VIEWPOINT DEPENDENT IMAGING:

An Interactive Stereoscopic Display

by

Scott Stevens Fisher

B.S., Skidmore College
(1973)

Submitted in Partial Fulfillment
of the Requirements for the
Degree of
Master of Science in Visual Studies
at the
Massachusetts Institute of Technology
October, 1981
(i.e. February, 1982)

Signature of Author . . . . . .
Department of Architecture
October 8, 1981

Certified by . . . . . .
Nicholas Negroponte, Professor
of Computer Graphics
Thesis Supervisor

Accepted by . . . .
Professor Nicholas Negroponte, Chairman
Departmental Committee for Graduate Students

VIEWPOINT DEPENDENT IMAGING:

An Interactive Stereoscopic Display

by

Scott Stevens Fisher

Submitted to the Department of Architecture on October 8, 1981, in partial fulfillment of the requirements for the degree of Master of Science in Visual Studies.

ABSTRACT

Design and implementation of a Viewpoint Dependent Imaging system is described. The resultant display is an interactive, lifesize, stereoscopic image that becomes a window into a three dimensional visual environment. The image is continuously updated for a user's changing viewpoint relative to the display in x, y, and z by means of computer controlled optical videodisc and body tracking technologies.

The development of non-programmed, exploratory media and their precedents are also discussed.

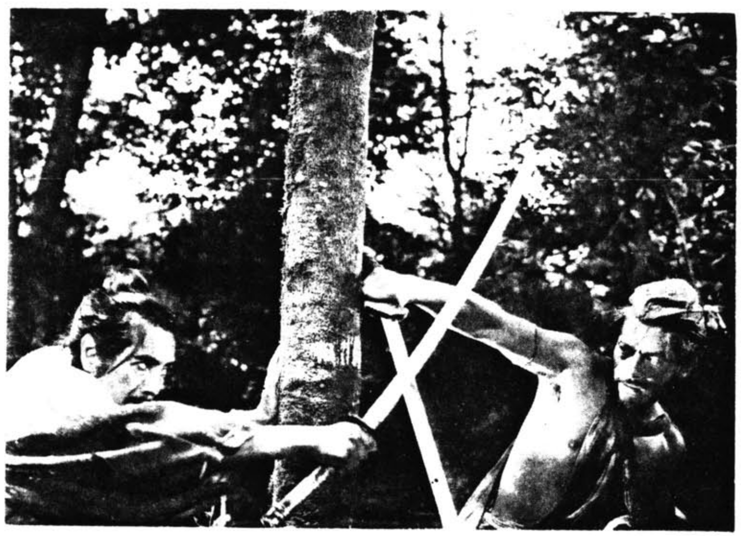
Thesis Supervisor: Nicholas Negroponte
Title: Professor of Computer Graphics

The work reported herein was supported by the Advanced Research Projects Agency of the Department of Defense, under Contract number MDA 903-81-C-0097.

TABLE OF CONTENTS

INTRODUCTION

In 1950, a Japanese movie director named Akira Kurosawa released a film called "Rashomon." Ostensibly a medieval Japanese murder mystery, the viewer's "image" of the actual criminal act must be synthesized solely from the disparate viewpoints of seven different observers and participants as they give testimony at the murderer's trial. No conclusion is given. (63)



At present, the increasingly rapid development of new imaging techniques in the field of media technology creates a similar situation for image viewers, users, and makers. For example, in the area of medical imaging, there exist numerous <u>different</u> techniques to image a particular malfunction.

A cardiac problem might be looked at with x-ray, ultrasound, tomographic scan, nuclear magnetic resonance, thermal imaging, dynamic spatial reconstruction, microwave imaging, radionuclide mapping, or by Stethoscope. Each method yields a slightly different type and amount of information- a slightly disparate viewpoint.

Today, the imaging specialist must be fluent with a vocabulary of imaging techniques - knowing which representations would prove most informative in a given situation. The number of possible media representations of an event, object or environment seem curiously related to Kurosawa's 'Argus-eyed' recount of the murder.

The research reported herein attempts to contribute yet another viewpoint for image representation and ways of looking at the world.

2. EXPLORATORY MEDIA

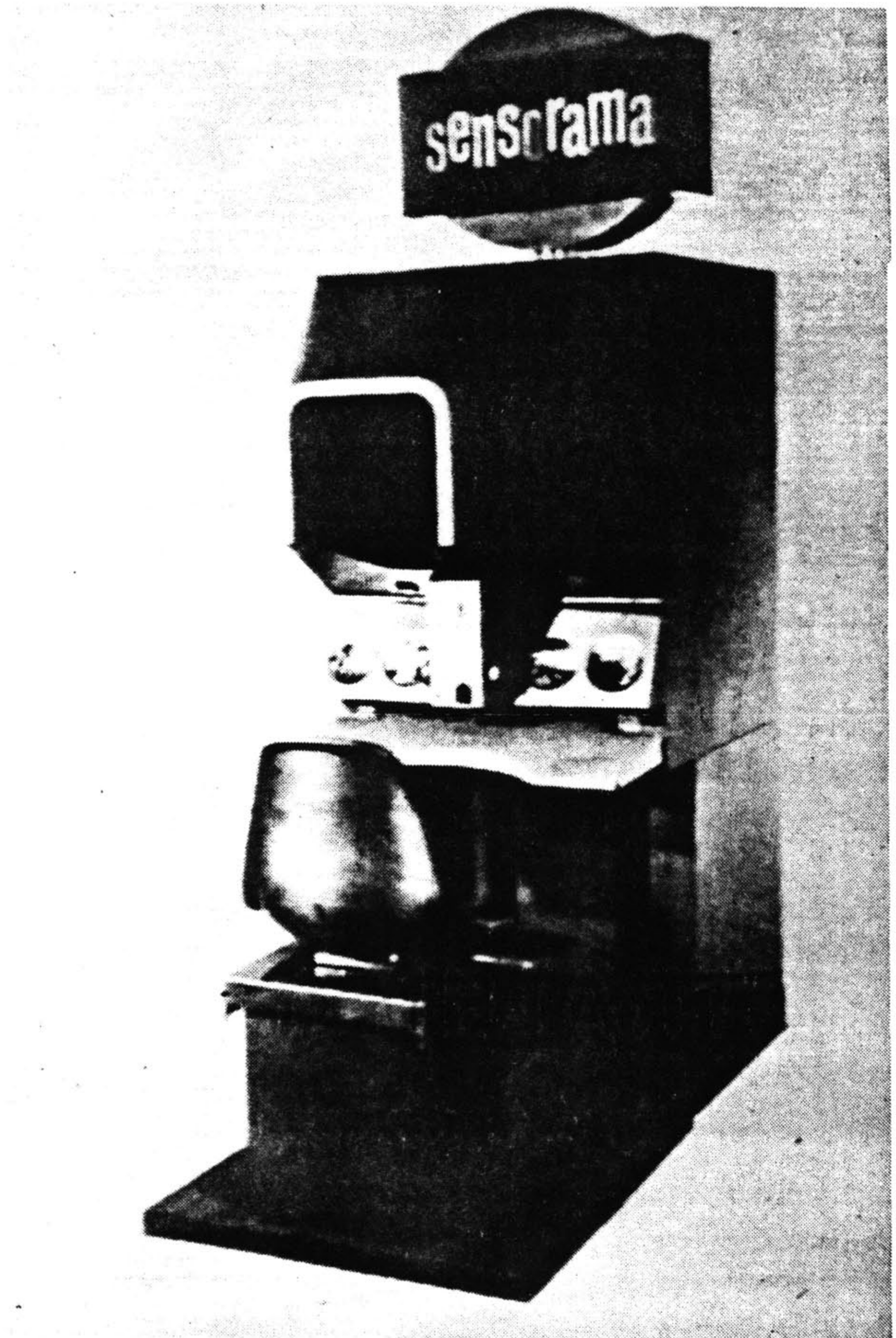
> "Watch out for a remarkable new process called SENSORAMA! It attempts to engulf the viewer in the stimuli of reality. Viewing of the color stereo film is replete with binaural sound, colors, winds, and vibrations. The original scene is recreated with remarkable fidelity. At this time, the system comes closer to duplicating reality than any other system we have seen!" [49]

Figure 1.

For most people, "duplicating reality" is an assumed, if not obvious goal for any contemporary imaging technology. The proof of the 'ideal' picture is not being able to discern object from representation - to be convinced that one is looking at the real thing. At best, this judgement is usually based on a first order evaluation of 'ease of identification'; realistic picutres should resemble what they represent. But resemblance is only part of the effect. In summing up prevailing theories of image-realism, Perkins comments:

> "Pictures inform by packaging information in light in essentially the same form that real objects and scenes package it, and the perceiver unwraps that package in essentially the same way." [60]

What is more important here is the <u>process</u> involved in 'unwrapping' the image. Evaluation of realism should also be based on how closely the presentation medium simulates dynamic perception in the real world. A truly informative picture would duplicate the act of confronting a specific scene in addition to merely being an informational surrogate. [16]



Sensorama viewing machine looks like
fugitive from a carnival, runs on quarters

Figure 1.

Additionally, psychologists refute the assumption that "when similarity [in image] reaches a maximum, it becomes identity". J.J. Gibson suggests that the notion of an image that is indistinguishable from reality is a myth, and points out that in all images, there will always exist a duality of image space; the space in which a picture lies and the space in which objects pictured lie. [24] The viewer must decide which information is relevant to picture recognition and which is not. More improvements, then, in creating a convincing virtual image will be a function of reducing awareness of space in which the picture lies.

Ways to implement these two additional factors governing image realism lie in the domain of media technology. Significant developments in this area are usually dictated by economics, available technologies and, as mentioned, cursory ideas about what the image should look like.

In 1978, John L. Baird demonstrated a viewing system called the "Baird Transmitter" [70]. In an illustration explaining the display, the user is labled as "subject undergoing television." Figure 2 . To date, Baird's description is still relevant. Television, as we experience it, plays to a passive audience. It has little to do with the ability to see at a distance other than in a vicarious sense; it offers only interpretations of remote events as seen through the eyes of others.

> "Except for rare instances, what is seen on network news is not the event itself unfolding before the live camera, or even a filmed record, but a story about the event reconstructed on film from selected fragments of it (or even from re-enactments of it)." [14]

Even though attempts are made at objectivity, TV news is the worst offender:

> "Our reporters do not cover stories from their point of view. They are presenting them from nobody's point of view." — R.S. Salant (CBS)

and definitively:

> "News is change as seen by an outsider (the correspondent) on behalf of other outsiders (the audience)." — Reuven Frank

Second hand information is probably better than none. But personal point of view is preferable:

> "We obtain raw, direct information in the process of interacting with the situations we encounter. Rarely intensive, direct experience has the advantage of coming through the totality of our internal processes--conscious, unconscious, visceral and mental--in processed, digested, abstracted second hand knowledge is often more generalized and concentrated, but usually affects us only intellectually— lacking the balance and completeness of experienced situations." [2]

In 1962, the Sensorama display previously cited was a remarkable attempt at simulating personal experience of an environment using state of the art technology. [31, 49] Figure 1. The display offered a barrage of environmental stimuli: 3D film, binaural sound, odors, wind and vibration. Despite Lipton's enthusiasm, even today, most people would regard odor, wind, and vibration cues as irrelevant information. At one time color information was thought to be an unnecessary addition to black and white images. What is considered a subtle perceptual nuance may become a media standard in the future.

As an environmental simulator, the Sensorama display was one of the first steps toward duplicating the viewer's act of confronting a real scene. The user is totally immersed in an information booth designed to imitate the mode of exploration while the scene is imaged simultaneously through several senses. The next step is to allow the viewer to control his own path through available information to create a highly personalized interaction capability bordering on the threshold of <u>virtual exploration</u>.

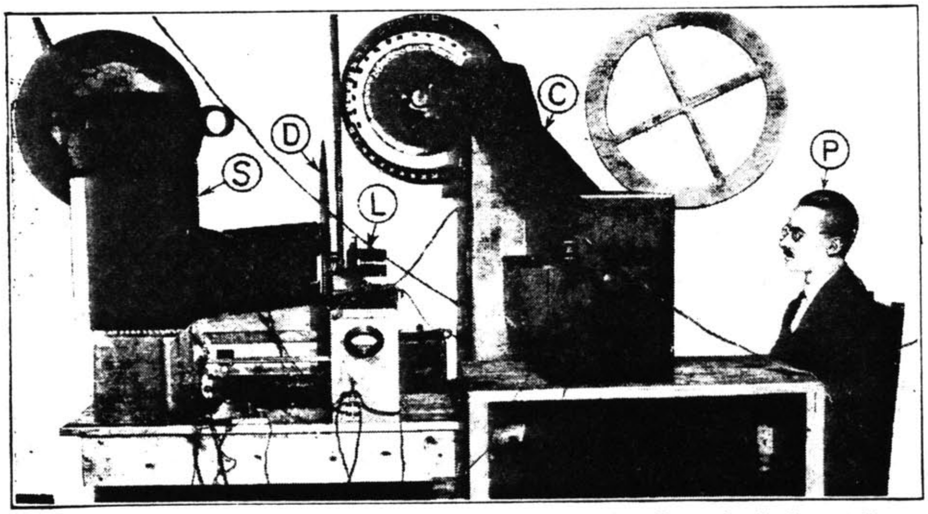
This has been the subject of recent research at the MIT Architecture Machine Group. A virtual travel system was implemented in which a user can virtually drive around the town of Aspen via computer controlled videodisc imagery presented on a touch sensitive television screen. Controlling speed and direction through graphic overlays, the user can drive up and down any street, turn corners and even access information inside buildings. [48] A similar system was also implemented in which the user can spatially navigate through stored information at many levels. [4]

A key feature of these systems is that the viewer's movements are non-programmed; that is, he is free to choose his own path through available information rather than watching a 'tour'. For this system to operate convincingly, a comprehensive information database must be available to allow the user sufficient points of view.

With the addition of live camera input to this system, television would finally become a true extension of the visual sense. Coupled with the remote extension of other senses, the result would be the ultimate in <u>exploratory media</u>- what Marvin Minsky calls "Telepresence". [51]

> "In the case of Venus if I recall correctly, the human operator in orbit wore an exoskeleton which controlled the movements of the body, arms, legs and hands of the device on the surface below, receiving motion and force feedback through a system of airjet transducers. He had on a helmet controlling the slave devices television camera - set, obviously enough, in its turret- which filled his field of vision with the scene below. He also wore earphones connected with its audio pickup. I read the book he wrote later. He said that for long stretches of time he would forget the cabin, forget that he was at the boss end of a control loop, and actually feel as if he were stalking through that hellish landscape. I remember being impressed by it, just being a kid, and I wanted a super tiny one all my own, so that I could wade around in puddles picking fights with micro-organisms." [74] See Figure 3 .

> "In life there were always surprises, but there were few in the laws of perspective." [27]



In this picture of the Baird transmitter, the photoelectric cells are in the box at the right: S, light-source; D, scanning disc; L, lens; C, cloth covering cells; P, subject undergoing television.

Figure 2.

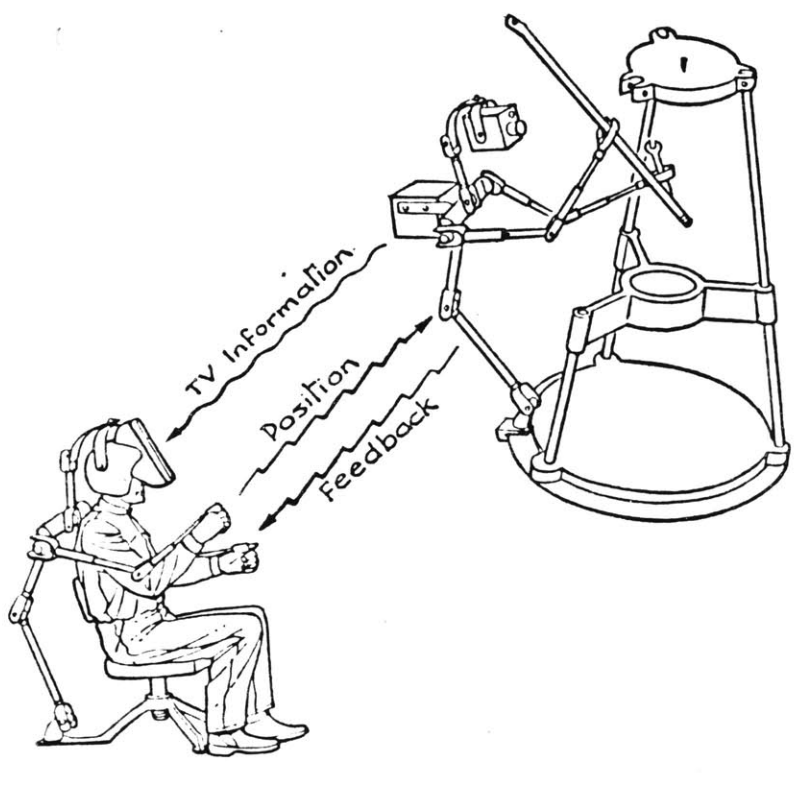


Figure 3.

The normal way of seeing an object in the real world is by moving around it, not by viewing from a fixed viewpoint:

> "To perceive from a fixed point of observation, one that is persistently stationary, is not the case with which to begin the study of perception, for it is atypical. The perspectives of the environment are unaturally frozen in time. It is an even more limited and restricted case than to perceive at a particular point of observation... The special case of the frozen optic array is of concern to painters; it is not relevant to the problem of how we see but only to the special problem of how we see by means of pictures." [25]

The most important feature of an "exploratory" display is the ability to liberate the user to move around in a virtual environment, or, on a smaller scale, viscerally peruse a scene. In essence, the viewer has access to greater than one viewpoint of a given image allowing him to synthesize a strong visual percept from many points of view. Availability of multiple points of view places an object in context and animates its meaning. As Merleau-Ponty observes:

> "Nothing speaks in isolation," [50]

and proceeds to make an exhaustive examination of the subtleties involved in viewing position:

"For each object, as for each picture in an art gallery, there is an optimum distance from which it requires to be seen a direction viewed from which it vouchsafes most of itself: at a shorter or greater distance we have merely a perception blurred through excess or deficiency. We therefore tend towards the maximum of visibility, and seek a better focus as with a microscope.[1] This is obtained through a certain balance between the inner and outer horizon: a living body, but a mass of matter as outlandish as a lunar landscape, as can be appreciated by inspecting a segment of skin through a magnifying glass. Again, seen from too great a distance, the body loses its living value, and is seen simply as a puppet or automaton. The living body itself appears when its microstructure is neither excessively not insufficiently visible, and this moment equally determines its real size and shape. The distance from em to the object is not a size which increases or decreases, but a tension which fluctuates round a norm. An oblique position of the object in relation to me is not measured by an angle which forms with the plane of my face, but felt as a lack of balance, as an unequal distribution of its influences upon me. The variations in appearance are not so many increases in size, or real distortions. It is simply that sometimes the parts mingle and become confused, at others they link up into a clearly articulated whole, and reveal their wealth of detail."

A vocabulary of 'viewpoint' - the awareness of visual 'surprise' and transformation that confront the active observer in the real world - is embodied historically the area of environmental design:

"The Chinese, in their large gardens, contrive different scenes for different times of day, disposing at the points of view, buildings which from their use point out the proper hour for enjoying the view in its perfections." [11]

---

[1]Schapp, Beiträge zur Phänomenologie der Wahrnehmung, pp.59 & ff.

and:

> "Observers move within architectural spaces, and their views change greatly with each movement...It is the implicit task of the viewer to organize all of these separate images into one coherent whole, in order to ascertain the sense of an entire architectural space from the juxtaposition of multiple exterior and interior views." [36]

This ability for object 'recognition' by synthesis of multiple points of view is mimicked on a different scale in the field of artificial intelligence. Research in robot vision simulation has developed an analogous method to locate an arbitrarily oriented object in a complex scene. Multiple filters are used that have a target object recorded on them as seen from many disparate viewpoints. Cycling through an array of 5 to 8 yields good recognition ability. [47]

Following in Section 3 is a summary of imaging technologies that have further contributed to a vocabulary of viewpoint dependent images. All have in some way been responsive to the viewing position of an active observer.

Section 4 is a description of the display system designed to respond continually to an active observer.

3. HISTORICAL PRECEDENTS

3.1 FIXED POINTS OF VIEW

Image making is almost always related to viewpoint. To create the kinds of 2D image representations we are used to today, man had to learn to see from a single point of view. [24] The invention of prespective was a method to see the world as a picture.

The most extensive examples of perspective imaging are works of art done in the post-renaissance. In most perspective rendered images, there is only one correct viewing position. Other points of view will yield distortions in the image. In much of anamorphic art, however, these distortions are so pronounced that the image is unreadable except from the proper point of view. Figure 5 . In many examples, this idea was integrated into architectural spaces so that, for example, a viewer moving through a hallway would suddenly come upon the correct access point for the image stretched along the wall. [46] Figure 6 . The most impressive of these works is the ceiling of the Church of St. Ignazio in Rome, done by Pozzo in the 17th century. [61, 46]. Viewed from the correct point in the center of the church, the virtual height of the ceiling is almost doubled by an elaborate 2D painting of a dome. As the viewer moves, gross distortions become evident and the illusion is destroyed, Figure 4.

Figure 4.

Figure 5.

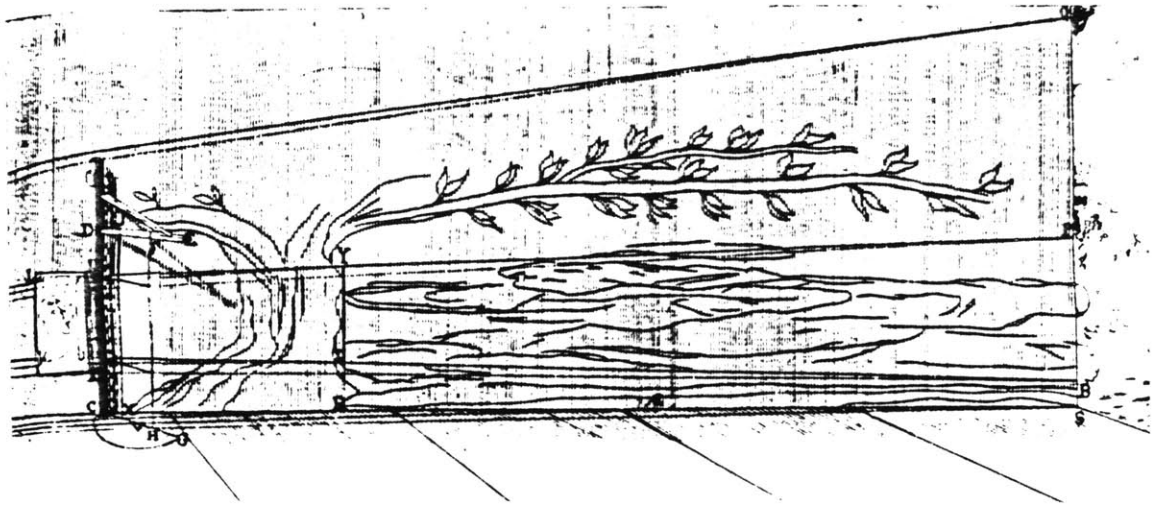


Figure 6.

Few people realize that photographs are also perspective projects and therefore have a correct viewing position point or 'station point'. When seen from this position with one eye closed, the image will usually appear 3D and, if the edges are concealed, lifesize. But our awareness of subsequent distortions from the wrong point of view is usually neglected:

> "People don't compensate for variations of viewpoint; they tolerate them." [60]

3.2 MULTIPLE POINTS OF VIEW

Perspective images from a single point of view are not the only way to represent the real visual world. There have always been attempts to multiplex more than one image or viewpoint into a single picture.

> "Natives of British Columbia represented a bear, say, in full face and profile, from back, above and below, from within and without all simultaneously. By an extraordinary mixture of convention and realism, these butcher-draftsmen skinned and boned, even removed the entrails, to construct a new being, on a flat surface, that retained every significant element of the whole creature." [8]

A modern version of this idea was the "Biscenorama", Figure 7., a favorite novelty in the 19th century. By making an accordian fold in a piece of paper, two images could be interlaced on a single sheet. Each image was visible only from a different point of view.

Figure 7.

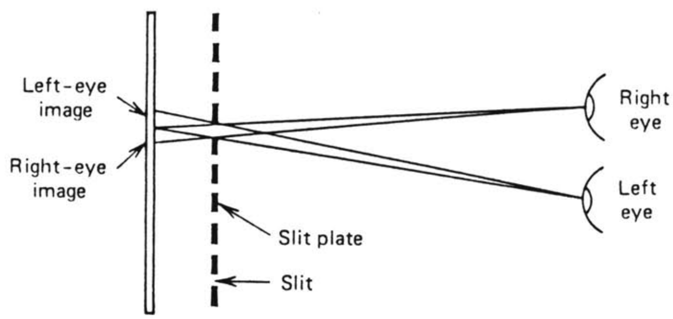


Figure 8.

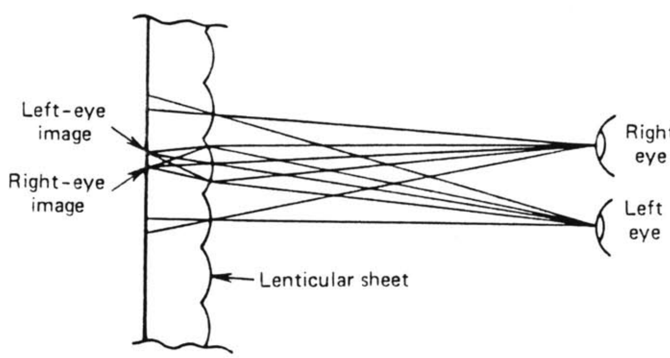


Figure 9.

Stereoscopic images.

Most of our visual experience is a constant synthesis of two disparate points of view. Because our eyes are horizontally separated by about 65mm, we see two slightly different images. This binocular disparity is a strong cue for depth perception and is essential in making 3D stereoscopic images.

Images taken translating along an axis parallel to that of the scene will yield disparate images that contain depth information of the scene when combined. For example, images taken at equal intervals along an airplane's flight path yield stereo images that are used for measurement-photogrammetry. A similar algorithm is used in the artificial intelligence field to derive depth information from the viewpoints of a monocular observer moving parallel to a scene. [45]

Research in stereoscopic imagery has developed methods to present different images to each eye without necessary viewing aids on the user. The simplest method involves a raster screen much like the bisceneorama. Disparate images are interlaced on one surface and the position and pitch of the screen in front directs only the correct image to each eye. Figure 8 .

In 1930, Herbert Ives substituted a screen of vertical cylindrical lenses in front of the multiplexed images to increase the resolution of the image and to allow more view-

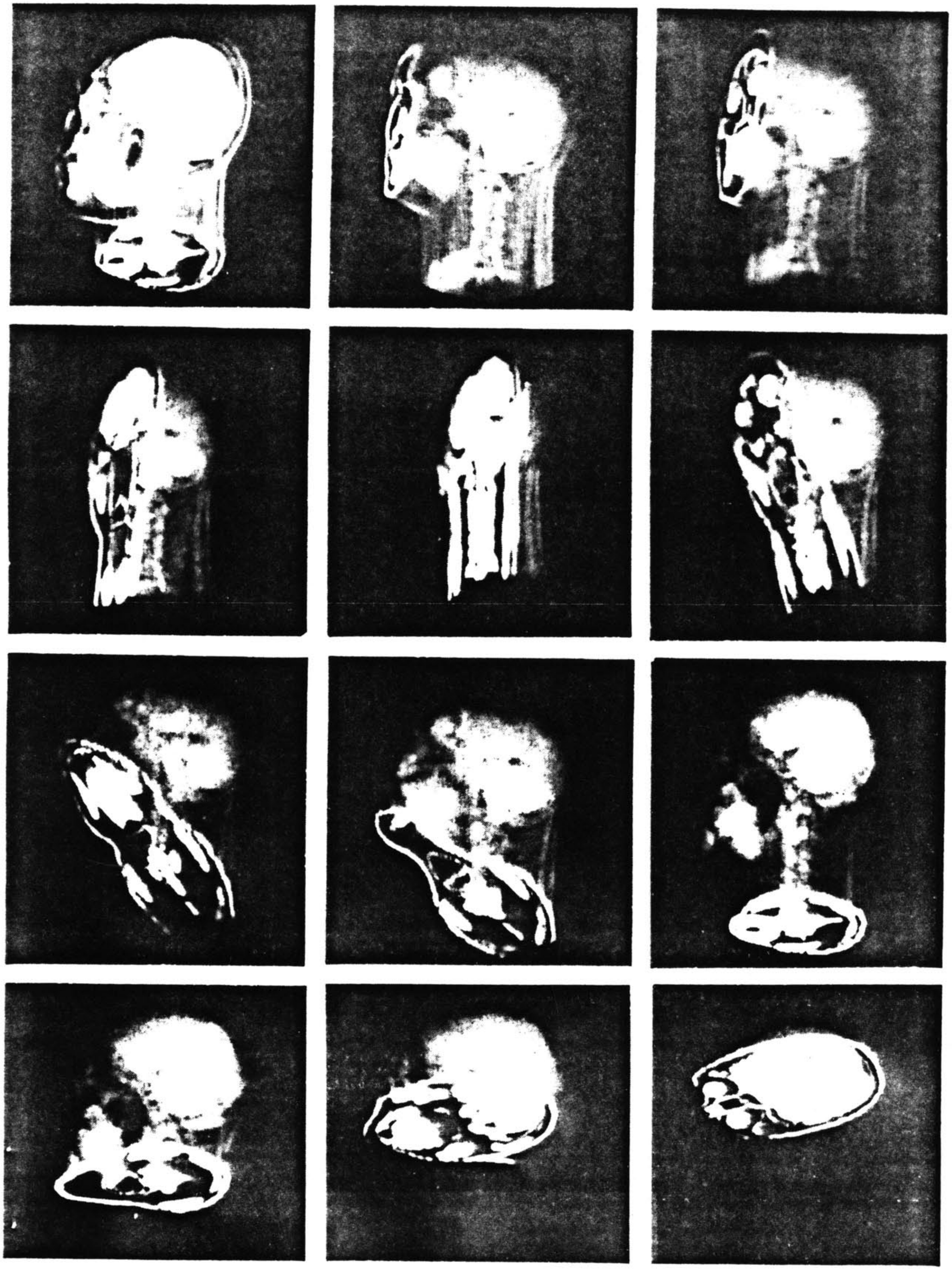


 Figure 10.

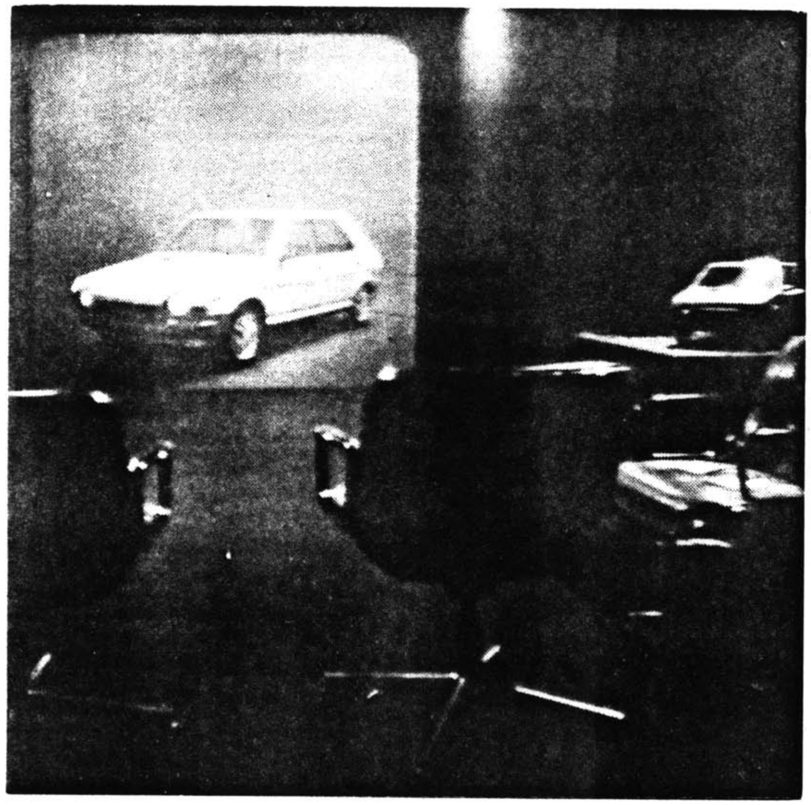


Figure 11.

Nov. 18, 1969 D. GABOR 3,479,111

THREE-DIMENSIONAL PICTURE PROJECTION

Filed Aug. 24, 1967 2 Sheets-Sheet 1

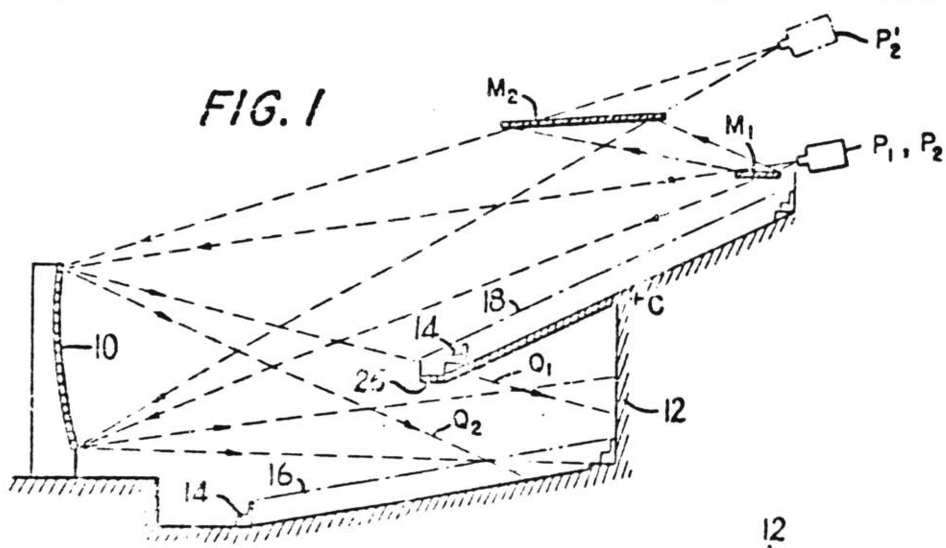


Figure 12.

points to be interlaced. [40] Figure 9 . This lenticular screen allowed the viewer to move slightly in front of the image and perceive motion parallax information as well as binocular parallax. Research was continued by the Bonnet studio in France and most recently by the Nimslo Corporation in Atlanta, Georgia.

Additional research in 3D imaging has been in volume displays which rapidly present sequential slices through an image space. Recent developments are a vibrating, varifocal mirror system and a rotating wedge shaped screen. Both systems stack up image slices so that a viewer can move around the display for varied points of view. Figure .

One of the most important new 3D display technologies is the hologram. By recording the interface patterns formed by objects exposed to coherent laser light, a 3D image can be produced that needs no viewing aids and which has horizontal and vertical parallax information. Compared to a stereo photograph taken from only two points of view, the hologram offers every possible point of view through a window defined by the film plane edge.

Projection environments.

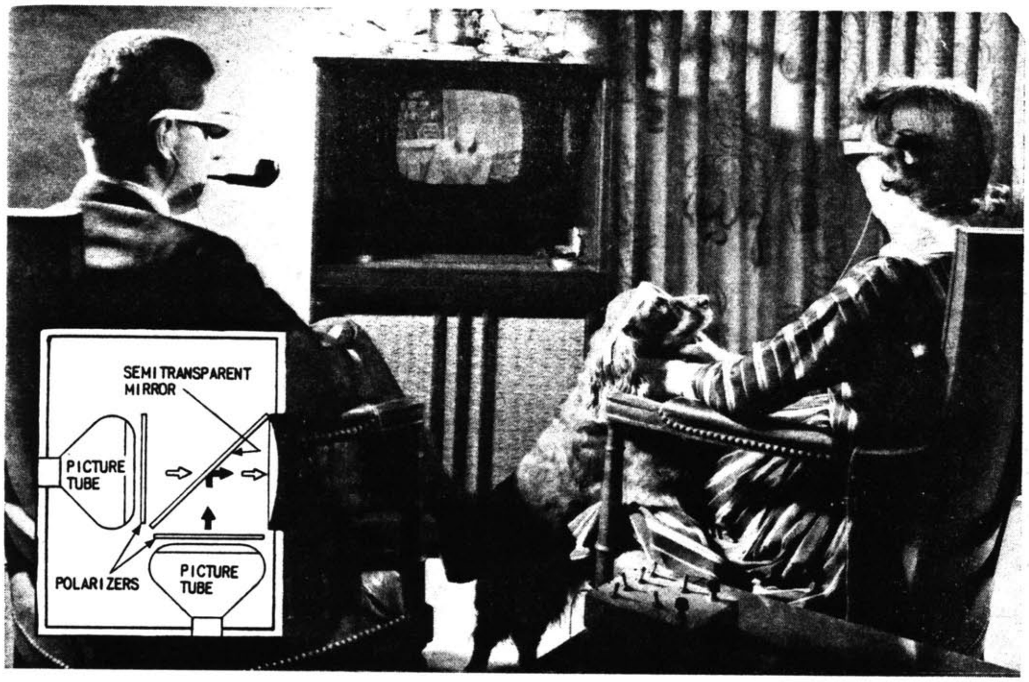
3D research has also been directed towards large projection environments. Unlike traditional three dimensional movies, development has been towards offering a position-corrected image for each person's viewpoint. In addition to the research cited in this paper, the author has implemented a 3D, lifesize display system for Fiat/Lancia as a design aid to enlarge 1:5 scale design models to virtual lifesize models that may be looked at from up to thirty points of view around the vertical axis. Figure 11.

Dennis Gabor, inventor of holography, has also patented several projection environments that employ holographic projection screens. Their direction selective properties allow presentation of different points of view to each member of the audience. Figure 12.

A mechanical version of the direction selective screen has been suggested by Robert Collender. Images are recorded by a hugh array of small electronic cameras and transferred to a special, single film strip. This is projected onto a screen composed of many vertical rotating, multifaceted mirror elements phased with the projector. Again each audience member sees his correct viewpoint for an angle of view up to 180°.

Television.

One of the first viewpoint dependent TV systems was the Duoscope developed in 1954.[37] Figure 13. Using a half-silvered mirror angled between two orthogonally polarized TV tubes, two viewers can watch different shows on the same screen area. (This can also be adapted to stereo-scopic images by presenting left and right eye images on each tube. The viewer wears orthogonally polarized glasses.)



Two-Headed TV Set Displays Two Different Shows at Once

Two people can enjoy different TV programs at the same time with a new set. The experimental Du Mont Duoscopic is actually two receivers in one cabinet, with two chassis, two sets of controls and two viewing tubes mounted at right angles (inset). A semitransparent mirror superimposes the two pictures, but each viewer sees only one show by watching through polarizing spectacles. Earphones handle the sound.

Figure 13.

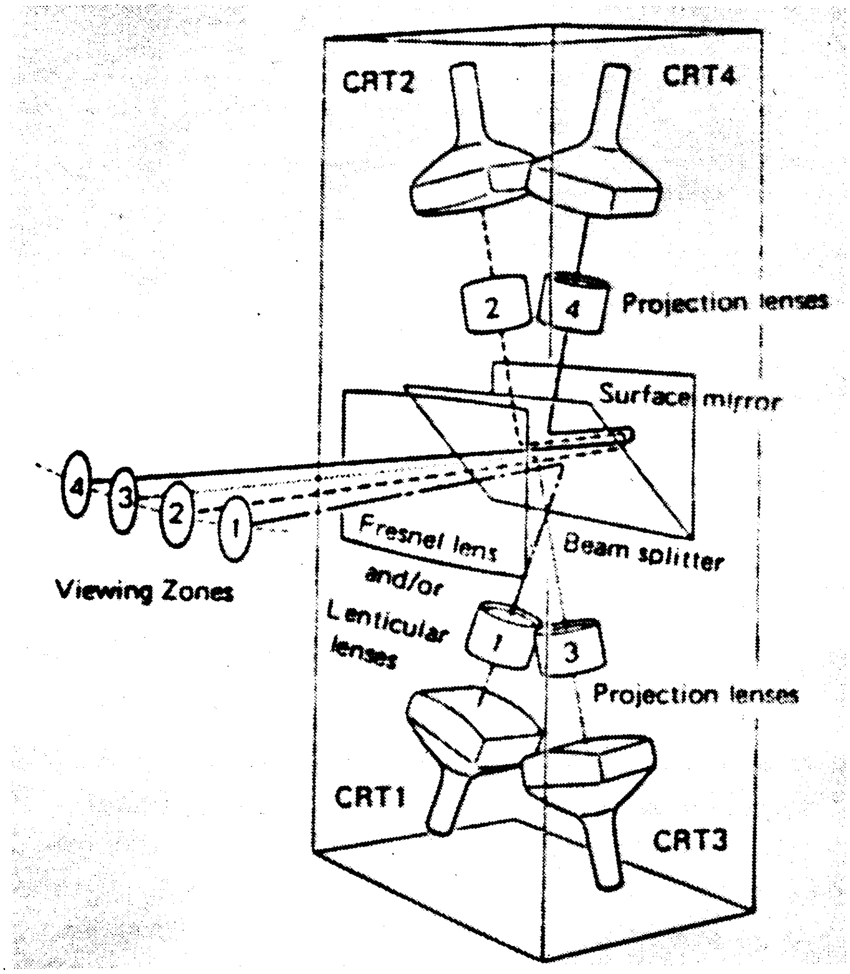

CRT2
CRT4
2
4
Projection lenses
Surface mirror
4
3
2
1
Fresnel lens
and/or
Lenticular
lenses
Beam splitter
Viewing Zones
1
3
Projection lenses
CRT1
CRT3

Figure 14.

Recent research by Sony has developed a lenticular 3D TV system with 4 viewing zones. Figure 14.

Research in the area of 3D computer graphics using digital storage techniques to access a comprehensive image database allows a user to view generated scenes from any given point of view. By writing one frame at a time, animation sequences can be built up.

3.3 VIRTUAL ENVIRONMENTS

As discussed in Section 2, virtual environments are a type of interactive display system in which the visceral process of a user's interaction with an environment is simulated in addition to visual information. The viewer is immersed in a virtual environment. The Sensorama display system and the virtual travel system are presented as first examples.

Much research in this area is in development of sophisticated flight simulators. Use of head mounted displays mounted on the helmet visor [15] or on eyeglasses, and head tracking systems, allows imagery to be projected only where the pilot is looking ("the area of interest"), within a wide field of view. [29]

A similar system for presentation of a visual surround is the 360°, non-programmed visual simulator in Orlando [56]. A helicopter pilot is surrounded by a spherical display

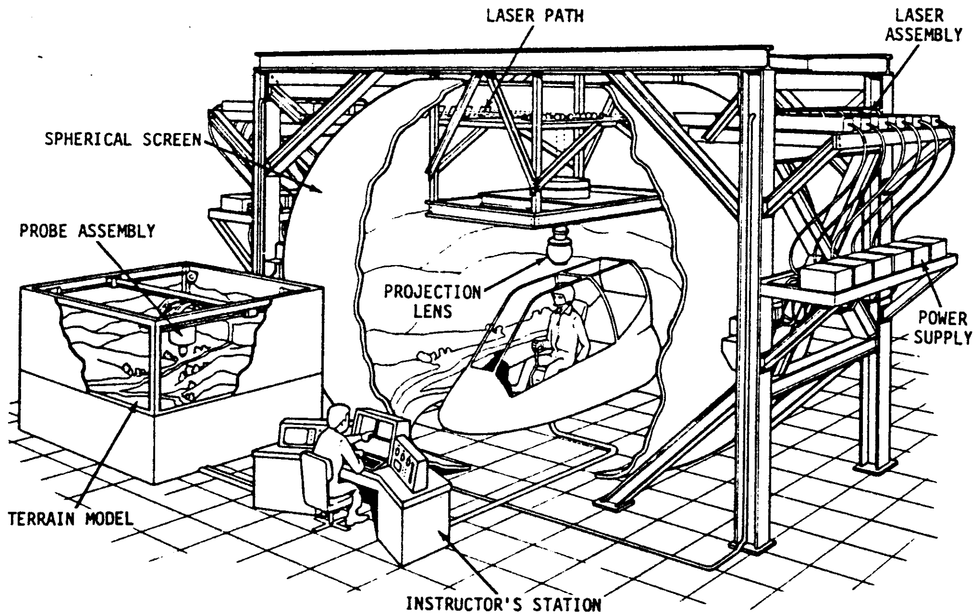


Figure 15.

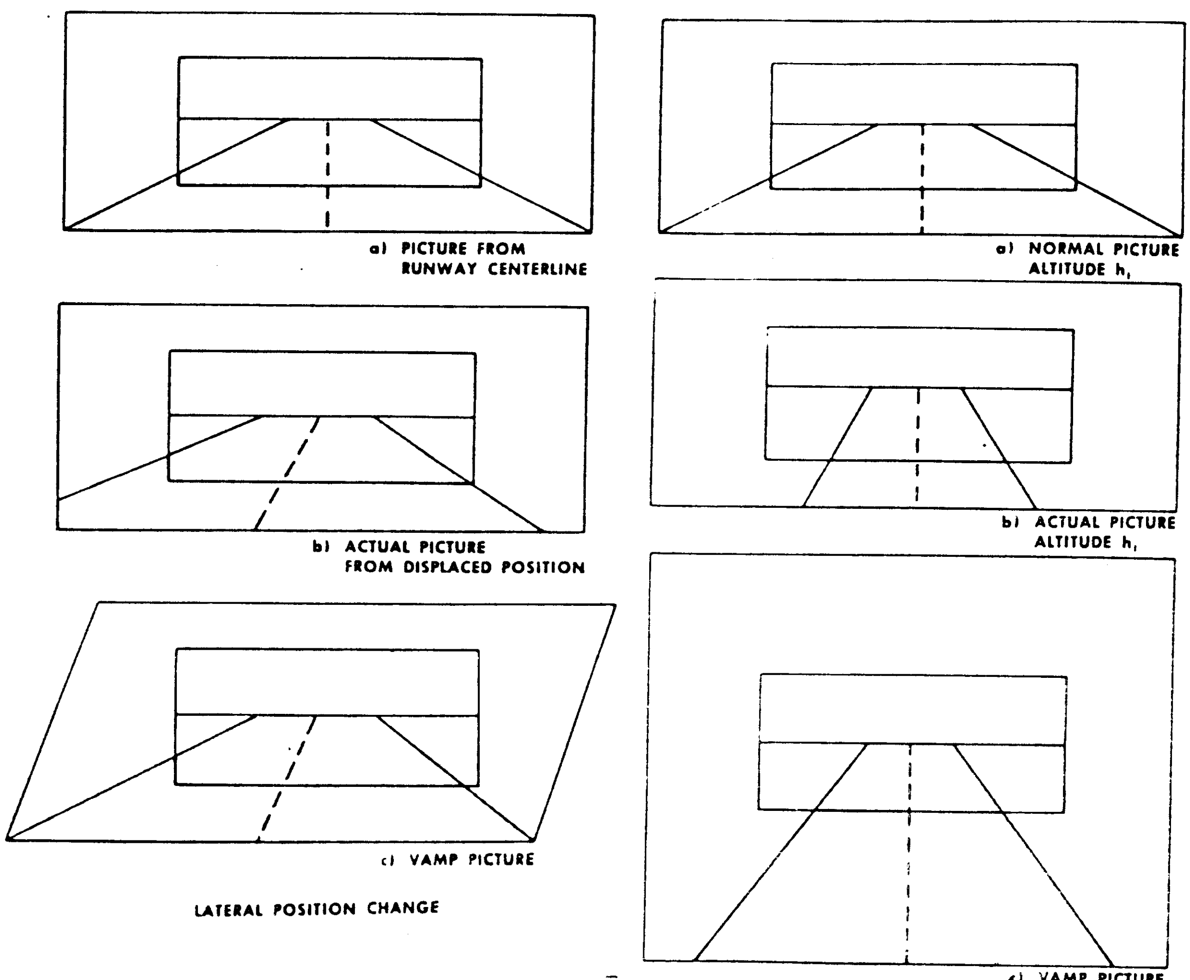


Figure 16.

surface on which images are laser projected. The source of imagery is a small TV camera probe which the pilot manuevers through a scaled terrain model as if in real flight. Figure .

Another presentation technique for simulating changing points of view in flight is use of "distortion optics". [13] Changes in lateral translation in front of a generated scene are simulated by transforming the rectangular image format to a parallelogram and in vertical translation by stretching the rectangle in that axis. By concealing the frame edges, the result is convincing. Figure 16.

To date, the most sophisticated displays have been computer generated virtual environments. In the late 60's at Harvard and MIT, Ivan Sutherland developed a head mounted three dimensional display that presented a perspective image that changed as the viewer moved. [68] Figure 17. Presented separately to each eye, a 3D vector graphic image is superimposed into the user's environment. Tracked first by a mechanical body tracker ('the sword of Damocles') and later by an ultrasonic tracker, the image was transformed according to the user's angle of regard. The user could, within limits, walk around these objects and later interact with them with a wand. [72, 73] . This was a non-screen dependent, 360° display system.

A similar system, "Stereo Matrix", was developed at the University of Illinois in 1973. [43] The viewer

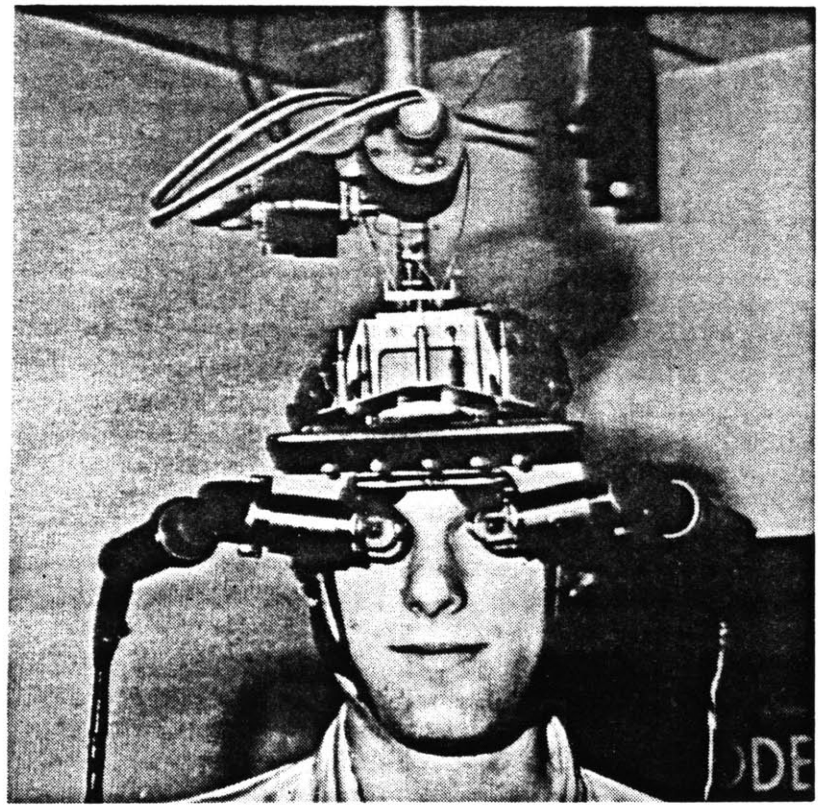


Figure 17.

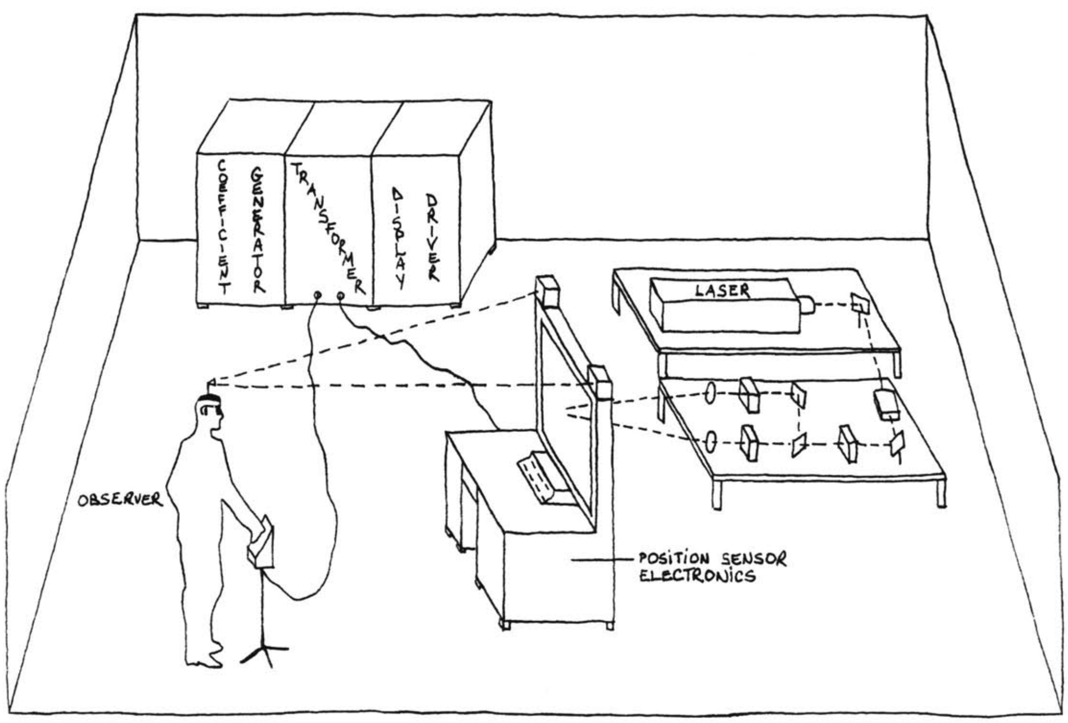


Figure 18.

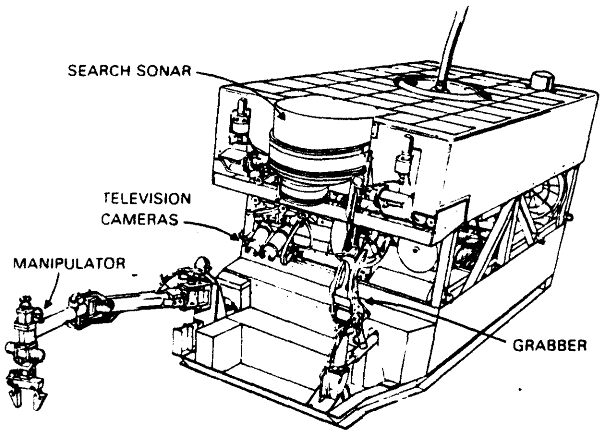
moved about in a 10 x 10 foot area in front of a 3 x 4 foot image rear projected by laser. The image was a polarized, 3D vector graphic display written at 60 fps and updated for the viewer's position (tracked by IR at 60 points per second). Translation, rotation and scaling of the image was under operator control and a 3D cursor was used for interaction. Figure 18.

3.4 TELEPRESENCE

The ultimate virtual environment is one which is a true extension of the user's vision through means of a remote camera probe. Most research has been towards exploration of adverse environments such as poison gas, under water or in space. The Remote Un-Manned Work System (RUWS) developed by the Navy has dual color video cameras for stereo viewing of underwater manipulator tasks. [64] Figure 19.

A similar system was devised at MIT's Lincoln Lab for the Mars exploration [69]. Figure 20.

Recently a system has been designed for British helicopter pilots in which a video camera on the nose of a helicopter is linked with head tracking to a display on the pilot's helmet. The pilot's direction of gaze controls the low light camera's movements for search operations in darkness. Similar systems are cited by Minsky, in which a camera mounted on top of a building is controlled by a remote user's head movements [51], and by New York artist



RUWS stereoscopic video system

Figure 19.

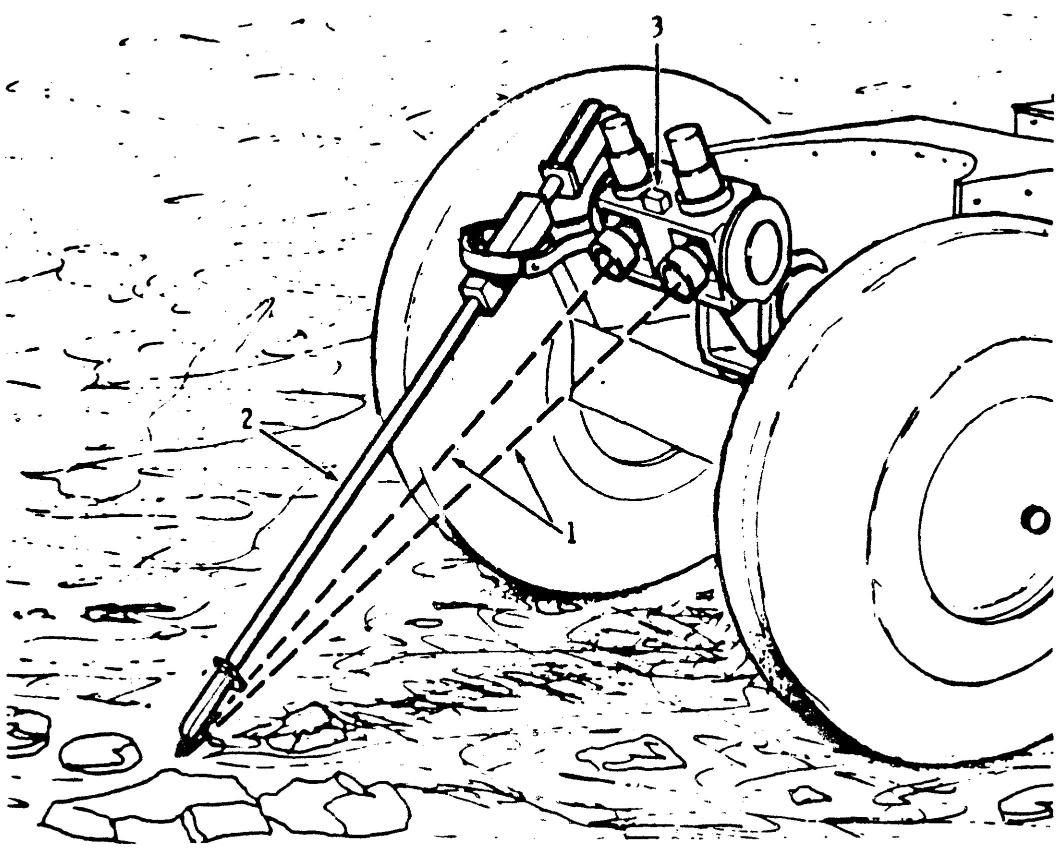


Figure 20.

Alphons Shilling, who built a video camera system for direct 3D viewing at a distance. Figure 21.

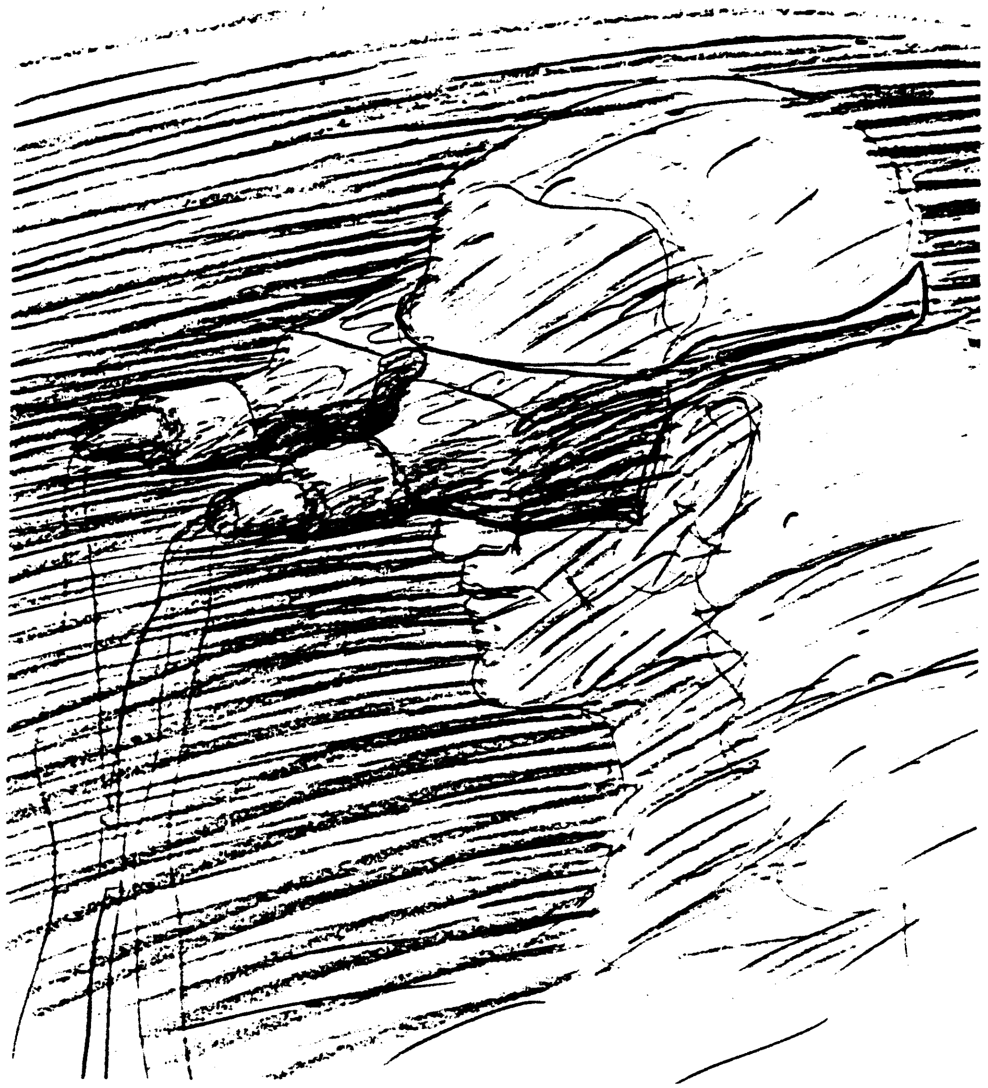


Figure 21.

4. VIEWPOINT DEPENDENT IMAGING DISPLAY

4.1 OBJECT

The object of this research is to design and implement a viewpoint dependent imaging system in which a video monitor or projection screen becomes a virtual window into a three dimensional visual environment. As a user physically changes his viewpoint of the represented environment in relation to the display surface, a series of stereoscopic images is retrieved from an image array stored on optical videodisc. All possible viewpoints of the virtual space are recorded and become fluidly available in coordination with the viewer's movements. The resulting display is continuously updated to present a perspective corrected, lifesize, 3D image that is under user-control.

4.2 TASKS

The design problem undertaken here is an attempt to more closely represent an observer's interactive perceptual experience of the visual world by presenting sensory information not offered by traditional media technologies.

In this display, necessary information relative to a user's point of observation is generated by three important cues for visual depth perception and is presented to the viewer by means of state of the art display technology. Although similar to the familiar vocabulary of film and video camera movements, the emphasis of this display is on user control as

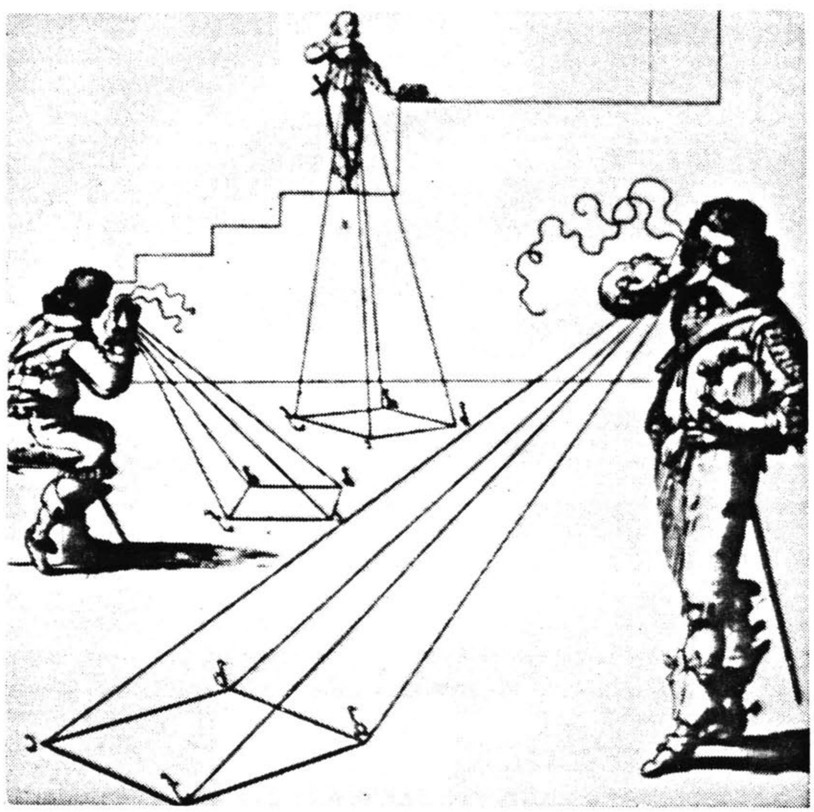


Figure 22.

opposed to vicarious, directed observation. These cues allow the viewer to virtually <u>explore</u> the image space as in a real environment - combining several disparate images into a coherent experience of that space.

A first priority is to define the viewer's position in coordinates relative to the plane of the display screen:

x = position along the <u>horizontal</u> axis parallel to the display surface

y = position along the <u>vertical</u> axis parallel to the display surface

z = position along the axis perpendicular to the display surface

Change in position information for an active observer is described as translation along these axes relative to the screen and yields three important cues:

Translation parallel to the plane of the screen in the x and y axes is called <u>motion parallax</u>. Equivalent to 'tracking' movements in cinematography [35] (and not to 'panning' moves), motion parallax is revealed as change in relative positions of far and near objects in a scene. The amount and direction of change is described as horizontal or vertical parallax and is the strongest visual depth perception indicator over distances up to 10 meters and beyond [30]. Ittelson describes it accordingly:

> "If one looks, with one eye closed, up into the dense foliage of a tree, the jumbled and disorderly array of leaves and branches, seen with head motionless, quickly assumes order and spatial localization if the head is moved continuously left and right six or eight inches." [39] See Figure 23 and Figure 33.

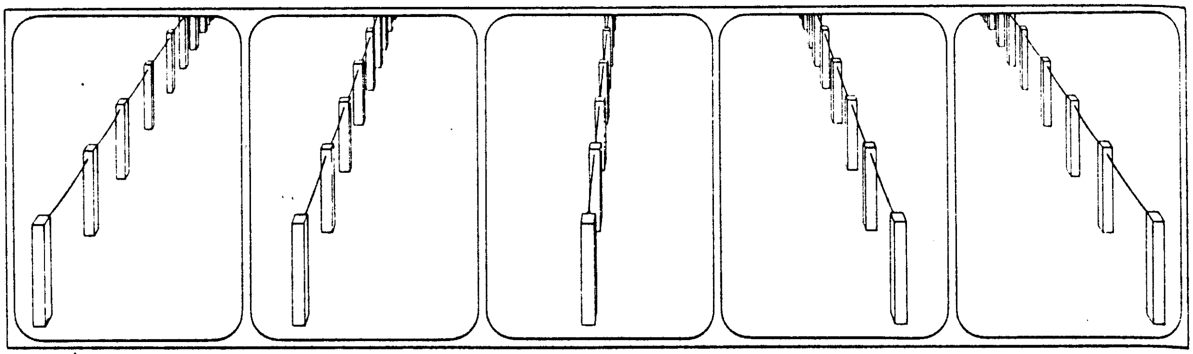


Figure 23.

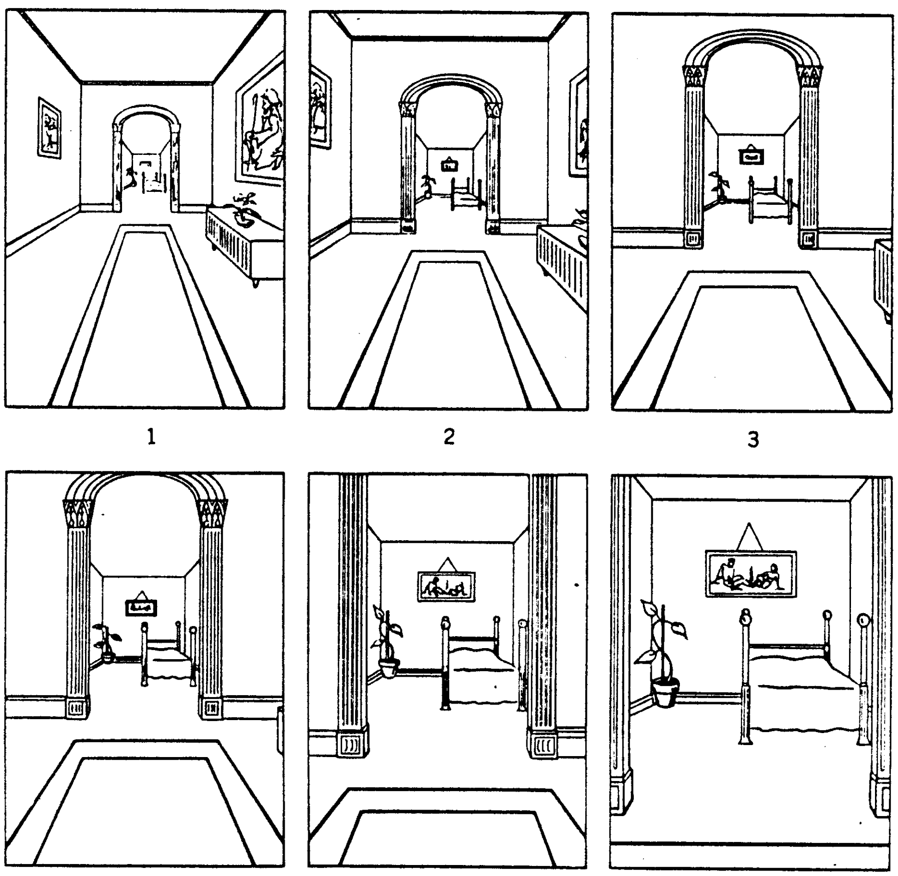
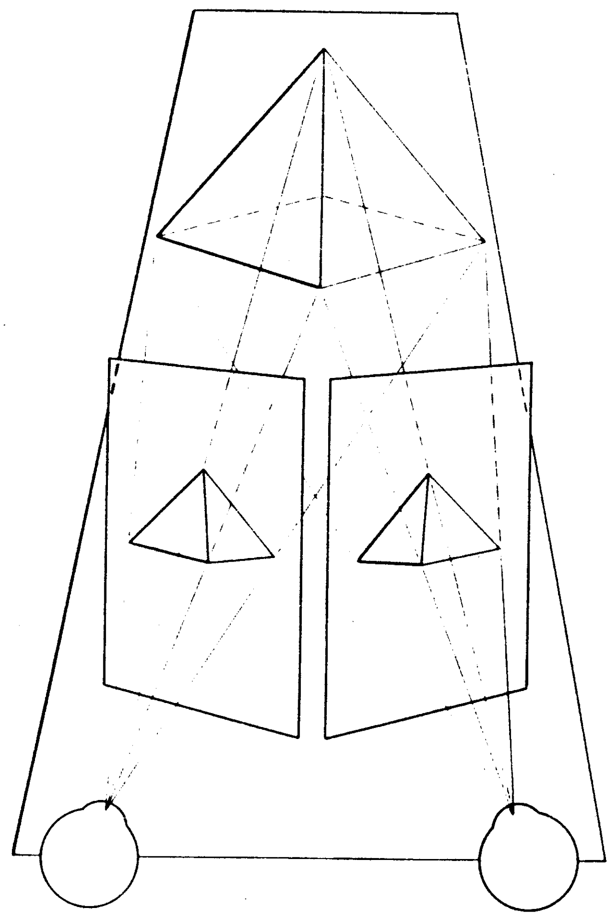
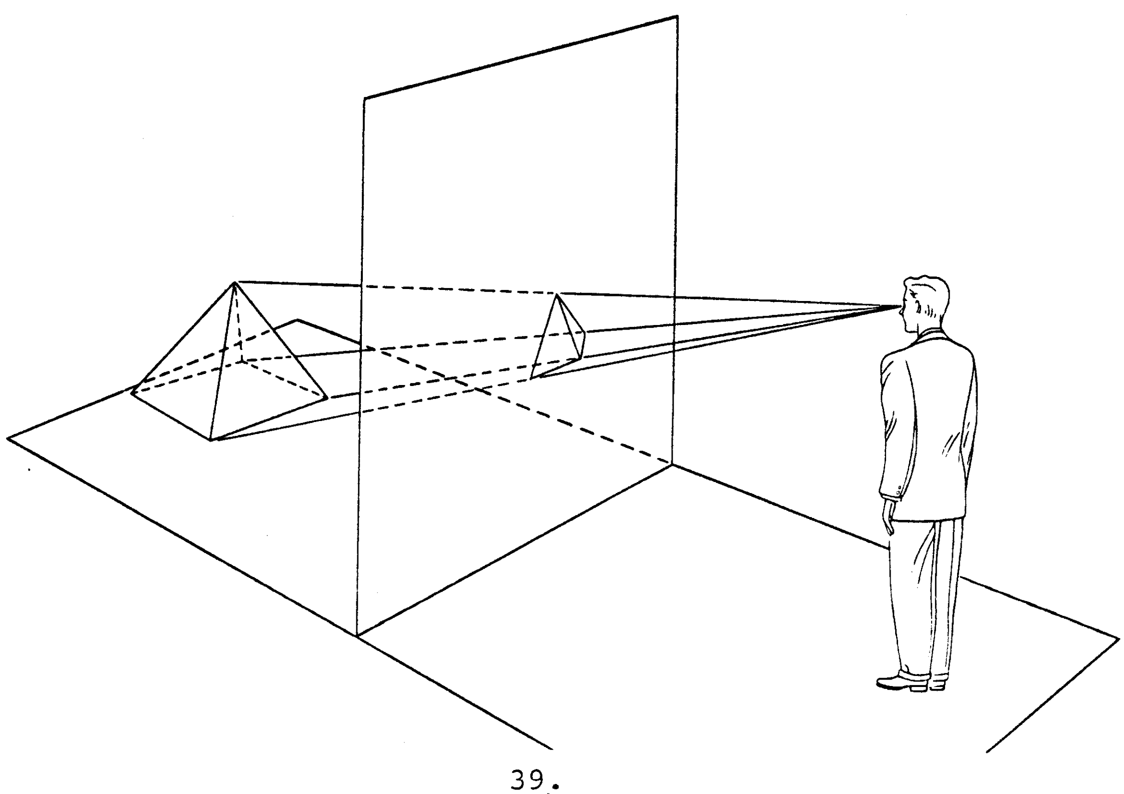


Figure 24.

Figure 25.





A subset of translation in the x axis yields binocular parallax or binocular disparity. Because our eyes are horizontally displaced at an average of 65mm, each sees from a slightly disparate viewpoint. An informative exercise is to station oneself in front of a large piece of glass and, holding steady, outline objects seen through the glass by one eye with the other closed. Without moving, reverse the process and trace the view seen by the other eye. The resultant image clearly illustrates the variations and similarities in the two images. As Pirenne mentions, for any scene, there exist two cones of vision, with one apex to each eye corresponding to the center of projection or center of perspective for that particular viewpoint. [61] See Figures 22 & 25. Intersection by a plane perpendicular to this axis of projection will yield images which may be optically fused as a stereoscopic image. The availability of two projectors of a given scene is particularly important with regard to resolving image ambiguities. Given only one cone of vision, as in traditional 2D displays, there exist an infinite number of possible object positions that the image displayed on the screen could be generated by. For example, a 5 foot high image of a face could be interpreted as either a closeup of a normal sized head or as the head of a giant. With the additional viewpoint in binocular vision, the image is specifically located in the z axis and the actual size confirmed. Depth information offered by this cue is effective to about 10 meters, beyond which image disparities are unresolvable. [30] Figure 26.

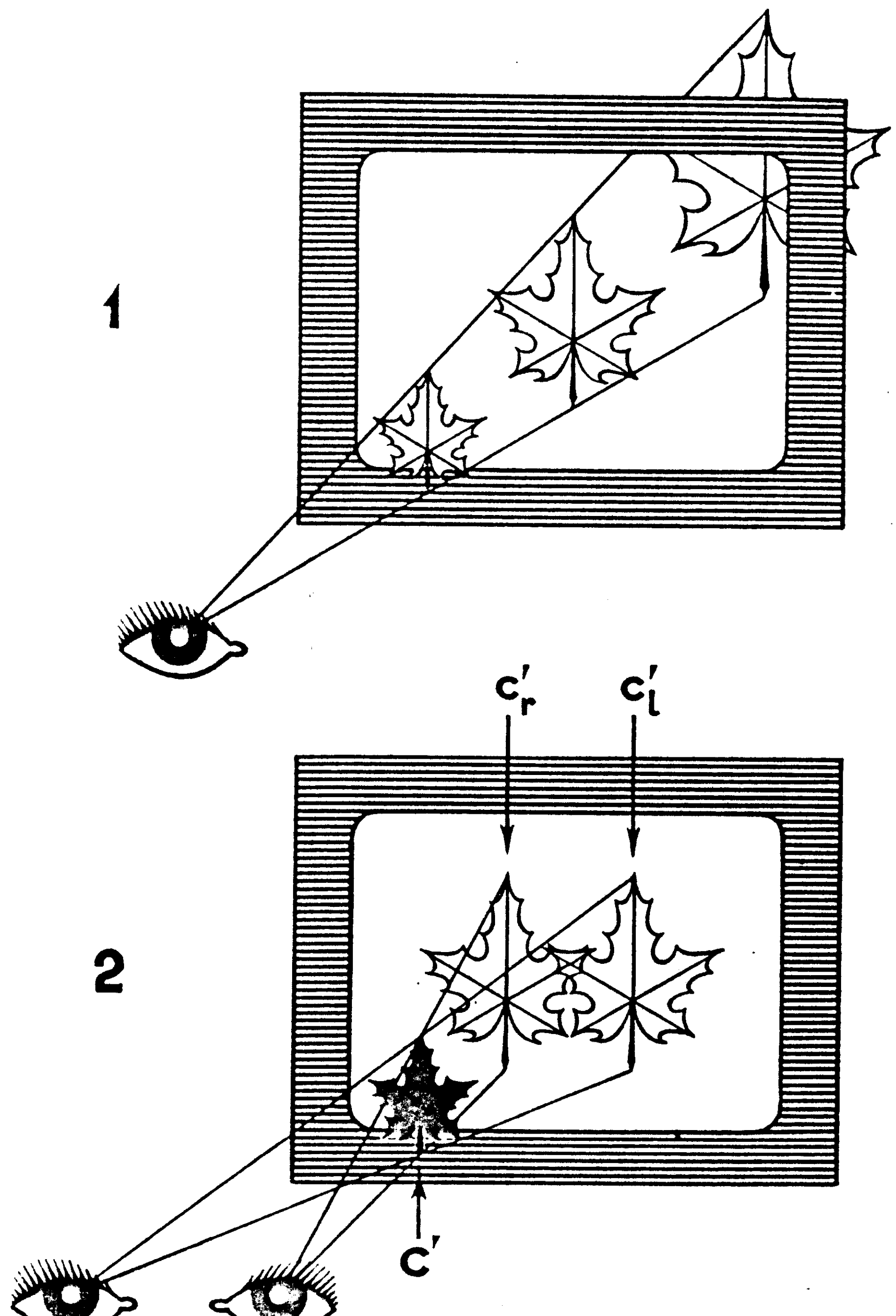


Figure 26.

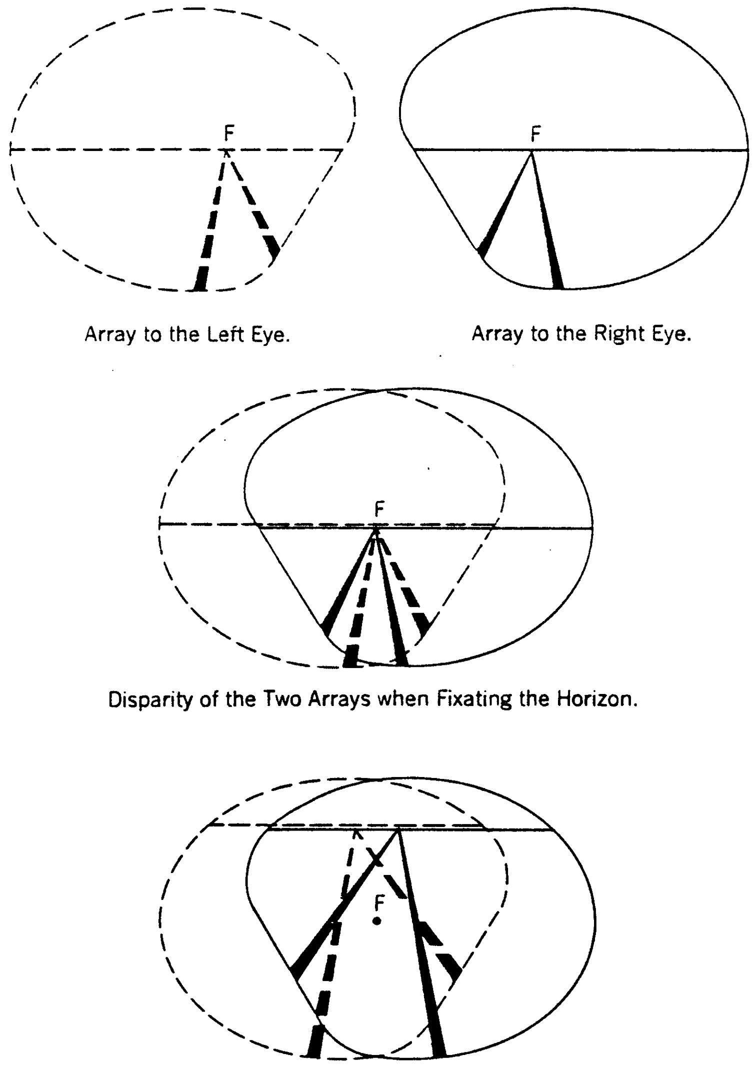
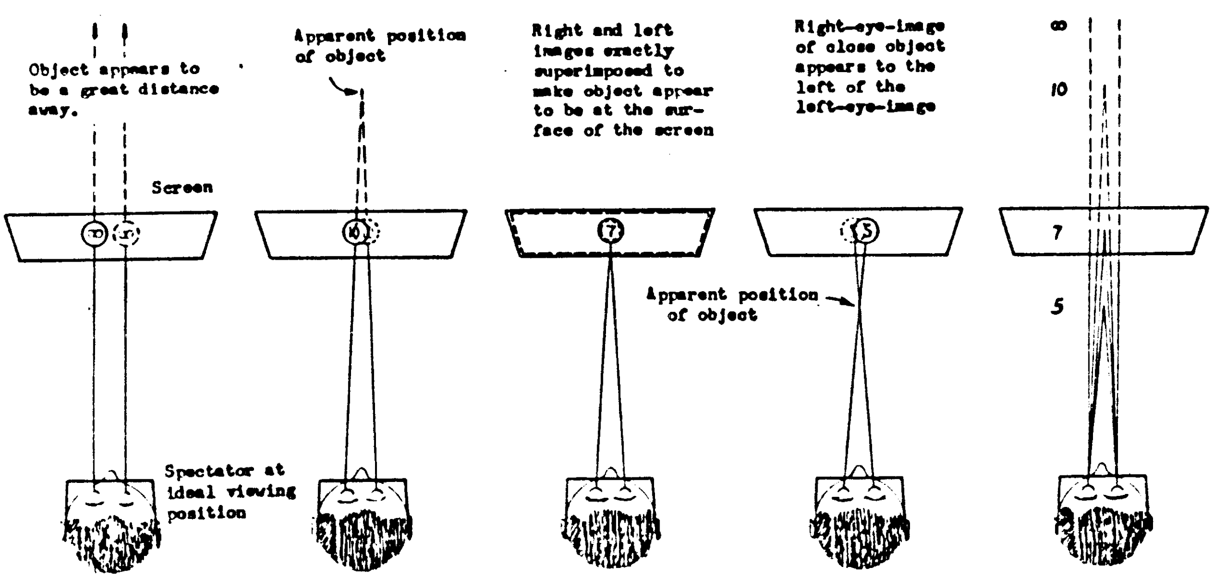


Figure 27.

Figure 28.



Finally, viewer translation in the z axis yields motion perspective [21] and is equivalent to 'dolly' shots in cinematography (not to zoom). [35] See Figure 24 . As mentioned, the plane of the screen is perpendicular to this axis and the projection size of the imaged scene is a function of the viewer's distance from that plane. As the viewer moves along this axis, object size relationships are formed. The displayed image appears lifesize when viewed from a position that is equal to the product of image magnification and the focal length of the original taking lens.

Viewpoint = M x F.

See Figure 25 .

At this position, the projected image will subtend the same angle of vision as for the camera in the original scene. When combined with binocular parallax cues (i.e., a stereoscopic image), the result is 'orthostereoscopic'. [44] Maintaining a constant magnification factor (size of screen) means that the proper viewing position for a moving observer is proportional to the focal length of the taking lens. As the viewer approaches the screen, a shorter focal length is needed to preserve the effect. Gibson, in researching visual cues necessary in flight simulators, is responsible for differentiating between motion parallax and motion perspective cues. [22].

The main task, then, of this display is to record and store sufficient visual information about a given environment so that these three cues are readily accessible. This is done by photographing every possible viewpoint within a given viewing area and storing them as viewpoint arrays. Description of methodologies follows.

4.3 DISPLAY CONFIGURATION

The process involved in implementation of this display falls in three main parts:

1. Recording the viewpoint arrays
2. Editing and formatting the arrays onto optical videodisc

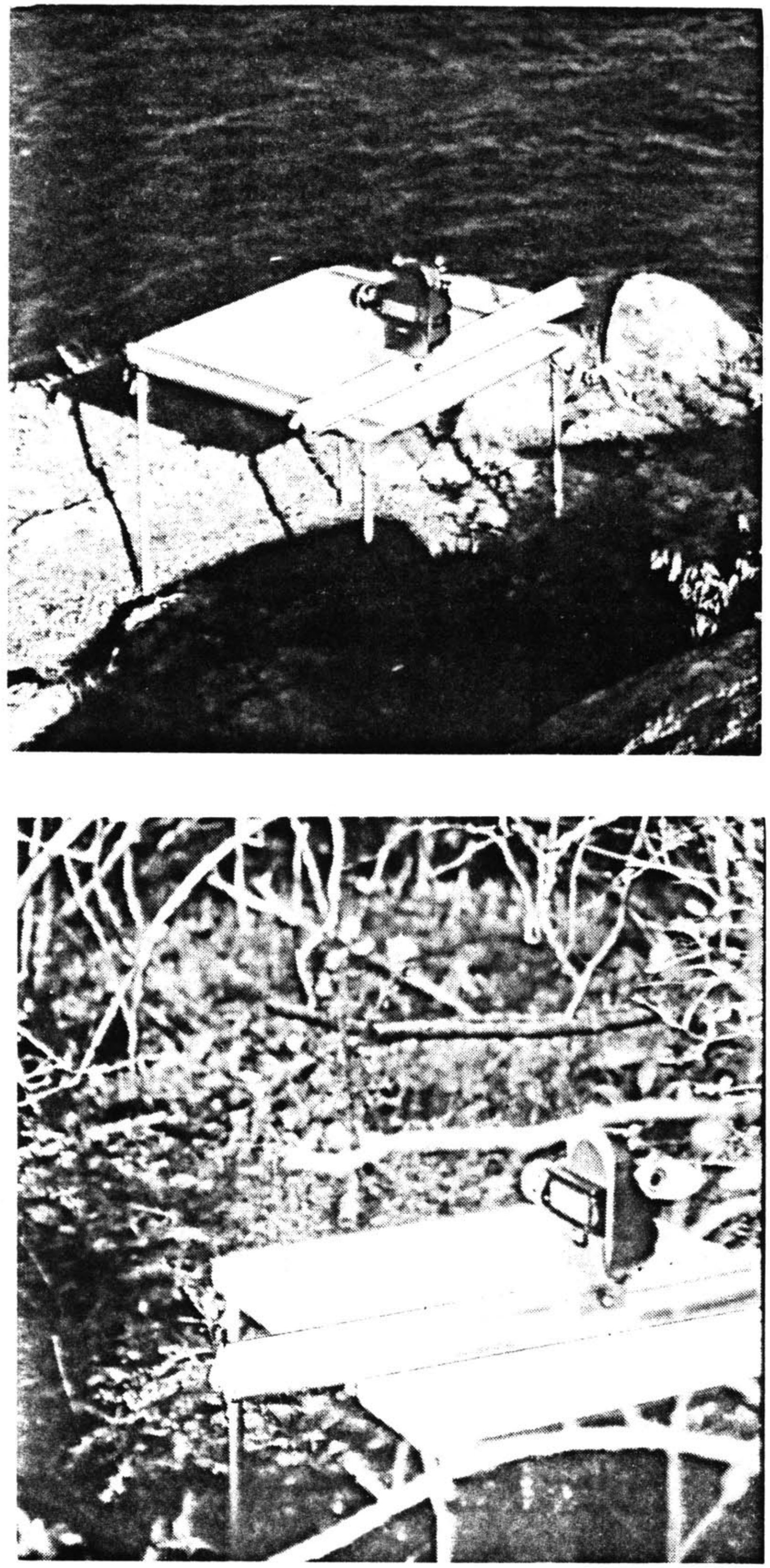
and 3. Programming the hardware configuration for display playback.

1. Recording the viewpoint arrays.

At a given scene, a mechanical camera track is positioned so that a 16mm movie camera is shuttled along the track in the x, y and z axes. One frame of film is shot at predetermined intervals until the complete matrix of viewing positions is recorded.

The first configuration of this track consists of a 1 meter long optical bench placed on a level table surface. See Figure 29. A Bolex movie camera is mounted on the track and triggered by a microswitch released at set intervals. The first set of image arrays is shot at 1cm intervals over a distance of 90cm, translating in x (i.e., horizontally parallel to the screen). Then the track is moved back in

Figure 29.



the z axis and, remaining parallel to the scene, the camera is again shuttled along the x axis. This is continued at 2cm intervals for a distance of 90cm in the Z axis. After each translation in x, the array is slated. The height of the array ia a constant 1 meter, due to movement inability in the y axis. The resulting x,z array is a total of 4,050 frames (equivalent to about 80 feet of 16mm film). Various shooting algorithms are tried, in an attempt to determine the most useful: the density of the array is changed by varying the camera trigger intervals (greater intervals for a less dense array also will result in an increase of apparent translated speed in the final display).

Scenes are also shot translating in the z axis with secondary shifts left to right along the x axis. Test arrays are also shot to evaluate parallel lens axis versus lens 'toe-in' information. Also, effects of changing focal lengths of the lens proportional to z translation are investigated.

A second set of image arrays is being shot on a larger track frame from which the camera is suspended. The shooting area is extended to a 3 meter square area with a 1 meter capability in Y. The frame construction is lightweight channel aluminum track supporting a camera shuttle on plastic roller skate wheels. The camera trigger mechanism is a photoelectric-pickoff switch similar to a bar code reader. Density of the image arrays is significantly increased and will be on the order of 20,000 frames from every level of y. A full array from a standing and sitting position will therefore almost fill one side of an optical videodisc.

2. <u>Videodisc formatting</u>.

When the viewpoint arrays are recorded, the 16mm film is edited and transferred to 2" videotape, from which an optical videodisc master is made. Several identical copies of the disc are then pressed. Use of the videodisc medium offers several important advantages:

> Storage density is 50,000 still frames per side, at relatively low cost.
>
> The frames can be randomly accessed at a worst case rate of 3 to 4 seconds. Under computer control, multiple discs may be used to reduce formatting problems.

A disadvantage in this application is that arrays must be stored linearly as opposed to their original 3D matrix. Because some information can be accessed quicker than other, a priority of access must be determined. In this display, it is first assumed that viewer motion in the x axis is most important. Yielding motion parallax and binocular parallax motion in the z axis is next; then y information. Therefore the arrays are formatted on disc as originally shot. This allows the viewer to translate at 30 frames per second with short search time for movement in direction of z or y. If three discs could be mastered with different sequences, each could be formatted around one axis priority. In playback, movement in x would be shadowed by the other two discs pre-cued to cut to a 30 fps translation in those axes. Search time is eliminated.

3. Playback configuration.

This display is essentially an interactive movie that is accessed spatially rather than temporally: 5,000 frames of a normal film runs about 3 minutes. Here an image of 5,000 frames offers as many viewpoints. The crux of this display is to match up these viewpoints stored on optical videodisc with the observer's position. The technologies used to achieve this are as follows.

The observer's position is tracked by a low frequency magnetic body tracking device manufactured by Polhemus Navigational Sciences, Inc. [62] The tracking range from the magnetic field source to the sensor worn on the user's head is approximately a 4 foot radius hemisphere. It operates at 40 points per second and reads three position coordinates as well as three degress of attitude. Values are then stored in the mainframe computer. See Figure 30 and Figure 31 .

A database is created for the information on videodisc to match up the image arrays with these position coordinates by defining start and end points for image array sequences in a specific axis. As the viewer moves, the videodiscs are computer controlled to match the viewer's direction and

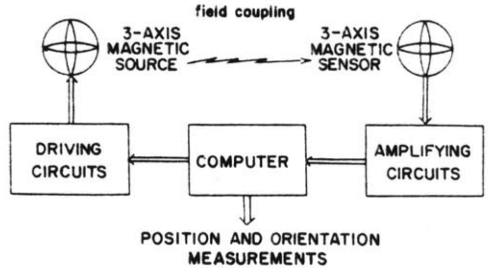


Figure 30.

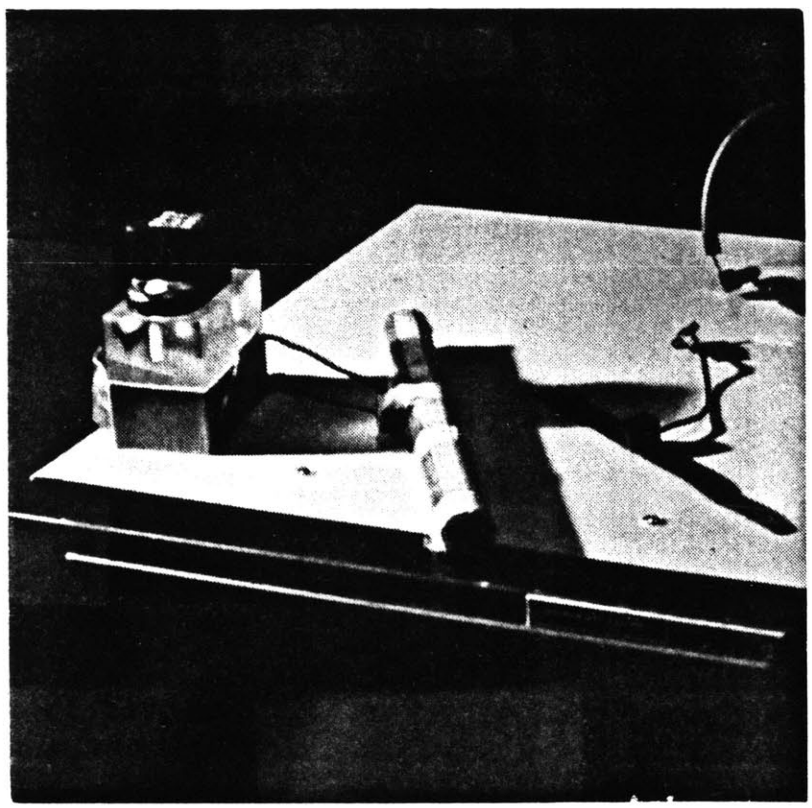


Figure 31.

speed (up to 30 fps). A typical rate is approximately 2-3 miles per hour moving in x.

The actual <u>display medium</u> can take several forms. Much of the imagery is shot to be viewed on a standard video monitor because of the small tracking range ability. Ideally, the display will be used with a large screen video projector in a rear projection format. In this configuration, the ability to present <u>lifesize</u> images is demonstrated. To represent viewer movement in the Z axis, the video projector should be adjusted to project an image 4 feet wide. Since the body tracking range is also 4 feet in z, the viewer's position is calculated to fall from 11 feet to 15 feet away from the screen. This distance, as mentioned, is the product of camera focal length and projected imaged magnification. With a constant magnification factor of 122, $(\frac{\text{screen width}}{\text{16mm film width}})$, at 11 foot viewing distance, a focal length of <u>28</u>mm is necessary to maintain proper perspective. A 13 foot viewing distance requires a 33mm focal length. A 15 foot viewing distance requires a 38mm focal length. The viewpoint array is shot as usual on the camera track with the added task of changing focal length for each new z value. According to determined shooting resolution in z (1cm→2cm), the total number of sequences is calculated and translated to appropriate focal length changes. The 28 to 38mm range of a 28-45mm zoom lens is divided into an appropriate number of positions to match the number of shooting points in z and changed at each repositioning.

For lifesize display sequences on a 10" video monitor, a similar process is used. Wtih a magnification factor of 25, closest viewing distance is 27" from the screen. These viewing sequences are shot with a 28mm focal length. Over a 3 foot range in z, a gradual focal length shift to 64mm is required at 63" from the monitor. The same procedure is followed to obtain the inbetween focal lengths.

As described so far, this system yields a '2D' display rich in motion parallax and motion perspective. These cues may provide sufficient information about depth relationships in the scene. But since each eye sees a slightly different viewpoint, we can slso use image arrays from the x axis to provide important binocular parallax cues. Two identical videodiscs are used - one for each eye. Frame input to the left eye from one disc is matched to the other eye by stepping forward on the second disc to a frame taken approximately 6cm from the first. The difference in these images matches the normal interocular distance of 6cm. At a shooting resolution of 1cm per frame, this is a 6 frame disparity. The discs are run conjointly with this constant disparity to each eye. Lesser or greater disparity may be easily set to accomodate various distances involved in the represented scene.

Viewing of this 3D image is by means of piezoceramic

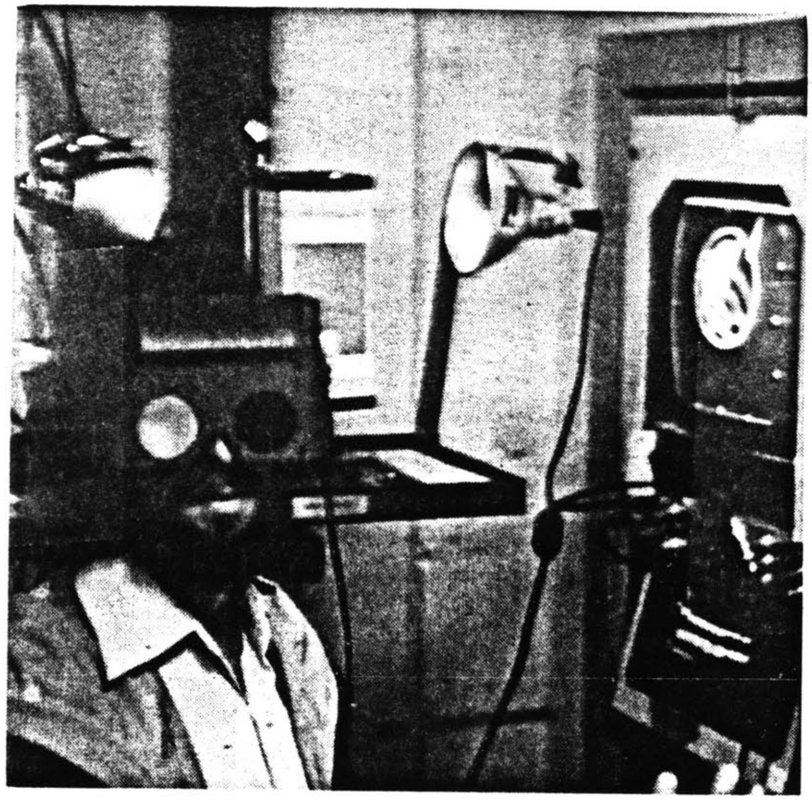
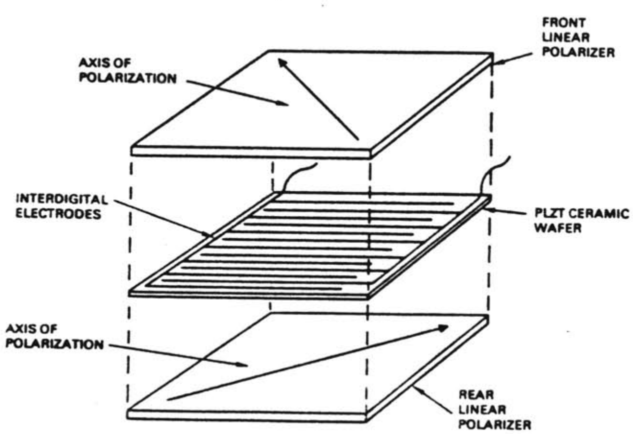
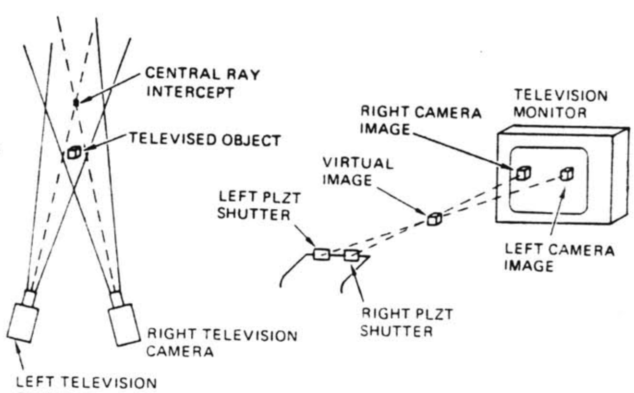


Figure 32.



PLZT electro-optic shutter assembly



Basic stereoscopic video display geometry viewing glasses, 'PLZT's', worn by the viewer. [62] See Figure 32. The two videodisc frames are mixed together to form one TV frame made up of the left eye image in the even interlace field and the right eye image in the odd interlace field. Each lens of the viewing glasses acts as a lightvalve, and opens and closes in synch with the fields of the TV frame every 1/60 of one second. As a result, each eye sees only its correct field and a stereoscopic image is presented. These viewing lenses are a sandwich of two orthogonally polarized filters around a piezoceramic wafer. In response to an electrical pulse, the PLZT wafer rotates the polarization of the incident light by 90 degrees, allowing transmission through the second polaroid filter. Light transmission is limited to about 20% and image resolution is one-half normal.

4.4 IMAGE CONTENT

Particular attention has been paid to the image content of scenes shot or scripted for the viewpoint dependent imaging display. Content ranges from scenes that are especially rich in response to viewer movement, to environments that are <u>'binocular specific'</u>, that is, cannot be adequately represented by 2D images. Simple examples of these are reflections, mirrors, or dense foliage. The <u>viewpoint specific</u> images make possible representation of other visual phenomena such as beat patterns of a fence with its shadow (as one moves along it) or flight simulator-like landing scenarios. Other phenomena such as camouflage revealed through movement or anamorphic images, spatially accessed, are relevant.

The object has been to build up a vocabulary of three dimensional, dynamic phenomena that cannot be represented in standard display environments. Specific content of first experiments are as follows:

1. Closeup arrays of a moving head. Variations are shot in which the eyes always follow the viewer as he moves through the viewpoint array. These arrays could also be programmed to simulate mirror reflection of the viewer's head movements.
2. Viewpoint arrays of an interior architectural space, filled with various objects.
3. Arrays of exterior landscape scenes and environmental phenomena.

Projected content to be recorded is:

1. Computer generated imagery
2. Medical imagery - CT scans
3. Time lapse phenomena activated by the observer's movement through the viewing space (similar to pixillation techniques)
4. Dynamic normal footage that can be accessed only from one specific viewpoint and which, when activated, runs at 30 fps.

Figure 33.

4.5 DEVELOPMENTAL PROBLEMS

Obstacles encountered and lessons learned in developing this display are an important consideration for future research in this area. Necessary refinements are mainly in the areas of viewpoint array shooting algorithms and overall hardware modifications.

Although shooting algorithms are partly dependent on the resolution and repeatability of the tracking device for the camera, more research must be done on the proper orientation of the camera lens axis relative to the scene. The problem encountered is in later placement of the 3D 'window' that the viewer looks through in the 3D version of the display. Figure 35. In viewing through a real window, the left and right eyes see disparate 2D areas of view flanking the central 3D overlap area. In viewing a stereoscopic image, these 2D areas are reversed and somewhat disorienting. The amount of non-overlapped, 2D area is a direct function of lens axis orientation to scene. Footage for the display is shot predominately with lens axis perpendicular to the scene always parallel to the successive shot. Scenes with little forground content are little affected by this while closeup shots tend to exaggerate this reversal effect. Solutions to this would be either to mask these non-overlapping 2D areas out of the image, or mask the screen with a kind of proscenium to simulate the 'real' window situation for the viewer.

An alternative solution is to 'toe-in' the axis of the taking lens for every shot such that all converge on a central point within the chosen scene. For example, the shooting array for a closeup head sequence would have every frame converged on the nose. Implementation of this requires modification of the camera track to also rotate the camera around its vertical axis at each increment along the x axis. But although this procedure will eliminate the original window problem, a new problem is introduced by the resulting trapezoidal distortion of any places in the scene parallel to the x axis; when presented to each eye, stereo fusion is difficult. A suitable solution requires more extensive testing.

Hardware modifications are needed in each phase of the display configuration and are founded in the exclusive use of prototype or state of the art equipment.

As mentioned, the viewpoint array shooting algorithms are dependent on the resolution and repeatability of the camera track. The two systems in use, the optical bench and the larger xyz frame, are sufficient in resolution with some improvement necessary in the camera triggering mechanism. A more accurate photoelectric pickoff switch is projected to increase the shooting speed along the track while offering easily adjustable resolution. The calibration tapes can be easily substituted. Another problem encountered is portability of the tracking systems. A large scale xyz frame that could be easily disassembled and transported with sufficient stability would be desireable. Precise repeat-

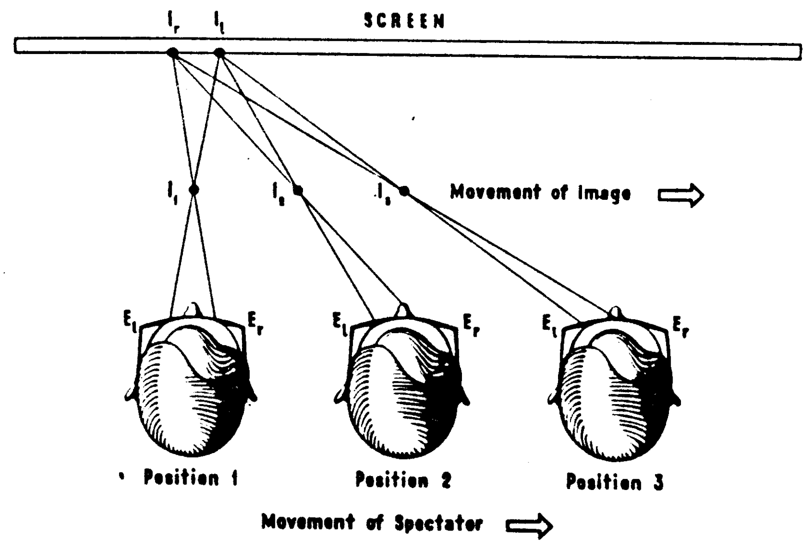


Figure 34.

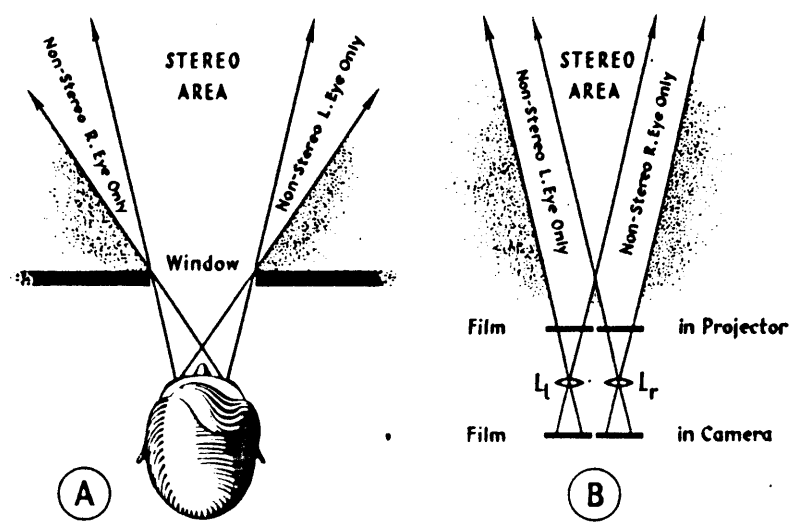


Figure 35.

bility of camera motion along the track in each direction is also necessary to facilitate different disc formatting options. Currently, the display is run with two identical discs assuming priority for the viewer's translation in x. As mentioned in Section 4.3, if several more discs were used, each formatted with a different axis priority, these could be cued to insert the proper sequences without screen blanking.

To shoot these arrays, each scene would be done three times in exact register, each time with a different axis priority:

1. Once with continuous translation in x, moving back in z and then up in y.
2. Once with continuous movement in z, moving across x and then up in y.
3. Once with continuous movement in y, moving back in z and then across in x.

A change in user motion from translating in x to continuous motion in z would cut to disc z and continue without distracting screen blanking from search intervals. The camera support to do this would have to be a computer controlled plotter similar to the 'ACES' system used by the Disney animation/special effects studios. This enables precise registration and programming abiltiy for varied camera paths. In addition, such a system would allow more comprehensive viewpoint arrays that could respond to a viewer's head attitude such as tilt or elevation as well as to xyz position information.

The major problem in this display is inability to quickly access three dimensional images stored in the almost two dimensional videodisc medium. The suggested use of many discs formatted in various priorities is an awkward solution. What is, in fact, necessary is a 3D storage medium in which a frame of the viewpoint array could be imbedded spatially correspondent to its position when filmed. A partial solution could use a system similar to the Thompson CSF videodisc, which uses a transparent disc and refocusable laser readout mechanism. Both sides of the disc can be accessed by only refocusing the laser. For this display purpose, the discs could be stacked and multiple reading heads utilized, to access a 3D matrix of viewpoint information. In the future, a more satisfactory solution might use a 3D information matrix built up of liquid crystal elements.

Another area for hardware refinement is the viewer tracking system. At present, the Polhemus unit installed at the Architecture Machine Group has a maximum range of 1 meter from source to sensor limiting the viewer in this display mostly to gross head movements. Ideally the system should enable whole body movement. The viscerality of walking into a scene is key to this display system. The ability to present the system on one 8' by 10' projection wall of a media room is projected. To adequately cover this area, the tracker range must be at least a 10' radius of a hemisphere projecting from the center of the room's ceiling. Recent developments indicate that this is feasible. A final goal is to create a one to one correspondence between viewing space and the virtual space of the projected images.

The mechanism for viewing the display in 3D alternately presents disparate viewpoints to each eye by mixing output from each disc into a single interlaced image. This signal is decoded by viewing glasses that are slightly heavier than normal glasses and also serve as a support for the body tracking sensor. One inherent disadvantage is the low light transmittance through the lens sandwich of tow polaroid filters and piezoceramic wafer. Total transmittance is 17-20% and requires that brightness be greatly increased on viewing monitor or projector. A solution to this problem is to move the PLZT wafers from the viewer's eyes to the lens of the projector. In this configuratin, only one polaroid filter is needed in conjunction with the PLZT wafer. The projected image would be polarized in the orientation of the polaroid filter for one video field and rotated 90° through the PLZT wafer for the next field. The viewer need only wear standard polarized 3D viewing glasses free from uncomfortable electronics and high voltage on the head. The size of this PLZT wafer to fit over the projection lens should be 4", and is only currently being developed. Usual wafer size is about 2" diameter. The type of video projector used with these glasses is also critical. Experiments with a GE light valve have been unsuccessful because of slight image depolarization from the valve's oil bath and/or because of long image persistence (i.e., longer persistence than phosphor on a monitor). Yet experiments with Sony and Advent projectors are successful.

Another critical problem with the PLZT viewing system is the reduced image resolution. Since left and right eye images are interlaced in one frame, resolution is cut by one half. A simple, yet costly, solution would be to use a 1000 line resolution display. A better alternative is to modify the mixing characteristics of the device to alternate fields displayed each 1/60th of a second. (Presently one field is thrown out in the mixing process, but is recoverable.) In this situation, for example, the left eye would see field one of its image, right eye would see field one of its image, then left eye would see field two and right eye field two, although this requires that each frame is looked at for greater than 1/30th of a second (i.e., at least two frames).

A final alternative would be to forego use of the PLZT viewers and instead use two light valve projectors. Each would be orthogonally polarized and their images superimposed on the projection screen. The viewer only wears matching polarized 3D viewing glasses with tracking sensor to yield a display with full resolution and brightness. This configuration also enables adjustments for the window problem previously mentioned. By 'toe-in' of the projectors, the

the 3D display 'window' can be manipulated. Objects in the scene which are exactly superimposed will fall in the plane of the screen, while other disparate objects will appear behind or in front of the screen. This is a critical adjustment not possible with the PLZT viewing system. (Figure 28.)

4.6 FEATURE SUMMARY

Most of the display 'features' stated in OBJECT OF RESEARCH (Section 3.1.) are not new display elements. Preceeding research on stereoscopic displays, random access displays and lifesize-position dependent displays (separately or in combination) has been cited in Section 3. On the other hand, none of these displays, alone or together, have yet become standard features in our traditional film and video media formats.

This research is an attempt to bridge the gap between awkward, costly research prototypes and sophisticated interactive display systems that are elegant, technically and economically feasible, and easy to use. Relative to efforts of the 60's and 70's, this synthesis of state of the art technologies in the viewpoint dependent imaging system is significant.

There are also several new display capabilities unique to this system that should be mentioned.

The optical videodisc mastered for this display is the first with exclusively stereoscopic content. More importantly, the lowcost, high density storage capabilties of the videodisc make 3D information parallax possible in the vertical axis as well as horizontal. Coupled with increased computer power, its rapid access time of rasterscan images far surpasses previous 3D vectorgraphic displays in similar environments.

Another important feature of this display is the relatively low bandwidth required for transmission. A major obstacle to widespread use of 3D TV previously has been the high bandwidth required. For a simple, good resolution, color 3D display configuration, at least twice normal bandwidth is required for use of two channels. In more sophisticated 3D displays such as holograms, every possible viewpoint would be transmitted at a cost of enormously high bandwidth. Okoshi's bandwidth calculations for a variety of 3D displays is shown in Figure 36. [57] In this display, by only transmitting the correct viewpoint for a user's position relative to the screen at any given time, the bandwidth remains equal to that of one normal broadcast channel. Tracking the viewer eliminates redundant and irrelevant information.

Means for continuously updating a 3D display for an active observer provides a solution to an insistent problem in all other 3D presentations, still and dynamic. Commenting on a proposal for this display, Dr. Richard Bolt observes:

TABLE 7.3 Bandwidths Required to Transmit Three-Dimensional Images Via Television[a]

	B (MHz)
Holography	
Two dimensional	
On-axis reference	30 (N=500)
Off-axis reference	120 (N=500)
Three dimensional	
Eye-piece type	6000
Wide viewing zone	1,500,000
Reduced-information	600,000 (N=500)
Integral photography	42,000 (N=500)
Multiple photography (unidirectional)	750 (N=500)
Unidirectional holography	3000 (N=500)

[a] f_F = 30 pictures/s, λ = 500 nm, Ω = 0.2, a = 200 cm, b = 20 cm, and w = 40 cm.

Figure 36.

"Stereo tends to reduce, do away with the specific awareness of the picture plane, knowledge of which by the viewer provides corrective information re: object relationships in picture: thus, if we introduce 3D stereo, and observer is free to roam about, then we need Polhemus body tracking to supply information for the proper projection point. I.E., it is less a luxury to body track than a necessity, when the observer can move about."

As A viewer moves in front of a typical 3D image, it appears to have reversed motion parallax. [67] The scene appears to pivot around the plane of the screen - making far objects move opposite to the viewer's direction of movement and near objects follow. See Figure 34. Because the viewpoint dependent imaging system continuously corrects for viewer translation, correct motion parallax is always perceived.

Along similar lines, the presentation of lifesized imagery is usually limited to one correct viewing position. In this display, the ability to quickly update the image maintains the effect as the user moves about. This is a key factor in creating a virtual window through which a user can explore a three dimensional image space.

4.7 APPLICATIONS

Aside from obvious simulation applications, this research should be considered as groundwork for development of more sophisticated interactive three dimensional image environments.

By substituting a remotely controlled stereo camera system for the videodisc stored viewpoint arrays, dynamic imagery would be possible, thus enabling real time exploration of inaccessible environments while under direct viewer control.

With multiplexing capabilities, i.e., several tracking and input stations, more than one user could access personalized or viewpoint specific images on a <u>common</u> display surface. In this configuration, the piezoceramic glasses are modified to present different viewpoints to each of several users rather than to each eye of a single user.

Finally, this display provides a foundation for development of a viewpoint dependent imaging system that is non-screen dependent. Through use of a headmounted display and eyetracking technologies, the user will have access to a virtual 360° visual surround.

4.8 FUTURE DEVELOPMENTS

In contrast to the refinements in immediate technology suggested in Section 3.5., significant future developments are pending technologies just emerging or, as yet, non-existent. Relevant areas of research are as follows.

In addition to the major visual depth cues of motion parallax and binocular disparity, there are two more important cues that should be implemented in 3D display environments. Accomodation, the focusing ability of the eyes, and convergence, directing each eye at a common target, are closely linked depth cues; the eyes always converge at the point of focus. But, in most 3D displays, objects appear to be far in front of or behind the display surface. In this situation, the eyes focus on the screen but are converged at the virtual position of the object, resulting in a subtlely unnatural visual experience. The effect is more critical for viewing distances less than 2 meters and 5 meters respectively. [30] Infinity optics used to collimate images in head mounted displays are one solution. Rectifying the problem for large projection surfaces is more complicated. This necessitates some kind of aerial image projection again by collimating optics or possibly concave mirror surfaces.

Another problem pending emerging technologies is rapid access of information for the 3D matrix of stored viewpoint arrays. Although a 3D storage medium as mentioned in Section 3.5 would be sufficient, an alternative solution under development involves computer controlled image interpolation. Instead of shooting a high density matrix of viewpoints, only a few 'boundary' views are taken. 'Inbetween' views are then synthesized from these 'boundary' views and the interpolation process is repeated on the fly as needed. [58]

A third critical goal for future development is the ability to access dynamic images. Aside from artifactual pixillation effects, the image in this display has been primarily static. A system could be devised where each viewpoint in the image array is tied into an auxilliary storage of dynamic imagery. When the user stops, the system cuts to the pre-cued dynamic discs. Action proceeds only when the user has decided on a propitious viewpoint. Shooting algorithms and the huge bank of storage units appear formidable at this time.

The most probable resolution for dynamic imagery would be to replace the videodisc input by live camera input that is controlled remotely.

> "...PERSONALIZED TELEVISION SAFARIS. When you can have a high quality cinema display in your own home, there will certainly be global audiences for specialized programs with instant feedback from viewer to cameraman. How nice to be able to make a trip up the Amazon, with a few dozen uknown friends scattered over the world, with perfect sound and vision, being able to ask your guide questions, suggest detours, request closeups of interesting plants or animals--- in fact, sharing everything possible except the mosquitoes and the heat." [9]

5. CONCLUSION

This display system makes use of available, state of the art media technologies to evaluate visceral viewer involvement as an essential element in virtual representation of visual environments.

The capability for a user to control viewing position proves to be an important interactive feature for exploring a virtual environment. Usually relegated to the realm of perceptual 'nuance', the visual depth cues of motion parallax, binocular parallax, and motion perspective, offered in combination, are central to the successful implementation of this interaction.

An important consideration is that not everyone desires control over what their TV is looking at; i.e. has a point of view. The emphasis here on non-programmed media and greater degree of viewer immersion in the display should not be interpreted as a suggestion to replace traditional imaging techniques, but rather as a necessary balance. Seeing from other points of view is even more informative when a personal viewpoint is established.

REFERENCES

1. Bateson, G. Mind and Nature, New York; Bantam Books, 1979.

2. Bender, T. Environmental Design Primer, Minneapolis; 1973.

3. Benjamin, W. "The Work of Art in the Age of Mechanical Reproduction"; Illuminations, New York; Harcourt, Brace & World, Inc., 1955.

4. Bolt, R.A. Spatial Data Management, Cambridge; MIT Architecture Machine Group report, 1979.

5. Bradbury, R. "The Veldt"; The Illustrated Man, New York; Bantam Books, 1951.

6. Bunker, W. "Training Effectiveness Versus Simulation Realism"; SPIE, Volume 162; Visual Simulation & Image Realism, 1978.

7. Carpenter, E. They Became What They Beheld, New York; Ballantine Books, Inc., 1970.

8. Carpenter, E. Oh What a Blow That Phantom Gave Me!, New York; Holt Reinhart, Winston, 1972.

9. Clarke, A. "Communications in the Second Century of the Telephone"; Technology Review, MIT; May, 1976.

10. Collender, R. "True Stereoscopic Movie System Without Glasses"; Information Display; July, August, September, & October, 1968; September, October, November & December, 1972.

11. Conder, J. Landscape Gardening in Japan, New York; Dover Publications, Inc., 1964.

12. Devich, R., and Weinhaus, F. "Image Perspective Transformations"; Imagery Data Systems; ESC, Inc.

13. Ebeling, W. "Fundamental Limitations in Visual Stimulation", New York; Singer-Link Division.

14. Epstein, ELJ. News From Nowhere, New York; Random House, 1973.

15. Ernstoff, M. "A Head-up Display for the Future", Proceedings of the SID, Volume 19 (4), 1978.

16. Farber, J., and Rosinski, R. "Geometric Transformations of Pictured Space"; Perception, Volume 7, 1978.

17. Fisher, S. "Eyetracking, 3D and Direction Selective Imaging in Interactive Display"; Cambridge, MIT Architecture Machine Group, 1980.

18. Fishlock, D. Man Modified, New York; Funk & Wagnalls, 1969.

19. Gabor, D. "Three Dimensional Picture Projection"; U.S. Patent 3,479,111; November 18, 1969.

20. Gerstein, M.S. GerDekkers, New Dutch Landscape, Cambridge, MIT Hayden Gallery, 1979.

21. Gibson, J.J. The Perception of the Visual World, Boston; Houghton Mifflin Co., 1950.

22. Gibson, J.J. et al. "Parallax and Perspective During Aircraft Landings"; American Journal of Psychology, Volume 68, 1955.

23. Gibson, J.J. The Senses Considered as Perceptual Systems, Boston; Houghton Mifflin Co., 1966.

24. Gibson, J.J. "The Information Available in Pictures"; Leonardo, Volume 4, 1971.

25. Gibson, J.J. "Visualizing Conceived as Visual Apprehending Without Any Particular Point of Observation"; Leonardo, Volume 7, 1974.

26. Gombrich, E.H. Art and Illusion, New York; Bollingen Foundation, 1960.

27. Hall, E.T. "Art, Space and the Human Experience"; Arts of the Environment, New York; George Braziller (G. Kepes, ed.), 1972.

28. Hall, E.T. The Hidden Dimension, New York; Doubleday & Co., Inc., 1966.

29. Harvey, J. "Current Trends and Issues in Visual Simulation", SPIE, Volume 162, Visual Simulation & Image Realism, 1978.

30. Hatada, T., Sakata, H., and Kusara, H. "Psychophysical Analysis of the 'Sensation of Reality' Induced by a Visual Wide-Field Display"; SMPTE Journal, Volume 89, 1980.

31. Heilig, M. Sensorama Simulator; U.S. Patent 3,050,870.

32. Herman, S. "Principles of Binocular 3D Displays With Applications to Television"; SMPTE Journal, Volume 80 (7), 1971.

33. Hochberg, J. "The Representation of Things and People"; Art, Perception and Reality, Baltimore; Johns Hopkins University Press, 1972.

34. Hochberg, J. "Art and Perception"; Handbook of Perception, Volume X, Perceptual Ecology; Cortevette, E., and Friedman, M. (ed.); New York; Academic Press, 1978.

35. Hochberg, J. "The Perception of Motion Pictures", Handbook of Perception, Volume X, Perceptual Ecology; Cortevette, E., and Friedman, M. (ed.); New York; Academic Press, 1978.

36. Hooper, K. "Perceptual Aspects of Architecture"; Handbook of Perception, Volume X, Perceptual Ecology; Cortevette, E., and Friedman, M. (ed.); New York; Academic Press, 1978.

37. Horzu, P. "Two Headed TV Set"; Popular Science, March, 1954.

38. Ittelson, W.H. The Ames Demonstrations in Perception, Princeton; Princeton University Press, 1952.

39. Ittelson, W.H. Visual Space Perception, New York; Springer Publishing Company, Inc., 1960.

40. Ives, H. "A Method of Projection in Relief and Color"; JOSA, Volume 22 (4), 1932.

41. Ivins, W.M. Prints and Visual Communication, Cambridge; MIT Press, 1953.

42. Krauss, R.E. Passages in Modern Sculpture, Cambridge; MIT Press, 1977.

43. Kubitz, W. and Poppelbaum, W. "Stereomatrix Interactive Three Dimensional Computer Display"; Proceedings of the S.I.D., Volume 14 (3), 1973.

44. Kurtz, H. "Orthostereoscopy"; Journal of Optical Society of America, Volume 27 (10), 1937.

45. Lavin, M. "Analysis of Scenes From a Moving Viewpoint"; PhD Thesis, Artificial Intelligence Laboratory, MIT, 1977.

46. Leeman, F. Hidden Images, New York; Harry N. Abrams, Inc., 1975.

47. Lerner, E. "Computers That See", IEEE Spectrum, October, 1980.

48. Lippman, A. "Movie Maps: An Application of the Optical Videodisc to Computer Graphics", Cambridge; MIT Architecture Machine Group report,1980.

49. Lipton, L. "Sensorama"; Popular Photography, July, 1964.

50. Merleau-Ponty, M. Phenomenology of Perception, London; Rautledge & Kegan Paul, 1962.

51. Minsky, M. "Telepresence"; Omni, Volume 2 (9), June 1980.

52. Negroponte, N. The Architecture Machine, Cambridge; MIT Press, 1970.

53. Negroponte, N. "Return of the Sunday Painter"; The Computer Age: A Twenty Year View, Cambridge; MIT Press, Dertouzos, M. and Moses, J., ed., 1978.

54. Negroponte, N. "The Impact of Optical Videodiscs on Filmmaking", Cambridge; MIT Architecture Machine Group report, 1980.

55. Negroponte, N., Lippman, A., and Bolt, R. "Transmission of Presence", Cambridge; MIT Architecture Machine Group proposal to Cybernetics Technology Division of Defense Advanced Research Projects Agency, 1980.

56. Oharek, F. and Harvey, J. "Component Performance of a 360^{o} Nonprogrammed Visual Display"; SPIE, Volume 162, Visual Simulation and Image Realism, 1978.

57. Okoshi, T. Three Dimensional Imaging Techniques, New York; Academic Press, 1976.

58. Okoshi, T. and Ohira, K. "Synthesis of an Autostereoscopic 3D image from Binocular Stereoscopic Images"; Applied Optics, Volume 18 (4), 1979.

59. Papert, S. "Computers and Learning"; The Computer Age: A Twenty Year View, Dertouzos, M. and Moses, J., ed., Cambridge; MIT Press, 1979.

60. Perkins, D.N. "Pictures and the Real Thing"; Project Zero, Harvard University, Cambridge, 1979.

61. Pirenne, M.H. Optics, Painting and Photography, Cambridge; University Press, 1970.

62. Raab, F.H. et al. "Magnetic Position and Orientation Tracking Systems"; IEEE Transactions in Aerospace and Electronic Systems, Volume AES 15 (5), 1979.

63. Richie, D. The Films of Akira Kurosawa, Los Angeles; University of California Press, 1970.

64. Roese, J.A. "Applications of PLZT Electro-Optic Shutter Stereoscopic Displays"; SPIE, Volume 199, Advances in Display Technology, 1979.

65. Sanders, B. "Stereoscopic Drawing by Computer-Is It Orthoscopic?", Applied Optics, Volume 7 (8), 1968.

66. Shepard, P. Man in the Landscape, New York; Ballantine Books, 1967.

67. Spottiswoode, R. and Spottiswoode, N. Stereoscopic Transmission, Los Angeles and Berkeley; University of California Press, 1953.

68. Sutherland, I. "A Head-Mounted Three Dimensional Display", AFIPS. Volume 33 (pt. 1), Fall Joint Computer Conference, 1968.

69. Sutro, L., and Lerman, J. Robotvision, Cambridge; Charles Stark Draper Laboratory, 1973.

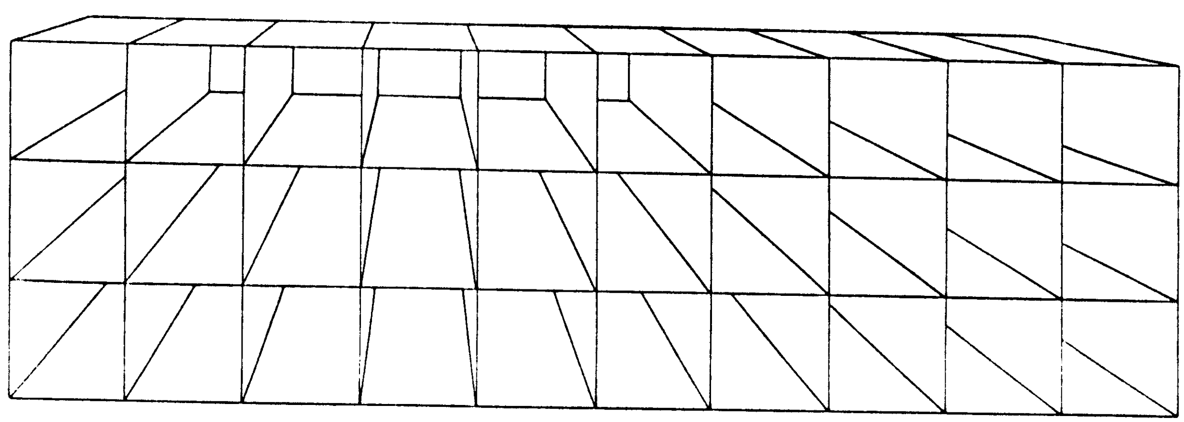
70. Tiltman, R.F. "How 'Stereoscopic' Television is Shown"; Radio News, 1928.

71. Tyrwhitt, J. "The Moving Eye"; Explorations in Communication, Carpenter and McLuhan, ed.; 1966.

72. Vickers, D. "Head Mounted Display Terminal"; IEEE: ICGS, 1970.

73. Vickers, D.L. Sorcerer's Apprentice: Headmounted Display And Wand, University of Utah; Doctoral Dissertation, Department of Electrical Engineering, 1974.

74. Zelazny, R. "Home is the Hangman" <u>My Name is Legion</u>. New York: Ballantine Books , 1975.



7. ACKNOWLEDGEMENTS

Thanks to:

Dr. Richard Bolt,

Professor Steven Gregory,

Professor Andy Lippman,

All at the Architecture Machine Group,MIT;

To John Borden, Peace River Films;

and particularly to Nicholas Negroponte for his enthusiastic support, innovative direction and especially for the accessibility of the Architecture Machine.