommodation always matches convergence, creating a painless experience of depth. Surprisingly, a scene displayed this way with perspective cues appears to have continuous depth and not be confined to only two planes, whether it is moving or still.

The foreground image consists of selected objects on a black surround. This black area appears as a blank or transparent space through which the background image is seen. The foreground image is brighter than the background image, hence, the foreground image appears solid, appearing to obscure or block the background image as it would if it really were solid.

For example, one could display a foreground image of a standing couple and a background image of a brick wall. The foreground image consists of the standing couple on a black surround. The background image consists of the brick wall. This is depicted in figures 1A and 1B. Because the virtual image of the background appears behind the image of the foreground, and because the foreground image is brighter than the background image, the image of the couple obscures a portion of the brick wall.

The foreground and background images appear in different planes in space. This creates parallax, in which the change of position of the viewer produces a different view in which background objects previously hidden by foreground objects become visible. Different bricks would be exposed or hidden as the viewer bobs horizontally or vertically. It also provides convergence and accommodation as the eyes aim at and change focus between the bricks and the people. This is represented in figure 2.

One embodiment of the Real-Depth™ 3-D imaging system uses two image sources (one as a foreground image source, and another as a background image source) with a partially reflective mirror disposed at an angle so that a virtual background image (the reflection of the background image source) appears behind the foreground image. The image sources can be any image-producing sources, still or moving, including transparencies, CRT or LCD.

A housing having a viewing aperture, a foreground image source aperture and a background image source aperture is desirably provided. Inside the housing is a partially reflective mirror. This embodiment provides line-of-sight from the viewing aperture through the partially reflective mirror to the foreground image source. The partially reflective mirror is disposed at an appropriate angle so that a virtual background image (the reflection of the background image source) appears behind the foreground image. The observable distance between the two images is created for instance, by placing the two image sources at two different distances from the partially reflective mirror. An example of this is shown in figure 3.

To completely eliminate visibility of background image information bleeding through foreground image information, a light valve, such as an LCD, is placed between the background image source and the partially reflective mirror at the same distance that the foreground image source is from the partially reflective mirror. This is shown in figure 4. Opaque images corresponding to the images displayed in the foreground plane are displayed on this light valve preventing bleed through of background image information.

To increase the appearance of connectedness between the foreground and background images, a third "floor-plane" image can be created with another image source placed along the "floor" between the two other image sources. A simpler method is to use a taller background image source and tilt the top of it backwards away from the foreground image source. This is shown in figure 5. Since, in a normal view, as objects get farther away, they get higher in a picture, the fact that higher points in the background picture actually are farther away will greatly enhance the illusion of depth, requiring the eyes to re-focus and accommodate when looking at areas in the background that appear at different depths.

A simplified optical hardware adapter using one image source, such as a CRT, is shown in figure 6. In this design, the top half of the screen displays the foreground image. The bottom half of the screen displays the background image. The combination of a mirror and a partially reflecting mirror creates the appearance of the background image being located co-axially a distance behind the foreground image, creating the Real-Depth™ illusion.

A thinner, more streamlined optical hardware adapter can be made from a single sandwich of molded micro-optics-containing plastic sheets. This plastic optical sandwich would be slightly larger than a single image source, such as a CRT, and would hang a few inches in front of it. This is shown in figure 7. This micro-optics adapter performs nearly the same function as the adapter shown in figure 6 which uses mirrors. Again, the image source would display the foreground image on the top half and the background image on the bottom half of the image source. The micro-optic adapter comprises three elements.

One element is a beam-combining element which can be either a Fresnel beam-combiner shown in figure 8A or a Fresnel semi-prism, which is shown in figure 8B. The function of this element is to bring the foreground and background images into co-axial alignment. The Fresnel beam-combiner bends light from the top and bottom images toward the viewer, creating the appearance that the two images exist simultaneously in the central area of the screen. The Fresnel semi-prism bends light from the top image while letting light from the bottom image pass straight through. This creates the appearance that the two images exist simultaneously at the bottom half of the screen.

The second element in this micro-optic adapter is a Fresnel semi-lens. This is shown in figure 9. The lens portions of this Fresnel semi-lens are the same size as, and coincide with, the areas of the Fresnel beam-combiner or the Fresnel semi-prism that send the background image to the viewer. The lens acts to create an enlarged viatual image of the background image, some distance behind the foreground image. This creates the Real-Depth™ illusion.

The third element is a Fresnel cylinder lens whose cylinders are oriented horizontally. This lens acts to expand the image in the vertical direction, creating a magnification of 2:1. Thus, a full-sized image, with the same aspect ratio of the full image source, is displayed. Both the foreground and background images must be compressed 2:1 in the vertical direction so that enlargement by this lens produces undistorted final images.

Alternately, a holographic optical element (HOE) may be constructed with equivalent functionality of the micro-optic adapter to accomplish the same function.

Dividing an image source, such as a CRT into four sections, combined with two mirrors, two partially reflecting mirrors, and two magnifying mirrors provides a Real-Depth™ 3-D display utilizing one image source which can be viewed by two independent observers depicting two different 3-D views or images. This is shown in figures 10A and 10B.

3. CONCLUSION

This new technology makes possible, for the first time, 3-D video games, without the need for glasses and without deleterious effects from long-term viewing, at a very low, consumer-affordable price. It also enables the production of Real-Depth™ animated VCR tapes and other imaging applications. Since this information can be transmitted with standard bandwidth, this technology also provides a viable method for transmission of 3-D images over the Internet, teleconferencing lines, and standard broadcast and high definition TV transmission.

REFERENCES

1. T. Okoshi, *Three-Dimensional Imaging Techniques*, Academic Press, San Francisco, 1976.
2. M. T. Bolas, S. S. Fisher and J. O. Merritt, *Stereoscopic Displays and Virtual Reality Systems III*, pp. 4-8, SPIE - The International Society for Optical Engineering, San Jose, 1996.

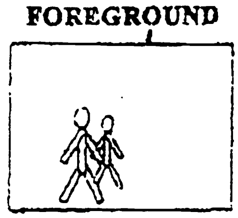

Fig. 1A

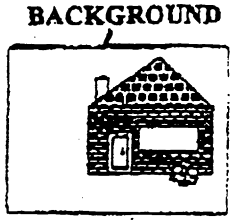

Fig. 1B

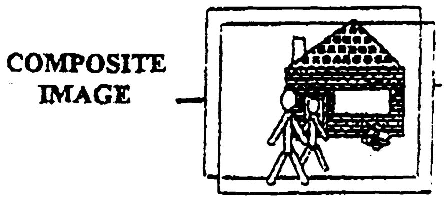

Fig. 2

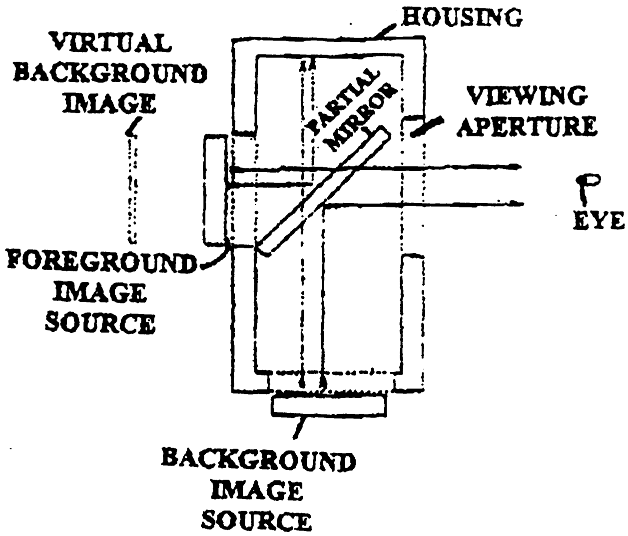

Fig. 3

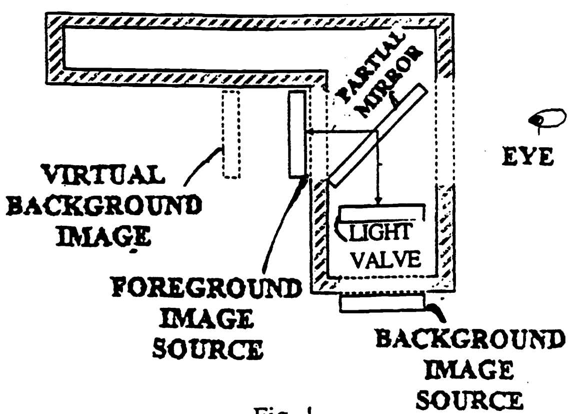

Fig. 4

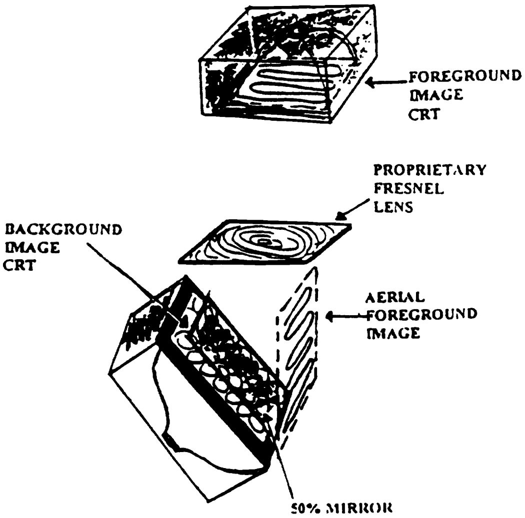

Fig. 5

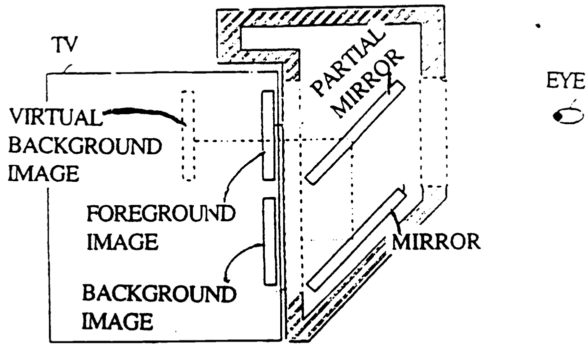

Fig. 6

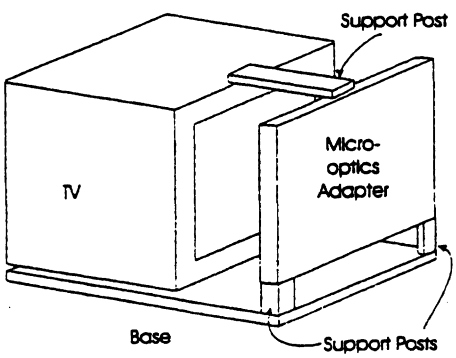

Fig. 7

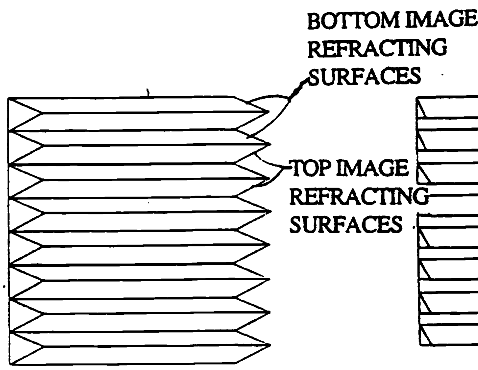

Fig. 8A

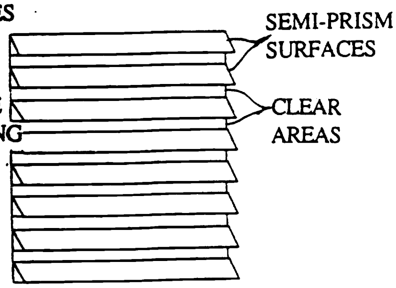

Fig. 8B

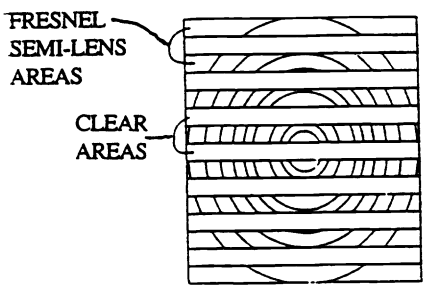

Fig. 9

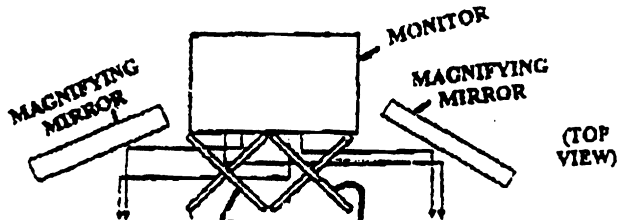

Fig. 10A

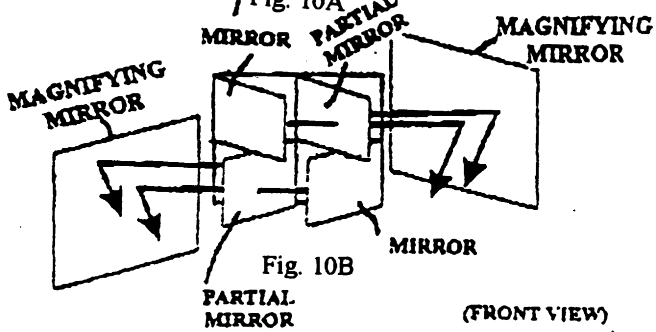

Fig. 10B