August 8, 2003

A View-Sequential 3D Display

Oliver S. Cossairt

B.S. Physics
The Evergreen State College, 2001

Submitted to the Program in Media Arts and Sciences, School of Architecture and Planning
in Partial Fulfillment of the Requirements for the Degree of
Master of Science in Media Arts and Sciences

at the

Massachusetts Institute of Technology

September 2003

Signature of Author:______________________________
Oliver Cossairt
Program in Media Arts and Sciences
August 8, 2003

Certified by:______________________________
Dr. Stephen A. Benton
E. Rudge ('48) and Nancy Allen Professor of Media Arts and Sciences
Program in Media Arts and Sciences
Thesis Supervisor

Accepted by:______________________________
Andrew Lipman
Chairperson
Department Committee on Graduate Students
Program in Media Arts and Sciences

A View-Sequential 3D Display

Oliver S. Cossairt

Submitted to the Program in Media Arts and Sciences School of
Architecture and Planning on August 13, 2003
in Partial Fulfillment of the Requirements for the Degree of
Master of Science in Media Arts and Sciences

Abstract

This thesis outlines the various techniques for creating electronic 3D displays and analyzes their commercial potential. The thesis argues for the use of *view-sequential* techniques in the design of 3D displays based on considerations of state of the art 3D displays and currently available technology. A view-sequential 3D display was built, which is described in detail. The performance of this display is given, as well as an analysis which characterizes its anticipated performance and comparison with measured results.

Thesis Supervisor: Stephen A. Benton
Title: E. Rudge ('48) and Nancy Allen Professor of Media Arts and Sciences

A View-Sequential 3D Display

By Oliver S. Cossairt

Accepted by:
Dr. Adrian Travis
Lecturer
Cambridge University Engineering Department

Accepted by:
Dr. V. Michael Bove
Principal Research Scientist
MIT Media Laboratory

Acknowledgments

This research is based on the work conducted at Cambridge University. The idea of view-sequential 3D displays originated from the work that was done there. The goal of this research was to build an improved system and characterize the improvements. This thesis presents an original implementation of a 3D display technique which has been thoroughly investigated already. The technique is revisited to demonstrate improvements in display prototypes that are the result of advances in 2D display technology.

I dedicate this work to Christian, Stephanie, and Dr. Benton, all of whom I have relied heavily on in the process of completing my degree. Thank you all. My parents deserve much credit also, as they have always supported me, no matter what I get myself into. This work would not be possible without the generosity of Adrian Travis and Allan Sullivan.

Contents

Chapter 1

Introduction

There is something that is naturally appealing about the prospect of three-dimensional displays becoming commonplace in our society. Much of this has to do with expectations built up by the media, and unreasonable demands made on our technology driven society. Researchers have been echoing that we will be capable of producing reasonable 3-D displays within the next few years for several decades. Results ensue, first impressions are made, but imaginations aren't tamed. The public is left largely unaware of what is within the reaches of our current technology.

There is, within the research community, a strong understanding of what is capable in this field. An esoteric few remain intrigued by the burgeoning electronic display industry, awaiting developments that may revive the potentials of 3-D displays. Only a handful remain though, and many more will need to be enlisted before 3-D displays become easily accessible to everyone interested in them.

Displays have become our primary interface for digital technology. Visual feedback has proven to be a necessary component of our interaction with the devices we use for complex computational tasks. Researchers working on virtual reality systems and 3D computer graphics visualization have realized that our current visual interface has not reached its full potential. There are others still who are interested in pioneering the new field of 3D displays. They can envision an interface designed to provide an experience which is as intuitive as navigation in the natural world around us.

Motivation for developing high quality 3D displays has faced much criticism, especially within the display community. To the 2D display engineer, it is difficult enough to develop products which meet the satisfactions of a highly critical customer base. Plus 3D costs more to develop, needs a new framework to reach mass consumer level, and doesn't really look that much different. Despite the valid criticisms that companies and researchers of 2D displays may have, there have been consortia developing to create standards for 3D technology which are supported largely buy companies that manufacture 2D displays [21]. Needless to say there is an interest in 3D, and that it may someday be realized as a strong market. Sharp forecasts that the Japanese LCD demand for cellular phones and PC LCD monitors will reach 70 million units in 2005 and 30 to 40 million units of them will be employing 3D LCDs.

This thesis describes the process of designing a 3D display, and is the culmination of my research at the MIT Media Lab Spatial Imaging Group as well as several years of experience within the community of individuals who are dedicated to the art and science of 3D imaging. I have included all the aspects of my research which I believe pertain to the design of 3D displays. The commercialization of 3D displays is particularly fascinating to me, so I have measured the success of displays versus their marketability.

There are and will remain to be a multitude of techniques and applications that 3D displays have found. This thesis focuses on the view-sequential 3D technique. I remain partial to this technique, but I do not consider it to provide the sole answer to the 3D problem. If 3D displays do finally achieve mainstream success, it will be due to the contributions from the entire set of artists and engineers working with them. If one of the many 3D display companies is able to strike it rich, they will owe a large part of their success to the advertising and outreach efforts which are the collaborative effort of all 3D companies and researchers.

1.1 The 3D Problem

1.1.1 Quantity of Information

3D displays almost exclusively require orders of magnitude more information than their 2D counterparts. This complicates three important aspects of the display industry: generation, transmission, and display of visual information. Despite the growing potential for 3D displays and their applications, there will need to be a justification for the extra expense that is inherent in 3D displays. As an example, consider the number of samples N in a full parallax computer generated hologram given by:

$$N = \frac{2\Delta h \Delta w \Delta\theta}{\lambda^2} \tag{1.1}$$

For a width Δw = 0.5 meters, height Δh = 0.5 meters, wavelength $\lambda =$ 500nm, and view angle $\Delta\theta = 30^o$ the number of samples required is about N = 10^{12}. Typically a computer monitor will have around 10^6 samples per frame, and this makes clear the complications involved in developing a full-parallax holographic television.

1.2 Depth Perception

What exactly causes the perception of 3D images is not entirely clear, though it is clear that there are several visual cues which assist the experience of a 3D object. These clues disambiguate the different categories of objects that exist in the familiar three-dimensional world. 3D displays typically attempt to manipulate these cues so that flat objects appear to take the shape of an arbitrary 3D form. This is a particularly interesting task because the visual perceptive processes dictates that 2D information be extracted from objects before any cognitive processing is issued by the brain. Thus there is room to engineer the mapping of visual information that gets sent to the eye, and the type of mapping used provides a distinction between many 3D display techniques.

A brief set of definitions is necessary to include in a discussion on the perception of 3D images:

Occlusion When an opaque object obstructs the view of another object, it is said to occlude that object.

Accommodation Using the ciliary muscle to cause a deformation in the lens of the eye. This allows an adjustment in focal length, which provides a mechanism for determining the depth of an object.

Kineopsis (Motion Parallax) The sampling of several sets of 2D perspective views from different vantage points.

Convergence When the muscles in each eye are cued to turn the eyes so that their optical axis converge to a point in space.

Binocular Disparity Each eye intercepts a different 2D perspective based on its unique location in space.

Typically electronic 3D displays have only been able to provide viewers with Binocular Disparity. The so called "Virtual Reality" displays are based on the same principle as early stereographic photography: each eye receives a unique view. These displays often cause viewers to suffer eye strain, presumably because of the difference between accommodation (on the display screen) and convergence (at point in front of the screen) cues. The displays also cause severe distortions if not viewed from the intended position.

Stereographic displays 3D displays are extremely appealing because it is easy to capture footage for them, they are inexpensive to produce, they have manageable bandwidths, and are sufficient for applications where viewers don't change position often. Stereographics Corporation offers several stereographic display solutions, including the new Mirage$^{\text{TM}}$ projectors based on DLP technology. Sony's Imax cinemas exist in most major cities in the U.S.

For applications where dramatic effect is not nearly as important as accuracy, stereographic displays are not a sufficient solution. Realistic 3D imagery can only be conveyed if sufficient motion parallax is present. Until recently, the bandwidth requirement for 3D displays was too great to include Kineopsis. Recently, several commercial displays have been introduced which include this feature, and they are described in the next chapter. The outlook of 3D displays has been greatly improved by the introduction of new pixel generating devices with greater bandwidths.

1.3 Converting to 3D

At the heart of any 3D display which provides Kineopsis is the *angular-multiplexing* scheme whereby an extra degree of freedom is encoded within pixel information. The *stereographic lenticular, parallax barrier, multi-view,* and *holographic* displays already mentioned all employ this technique. Volumetric displays encode pixels with an extra degree of freedom in the form of

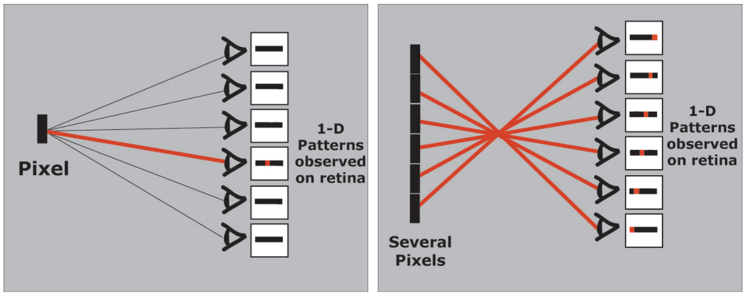


Figure 1.1: *angular-multiplexing of a single pixel allows an observer to intercept a unique ray at different locations (left). Several pixels that have been angularly multiplexed can be utilized to give the impression that rays are converging to a point in front of the array of pixels (right).*

an additional Z-coordinate. Though this is the simplest imaginable mapping, nature does not use this technique to encode depth information. The most realistic looking 3D displays will not utilize this technique either, though there remains much practical importance to volumetric displays at this stage in the evolution of 3D display technology.

angular-multiplexing is the method by which a single pixel appears a different color depending on the relative position of an observer. This method most closely resembles the natural world, where any isolated point in space can have rays emanating from it in any direction, and will appear different depending on which of these rays an observer intercepts. Multi-view and holovideo displays both directly execute an *angular-multiplexing* regime, whereas the other techniques encode information differently and then map from one regime to another.

In Fourier optics, there is a relationship between the direction of light propagation, and the spatial frequency pattern of its amplitude on an intersecting surface. This relationship introduces the analogy between *angular-multiplexing* and *frequency multiplexing*, and a mathematical relationship between the analysis of holographic displays and other 3D displays.

1.3.1 Spatial-Multiplexing

The lenticular, parallax barrier, and color selective filter techniques all employ *spatial-multiplexing*. These techniques sacrifice the pixel information

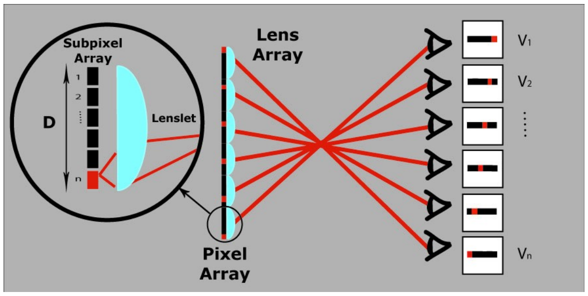


Figure 1.2: *Lenticular displays use the focusing action of lenses to encode directional information in multiple sub-arrays of pixels. The result is that the effective pixel size is reduced to the size D, and the number of views is determined by the number N of pixels within a sub-array.*

that is generated by a modulator/emitter in order to achieve direction dependent viewing conditions. An example of this is demonstrated in Figure 1.1. This technique has the appeal of minimal size, because a flat SLM and an extra layer to convert from *spatial* to *angular* multiplexing are often all that is needed. There is a trade-off between the number of views and the effective pixel size when *spatial-multiplexing* is used. The technique relies on both small pixel sizes and large pixel counts, a combination which is not generally justifiable in the eyes of commercial SLM manufacturers.

1.3.2 Temporal-Multiplexing

temporal-multiplexing is used in devices which can generate a sufficient amount of pixels within the temporal resolution of the eye. This technique is heavily dependent on the bandwidth of the SLM that is used. If several multiples of the desired Space-Bandwidth Product (SBP) can be generated within about $1/60^{th}$sec, then it is possible to generate multiple views with a single SLM. The Cambridge View-Sequential display relies on this technique to encode depth information with a fast CRT. As with *spatial-multiplexing*, it is necessary to dedicate a device to the task of converting multiplexing regimes. However, it is not necessary to sacrifice the SBP in order to for this to be accomplished.

Time multiplexing techniques offer great promise for academic researchers because they can provide impressive immediate results as well as opportunity for unlimited future developments. SLMs have become much faster over the years. Current SLM technology provides frame rates at remarkable speeds, fast enough to be utilized in commercially available volumetric displays, and optical switching technology.

The elegance of using a single modulator is the greatest appeal of time-multiplexing, one which could increase the simplicity and reduce the size and cost of a display. A view-sequential display could potentially be designed with a microdisplay and a Wedge$^{\mathrm{TM}}$ lightvalve to create a compact flat panel 3D display [5]. A view-sequential approach could be adapted to create a projection engine for a multi-view 3D display with wide fields of view [3]. A full parallax view-sequential display has not yet been built, and could be implemented quite nicely in coming years. Commercial graphics cards may soon be able to render enough content quickly enough to allow real-time interaction with computer graphics imagery. The possibilities for new research seem to be extremely fruitful with this technique, and the hopes of commercialization realizable.

1.4 The Cambridge/MIT Display

Electronic 2D displays have only developed commercial potential in the last few decades, and the information revolution has allowed them to develop at an alarming rate since they were first introduced. The microdisplay market is a burgeoning new industry based on consumer interest in data projection devices. Displays have become portable, ubiquitous, and to a certain extent, even dispensable. The expense of displays has been a concern of researchers, and most effort has been directed toward cheap and effective solutions. Advances in SLM technology have been utilized in recent 3D display systems, which are described in the next chapter.

Cambridge University has been developing view-sequential 3D displays for the past two decades, and achieved great success. Their displays use Cathode Ray Tubes (CRTs) to generate pixel information. Interest in this research was sparked when my colleague Christian Moller began his studies at MIT as a transfer student from Cambridge University. Suggestions by Christian's advisor Dr. Travis eluded that the view-sequential technology might be appropriately revisited in light of recent advances in display technology. Discussion quickly developed on how improvements could be made to previous view-sequential systems, and plans to build a new system were made.

It became immediately apparent that a very impressive 3D display could be built which utilizes new SLM technology instead of CRTs. The following improvements were considered:

- Increased Pixel Bandwidth
- Decreases Physical Size
- Flexible Optical Design
- Increased Brightness

The most important reason for considering the use of SLM technology is that it continues to develop very rapidly. It is of critical importance to 3D displays that pixel bandwidth continue to increase, for this bandwidth is necessary for continuing to increase the amount of parallax information. Current 3D displays aren't capable of providing nearly enough parallax information, so improvements in this category alone would demonstrate a successful improvement of the technology.

Chapter 2

Background on 3D Displays

2.1 Applications

It is hard to put a price on the effectiveness of 3D displays because the empirical factors in 3D perception are not yet entirely understood. Instead there ends up being a risk/reward trade-off that companies consider when deciding to make the 3D plunge. There have been several industries that have been targeted for 3D display use, but none of them have really taken a big bite yet. Competition between companies is fierce. Because 3D technology is often misunderstood, marketing tends to play a critical role in the success of 3D display companies. The most critical decision a company must ask themselves is: ”Who has the incentive to meet the increased cost of using 3D displays over 2D displays.” If this question is not answered, the company's target audience is likely have unrealistic or uninformed expectations of what the company can provide. 3D displays cannot be sold cheap yet, and it is unrealistic to expect this until there is a greater awareness of their existence and the advantage they can provide.

The large variety of 3D solutions that have been developed may hinder a loyal customer base, simply because it is difficult to differentiate the technologies from each other. There is also speculation that the variety of technologies is favorable because it ensures a diverse market. At any rate, it makes clear the fact that there isn't any one way to do 3D, and that there are improvements which we can look forward to in the coming years. This is the greatest potential that exists for a 3D display industry, that it can continue to grow... if it is fed.

The most likely candidates for industries that will benefit from the use of 3d displays are:

- Medical Imaging
- Simulations and Heads-Up Displays
- Design, Modeling, and Scientific Visualization
- Gaming and Entertainment

Visualization plays an important role in the diagnosis of medical conditions, and volumetric data sets are commonly used to interpret conditions about patients' health. Simulations may be the most profitable application outfitted for application of 3D displays. Military simulations can be an extremely lucrative business, as companies such as Ethers and Sutherland have demonstrated [7]. Immersive displays which offer extremely lifelike experiences are important for combat, emergency, and (the big one right now) security preparedness. Companies such as Honda Motors and Ford Motor Company have supported development of holographic stereogram printers in hopes that this technology will foster more efficient industrial design practices. Standards such as Virtual Reality Markup Language (VRML), and the new X3D offer ease of integration with stereographic viewing techniques, and popularity of 3D display devices for gaming offered by Stereographics is growing. The most successful 3D display application in the entertainment industry today is the 3D Imax theaters, which exist in several major cities worldwide.

2.2 Generating Pixels

The most important aspect of any 3D display is how it generates picture information. There are two ways to do this: either by eliciting a matrix of spherically emitting points upon a surface, or by altering the amplitude characteristics of a uniform beam of light. The former technique falls into the category of emitter displays, while the latter is known as Spatial Light Modulators (SLMs).

The modulator/emitters that have been considered for use in 3D displays are:

- Emitters

  - CRTs
  - Organic Light Emitting Diodes (OLEDs)
- SLMs
  - Pneumatic Liquid Crystal Displays (LCDs)
  - Acousto-Optic Modulators (AOMs)
  - Surface Acoustic Wave Modulators (SAWMods)
  - Optically Addressable Spatial Light Modulators (OASLMs)
  - Ferro-electric Liquid Crystal Displays (FLCDs)
  - Digital Mirror Devices (DMDs)
  - Grating Light Valves (GLVs)

Recent work by Hong[12] has introduced the possibility of developing a new type of acousto-optic modulator based on surface acoustic wave (SAW) transmission in crystals. SAWMods can be produced very inexpensively, but SAWMods, OASLMs, and GLVs are not commercially available yet. Commercially available CRTs and OLEDs which are capable of running at faster than video rates are not available either. These devices are capable of running at much faster rates, but development of custom driver electronics is necessary in order to achieve the desired effect. In the case of SAWMods, GLVs, OASLMs, and OLEDs, it is necessary to fabricate a device according to the desired specifications.

Since a Modulator/Emitter must be used to generate pixel information in 3D displays, it is worthwhile to consider a comparison of characteristics. The most important feature is bandwidth, which has already been mentioned. There are also several other important factors, including contrast ratio and color depth. All of these factors, which are outlined in Table 2.1, will contribute to the commercial success of a 3D display.

The limiting factor in commercialization may be the cost of a display, which should cause 3D display designers to choose their pixel generating device carefully. All the devices mentioned have cost/pixel bandwidth ratios in the range of MHz/$. It is apparent that CRTs give the greatest cost/bandwidth ratio, which makes them very attractive candidates for 3D displays. However, DMDs and FLCDs are next in line, and there greater bandwidth makes them extremely attractive candidates for 3D displays.

Other important factors in modulator/emitter characteristics include SBP (The number of pixels) and binary/analog mode of operation. Commercially available CRTs and LCDs have large enough SBPs to be able to sacrifice spatial resolution for encoding 3D information. DMDs and FLCDs are binary devices, which allow both the color depth and spatial resolution to be optimized to encode the desirable quantity of 3D information. These devices make great sacrifices in bandwidth to achieve large color depths, but it may prove to be more effective to compromise this sacrifice between color depth and 3D information. The regime which Actuality system uses is quite intelligent: all the bandwidth of their DMDs is used to encode 3D information and the spatial resolution is sacrificed if greater color depths are desired.

Company	**SLM**	**SBP**	**Rate**	**Colors**	**BW**	**Cost**
Viewsonic	CRT	3.1Mpix	68Hz	24bit	5GHz	$500
Cambridge	CRT	153kpix	1020Hz	8bit	1.2GHz	-
Kodak	OLED	75kpix	60Hz	24bit	108MHz	$10k
Brimrose	AOM	32x2kpix	50MHz	+ 8bit	+12.8GHz	$35k
IBM	LCD	9.2Mpix	56Hz	24bit	16.5GHz	$8.5k
CRL Opto	FLCoS	1.2Mpix	2.5kHz	1bit	3GHz	$6.2k
TI	DMD	768kpix	10kHz	1bit	7.68GHz	$10k
MIT	GLV	256pix	50MHz	8bit	10.2GHz	-
MIT	SAW	+10kpix	+500MHz	+8bit	+4GHz	-

Table 2.1: *Comparison of SLM technology. SBP = Space-Bandwidth Product , Rate = Refresh Rate, BW = effective BandWidth in bits/sec, CR = Contrast Ratio.*

2.3 Holographic Stereograms

It is worth mentioning holographic stereograms because they perform the same function as electronic 3D displays, and use the same optical method to encode depth information. The difference is that the method of storing image information is permanent with holographic stereograms. Once an image is recorded as a print, it cannot be altered. If alternate images are desired, more prints must be produced. Analysis of holographic stereograms can be used to characterize most 3D displays which utilize kineopsis.

The method of angularly multiplexing 2D intensity information can be achieved without sacrificing resolution by recording multiple 2D images onto a single hologram. Indeed, both the minimum resolution and parallax sampling requirements can be met so as to produce very convincing 3D scenes. Much of the work at the Spatial Imaging Group has been dedicated to refining these techniques. There is a diffraction limit on the amount of sampling in the parallax field, and consequently images of less depth than analog holograms must be used. In this case, the coherent imaging properties of a hologram are not preserved. It is the focusing properties and large storage capacity of holograms that are utilized.

There have been several companies which have attempted to provide commercial solutions for printing holographic stereograms by developing a custom hardcopy printer. Zebra Imaging offers a printing service for full parallax holographic stereograms that can be tiled to an extent of several meters [13], and Liti3D offers services for printing horizontal parallax stereograms in the same price rang as custom portraits [16]. Sony also put some work into a photo-booth style printer/capture system which offered a complete stand alone solution and produces wallet size images [1].

2.4 Holovideo: The Ideal 3D Display

Holovideo systems provide the ultimate solution for 3D display systems. However, it is at the expense of enormous data rates. The most significant hindrance that holovideo currently faces is the lack of display devices with sufficiently small pixel sizes. The angle of diffraction θ of a ray of light is given by the equation:

$$\lambda f = \sin\theta \tag{2.1}$$

This dictates that with a wavelength $\lambda = 500$nm, a diffraction angle of 30^o requires a spatial frequency of f = 1,000 lp/mm, or a pixel pitch of 500nm. Current microdisplay devices are capable of achieving pixel sizes in the 10μ range, but it isn't promising that they will go much smaller.

Holovideo systems can be thought of as having programmable optical properties, and are flexible enough to implement any of the techniques for 3D image creation. Holographic computation may ultimately provide the most efficient means of displaying 3D images, because it can allow a hybrid approach of of optical encoding methods, and can be used in an application specific manner. Recently, dramatically deep scenes have been demonstrated by Halle and Plesniak via the Real Image Plane (RIP) stereogram encoding method, which continues to be developed. MIT Spatial Imaging[11] [20] and Qinetiq[6] have both developed holovideo systems which approach commercial quality.

The largest drawback to holovideo systems is the inclusion of an extra processing step. Holovideo systems are capable of producing piecewise representations of 3D objects by displaying geometric primitives (namely 3D vertices), as well as the stereographic technique of angular-multiplexing of 2D images. Unfortunately, in either case, a fringe pattern must be generated which in some cases can cause unwanted redundancies. Thus holovideo offers greater flexibility in optical encoding methods at the expense of reduced flexibility with bandwidth limitations.

2.5 Survey of 3D Displays with Kineopsis

Many 3D displays have already been developed, some to the point of commercialization. This chapter provides a comprehensive list of all the displays that have been built in recent years, and begins a discussion and categorization of the different techniques that have been developed so far.

2.5.1 Volumetric Displays

Volumetric displays fill an entire volume with spherically emitting points (voxels). The resulting image lacks any occlusion information because none of the voxels are capable of blocking information from any other. It is possible to include an occlusion mask within a volumetric display to limit the viewing angle of each voxel, but this has not yet been demonstrated. Volumetric displays have the advantage of providing a large viewing angle about which to view 3D images. This has proved to be particularly useful for medical imaging applications, where it is desirable to see through different layers of data (e.g of tissue) and multiple viewers are required. Typically, a *time-multiplexing* technique is employed whereby several surfaces are addressed sequentially in time and each updated within the temporal resolution of the eye.

Actuality

Actuality Systems has designed the Perspecta® series volumetric displays which are based on TI's DLP projection technology [8]. Their display is capable of generating up to 100 million voxels with 3-bit color. They also implement spatial dithering to achieve more color values at a reduced voxel count. Their system consists of a custom built rasterization engine which receives 3D graphics primitives, a DLP projection engine with custom optics, and a spinning 50/50 diffuse/transmissive disc which traces out a sphere at 24Hz. The display can be viewed from 360^o horizontal and 270^o vertical, but the edge of the projection disc is always viewable. Because of the considerable work that the company has put into graphics acceleration hardware, the display can easily be interfaced with standard OpenGL applications by intercepting 3D primitives from conventional desktop graphics cards via a PCI bus.

Vizta3D

Vizta3D has developed the DepthCube$^{\text{TM}}$display based on twisted-smectic crystals and TI's DLP projection technology. Vizta3D developed proprietary firmware with the help of TI that allowed their projection engine to run at 15-bit grayscale at 1,200 frames/sec. They stack 20 individually controlled liquid crystal layers on top of each other, and synchronize them with their projector. When a single liquid crystal layer is activated, it becomes diffuse, creating a plane of voxels. Within a period of $1/60^{th}$ of a second, all 20 planes have been addressed and are cycled through again to avoid flicker. The result is a cube with 1024x768x20 = $15.3\text{x}10^6$ voxels. To avoid depth discontinuities between the 20 planes, Vizta3D implements what they call multiplanar anti-aliasing. The process involves addressing points which occur between two planes as a fractional amount to both planes. The result is a perceived $465.7\text{x}10^6$ voxels.

Vizta3D has managed to compress their system to the size of a conventional Cathode Ray Tube monitor. There system does not have the same viewing extent as Actuality (90^o field of view horizontal and vertical) and its depth range is limited to about four inches. Their display does offer interactivity without any additional graphics hardware because of the similarity between the coordinate system which the 3D images occupy, and the frame buffers of conventional graphics cards. The result is a highly flexible system which is easy to integrate with conventional 2D graphics solutions.

2.5.2 Lenticular, Color Selective, and Parallax Barrier Displays

Though they are technically multi-view displays, these techniques deserve their own category because they have earned a competitive place in the commercial market. These displays are capable of displaying all the same depth cues as multi-view displays. However, the number of views they are capable of displaying is usually limited because they employ *spatial-multiplexing*, whereby the resolution of the display is sacrificed to include parallax information. The minimum pixel size is consequently the limiting factor in these displays.

Flat panel 3D displays have been developed based on these techniques because they offer the most compact optical arrangement available for 3D display technology. Though pixels sizes on flat panel screens cannot be reduced below several hundred μm, larger displays with generous pixel counts have recently been introduced. These displays can be viewed from much further back than the typical viewing distance, suggesting the possibility of greater parallax sampling without loss in image quality.

These techniques have the disadvantage of repeating viewzones, and reduced horizontal resolution. Consequently, less horizontal detail can be displayed, and images appear to jump when the viewer moves out of the correct viewzone. Sometimes this can be considered an advantage because it allows viewers who are positioned off-axis to see an image on the display, which often draws them in and encourages them to observe the display correctly. However, it this is a misleading feature because it gives the impression that more information is being conveyed by the display than it actually is.

Stereographics Corporation

Recently, Stereographics Corporation has released their Synthagram®flat panel monitor series which is a lenticular-based 3D display[19]. The series ranges from XGA (1024x768 pixel) to UXGA(3840x2400 pixel monitors), and a custom fabricated diagonal lenticular screen, which divide pixels into 9 different views. The monitor is fed from the DVI output of a graphics card. The lenticular screen is designed to eliminate moire fringing, which can occur in lenticular flat panel screens, and divides pixels on the RBG level. Stereographics also offers an SDK, plugins for 3D Studio Max and Maya 3D modeling applications, and a windows viewer to playback content in the appropriately encoded format.

4DVision

4DVision has built an 8-view display system that is based on a color selective filter which limits the angular extent each pixel based on its color. The filter is applied to the individual RGB cells of a flat panel TFT or Plasma screen. Their filter offers less reflections off the surface than lenticular arrays, reducing glare which distracts from image quality. The system also offers the potential to switch between 2D and 3D applications. 4D Vision has built displays as large as 50" diagonal (the 4D-50$^{\mathrm{TM}}$display) and offer OpenGL drivers for easy integration with 3D graphics applications.

NYU

The Media Research Laboratory has built a parallax barrier display based on a DLP projector, a fast switching PI cell, a fast-switching ferroelectric shutter, and eye tracking technology[15][14]. The display provides stereoscopic viewing which is updated in real-time according to viewer position (including depth) and head rotation. The system encodes view information via the PI cell, which consists of several columns of shutter elements that can be addressed simultaneously and refreshed at 180Hz. The PI cell is placed in front of a projection screen and each shutter element provides the limiting aperture for a group of pixels. The shutter arrangement is computed and updated every $1/180^{th}$ second to allow greater flexibility in matching viewer position with the correct perspective information. The projector color wheel is removed and image information is encoded as color information before it is sent to the projector. The result is a time-sequential series synchronized image/PI cell bitmap pair that is updated at 60Hz.

2.5.3 Multi-View Displays

Multi-view displays offer the strongest similarity to holographic displays, and come with the added requirement of handling more information than stereographic, lenticular, or parallax barrier displays. Multi-view displays with enough views can provide sufficient image information to relax the competition between convergence and accommodation that exists in other displays. It is much more challenging to include the feature of interactivity with multi-view displays because the amount of imagery which must be generated in real-time is much greater. Multi-view displays also commonly implement multiple 2D display devices, which increases the cost of manufacturing. For these reasons, multi-view displays have yet to achieve commercial success, and have been restricted primarily to the research domain. However, they require less bandwidth than holographic displays, so they may be the most desirable temporary solution to comfortable 3D viewing.

Tokyo University

Tokyo University of Agriculture and Technology Department of Electrical and Electronic Engineering has produced the most impressive multi-view display, which consists of 64 separate QVGA (320x240 pixels) LCDs running synchronously, each viewable from a different direction. In most multiple projection systems, the boundary between views is set by the distance between stops of adjacent projection lenses. This generally makes it difficult to avoid discontinuities between views, which is distracting to viewers [3]. The team at Tokyo University has avoiding this problem by placing their LCDs at different heights and designing a honeycomb projection lens array. This allows the horizontal position of each projector to overlap so that the exit pupils can abut one another. A one-dimensional diffuser increases the vertical extent of each exit pupil so that there is a comfortable variation in height that all views can be observed from. This system provides a very satisfying amount of parallax information, but the cost of manufacturing 64 separate displays for each system makes commercial viability seem unlikely.

Cambridge University

The Cambridge University view-sequential displays were developed by implementing a *time-multiplexing* principle with a fast optical modulator, active shutters, and projection optics. Time sequential information is angularly multiplexed by restricting the pupil of a projection system in conjunction with images being displayed on the modulator. In the original systems, custom designed ferroelectric shutters and Cathode Ray Tubes (CRT) were developed. Full color images with as many as 26 views updated at 50Hz [2], and images as great as 50” [18] were demonstrated. View-sequential systems have the advantage of using a single modulator, which is much more economically feasible than other multi-view systems. They also provide viewing zones which naturally abut one another, which is a subject of difficulty for other multi-view displays [3]. However, they require large projection optics, which limits the field of view. They also throw away a lot of light because of the active shutter, which increases as more views are included in the system.

Chapter 3

Design Proposal and Implemetation

3.1 Synthesis of Current Work in Field

There is no doubt that a 3D display market is developing, but the products that are currently available are severely restricted by the cost of manufacture. Typically, the time that is put into R&D causes the price of first generation systems to be far out of the reach of mere enthusiasts of 3D. The line is quickly drawn between those who really need to have effective and realistic visualizations, and those who can do their work just fine with a desktop monitor. Table 3.1 shows some of the characteristics of some commercially available 3D display products. These displays consume little more power than there 2D counterparts, and maintain similar contrast ratios. All products have solutions which allow easy integration with legacy 3D graphics applications, as well as software development kits.The prices however, range from 2-20 times the price of 2D displays of similar dimensions.

Table 3.2 shows a comparison of important features for several 3D display. It is apparent from this list that a large variety of features are available, and any new technology will have to demonstrate a strong competitive advantage in order to make find its place as a commercial product.

The idea behind eye-tracking displays is extremely intuitive, and the MIT and NYU systems show remarkable promise for this technique. However, these systems require the viewer position to be precisely determined. Though very small latency between viewer detection and image refresh has

Display	**Type**	**Power**	**Size**	**Depth**	**CR**	**Cost**
Mirage 6000	St	1.7kW	Theater	limited	1000:1	\$80k
Synthagram^{TM}222	Len	300W	55cm dia.	10cm	400:1	\$15k
4D-50^{TM}	CSF	300W	125cm dia.	10cm	400:1	\$10k
Perspecta®	Vol	350W	60cm dia.	44.5cm	-	\$50k
DepthCube^{TM}	Vol	-	47.5cm dia.	10cm	-	\$50k

Table 3.1: *Comparison of commercial 3D displays. St = Stereographic, Len = Lenticular, C.S.F. = Color Selective Filter, Vol = Volumetric.*

Display	**Type**	**Size**	**Depth**	**Views**	**FOV**
IMAX	St.	Theater	limited	2	limited
Mirage 6000	St.	Theater	limited	2	limited
NYU	P.B., E.T.	60cm dia.	limited	2	90^o x 90^o
MIT	E.T.	60cm dia.	limited	2	limited
Synthagram^{TM}222	Len.	55cm dia.	2”	9	90^o
4D-50^{TM}	C.S.F.	125cm dia.	2”	8	90^o
Sharp 2D/3D	P.B.	37.5cm dia.	limited	2	limited
Perspecta®	Vol.	52.5cm dia	24”	10^8	360^o x 270^o
DepthCube^{TM}	Vol.	47.5cm dia.	4.1”	10^7	90^o x 90^o
Cambridge	M.V.	125cm dia.	5cm	32	30^o
Tokyo Univ.	M.V	15 x 15cm	5cm	64	NA
MIT	H.V.	15 x 7.5cm	50mm	∞	30^o
Qinetiq	H.V.	.5 x .5m	∞	∞	3^o

Table 3.2: *Comparison of 3D Displays. Len = Lenticular, C.S.F. = Color Selective Filter, P.B. = Parallax Barrier, E.T. = Eye Tracking, Vol = Volumetric, St = Stereographic, M.V. = Multi-view, H.V. = Holovideo.*

been demonstrated, it is the author's opinion that suitably precise viewer detection devices are currently unsuitable for this technique to compete with other 3D displays.

Volumetric displays are full parallax, have wide viewing angles, and are fairly easy to make interactive. The mapping of depth information that is made by computer graphics hardware is very similar to volumetric displays, and allow the computational tasks that displays must perform to be reduced. These displays do not produce the most realistic 3D imagery though, and the task of including occlusion into volumetric displays seems to greatly reduce the simplicity of this technique.

SLM technology is not yet capable of producing sufficiently small pixel sizes to make holovideo systems commercially viable anytime soon. Small pixel size, scanning optics, large output lenses, and supplemental processing all weigh against holovideo systems.

As resolution increases, displays that implement spatial-multiplexing will improve. However, it is presumable that small enough pixel sizes will eventually be achieved. If this is the case, diffraction effects will dominate. Inevitably, spatial-multiplexing techniques will be absorbed by diffraction-based imaging techniques simply because they depend on both the non-diffractive properties and minimum feature size of SLMs.

3.2 Deciding which Technology to Use

It became apparent that developing custom technology, was not conducive to the task of developing a 3D display because of time constraints. Originally, it was considered that OLEDs could provide the most favorable results, but would have to be custom fabricated in order to meet our desired specifications. Ultimately we decided on well-developed commercial technology which would most reliably provide the features that we were looking for.

3.2.1 DMDs

DMDs are receiving much success in the film industry with the success of their DLP projection engines. Film distribution is undergoing a great change, and for the first time in history, films are being released in digital form instead of on cinematographic film. If DLP and other digital projection technology can hold its weight, then the film industry may come to rely on these SLM devices to reduce printing and transportation costs. The greatest advantage that this provides 3D display designers with is a reliable SLM which is well tested as a durability and long lasting solution.

Because DMDs are binary devices, they also have great appeal to the optical communication industry. Though there is not as much immediate financial gain here for DMDs, TI has demonstrated a strong interest in providing solutions which are flexible enough to be developed for optical switching and computing solutions, which both require greater flexibility over DMD operation. The consequence is that TI has developed a reliable architecture for programming the speed of DMDs, which is also a great advantage to 3D display developers.

3.2.2 FLCDs

FLCDs are capable of binary switching times as fast as 50μsec, which makes them a great contender as pixel generating devices for view-sequential displays. However, microdisplays based on FLCD technology have only demonstrated frame rates in the 1-3kHz range, placing them a notch below DMDs in bandwidth. Transmissive FLCDs are ideal shuttering devices though, and are used extensively in optical switching systems.

3.3 Acquiring the Devices

We were loaned an 800x600 pixel DMD and driver board that was bought from Rochester Micro Inc. The devices came with custom software which allowed programmable control over loading sequences of images onto the DMD and controlling the speed that the images were displayed at. The driver board had 512Mbytes of on board RAM and allowed bitmaps to be loaded to memory from a desktop via ethernet. Rochester Micro claimed that the device was capable of displaying bitmaps at 4kfps, however, we were not able to run the device faster than 500fps with reliable results. After much phone consultation and repairs from Rochester Micro, we severely damaged the driver board while trying to debug the problem. We decided to find another source for DMDs rather than wasting any further funds on repairing this unreliable device.

These experiments were extremely daunting because the availability of DMDs with programmable boards is not great. We began a discussion with Productivity Systems, which offers exactly that solution, but at a large cost. The discussions with Productivity made me aware of two of their customers: Actuality Systems and Vizta3D. We contacted these companies to implore about any spare devices that they might be able to offer, and to our good fortune received a positive response from Vizta3D. Vizta3D was in the process of cleaning their storage facilities, and in need of removing some of their earlier prototype systems which they no longer needed. They gave us an entire 800x600 pixel DMD projection system, complete with drivers, a PCI data transfer card, and proprietary software to drive the display. The projection engine consists of:

- Mercury arc lamp.
- RGB separator and combiner optics.
- 3 800x600 pixel DMDs, one for each color.
- Projection optics.
- Driver electronics which display 5-bit RGB images at 800fps on each DMD.
- Framebuffer to store images with high speed connection to the driver board.
- PCI data transfer card to load bitmaps from a computer to the display.
- 80 volt, 5 amp power supply for the DMDs and projection lamp.
- 3' x 3 x3' storage and transportation container for the projection engine.

We also received a custom designed FLCD shutter and driver from Cambridge University. The shutter is 100x100mm and consists of sixteen columns which are individually addressable. The device is specially designed for use in view-sequential displays, where each column limits the viewing angle of a single view. Each column can be refreshed as quickly as 1-2Khz. As it is used in this display, each column of the FLCD is switched open for 1/800sec and then waits 1/50sec to switch open again.

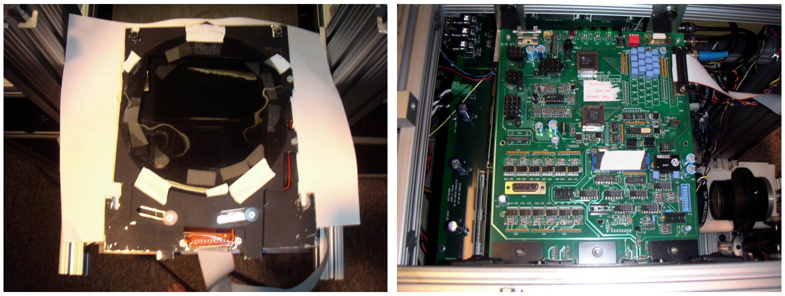


Figure 3.1: *The FLCD is shown (left) as well as its driver electronics (right).*

3.4 System Design

The system that was decided upon uses a DMD projection engine to generate pixel information, and an FLCD to encode depth information. The system is identical to the earlier view-sequential systems built by Cambridge University with the CRT replaced by DMDs. The proposed system offers the potential of increased brightness, greater pixel bandwidth, and decreased display size. These improvements were considered as extremely beneficial to the success of this technology, so the construction of a prototype system was built to analyze image quality, measure display characteristics, and open discussion for future developments.

3.4.1 Optical Design

The optical design of the proposed system is perhaps the simplest aspect of the display. There are only five components: DMD, diffuser, FLCD shutter, projection lens, field lens. The optical design that we implemented is identical to the design of the 10” and 25” cambrige view-sequential displays. This design was chosen for simplicity. The use of a convex mirror to replace the output lens was considered, but the expense of this item prevented us from including it in the first generation system.

GIRS

To begin a discussion of the optical design, consider the General Imaging Relay System described by Benton[4] , which provides a framework for analyzing the optical characteristics of a 3D imaging pipeline.

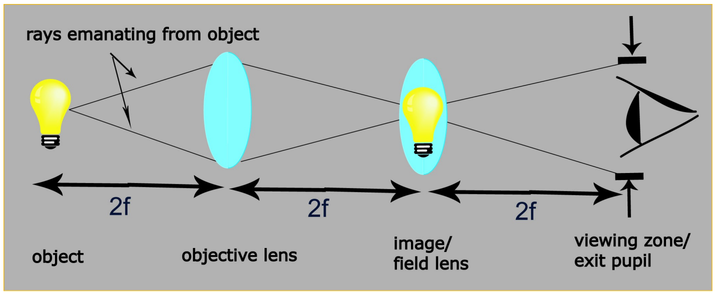


Figure 3.2: *The GIRS can be used to analyze information content in an arbitrary 3D display. An objective lens generates an image that comes to focus on a field lens. The objective lens is the entrance pupil for the system, and the image of it that is created by the field lens limits the field of view.*

The GIRS consists of an object, an objective lens, and a field lens. The lenses have the same focal length, and each element is separated by twice that distance. When a viewer looks through the field lens, they see a non-distorted image of the object centered at the field lens. The image disappears when the viewer moves too far off axis. The extent of the viewing zone is determined by the size of the objective lens. The objective lens forms the limiting aperture of the optical system, or the *entrance pupil*. The image of the *entrance pupil* created by field lens defines the viewing zone, or the *exit pupil*. Since the GIRS does not magnify the image of the *exit pupil*, the viewzone is the same size as the objective lens.

View-Sequential Display Optics

View-sequential display optics differ from the GIRS system only slightly. In the simplest case, the object is replaced with an SLM, and a shutter is placed in front of the projection lens to further restrict both the *entrance* and *exit pupils*.

This design allows the viewing of the SLM image to be restricted to a limited portion of the viewzone. As soon as the SLM generates a new image, the portion of the viewzone that is viewable is shifted, and this is repeated until the entire viewzone is filled. The shutter and the SLM must be precisely synchronized in order to ensure that each view is correctly displayed,

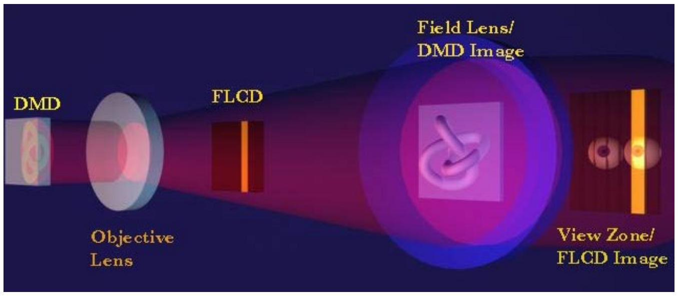


Figure 3.3: *view-sequential displays are a variant on the GIRS system, whereby the entrance pupil is limited by an active shutter.*

and an entire set of views must be refreshed within 1/50sec.

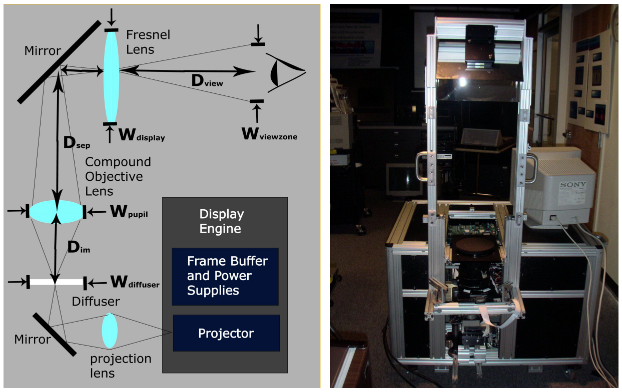


Figure 3.4: *The left image shows a side view of the final optical design for the view-sequential display, listing all the components and the relevant distances. The right image shows a photograph of the optics from the back of the display.*

The final system utilized the long path lengths that are inherent in the display to raise the viewing screen of the display to standard viewing height. The optics are shown in Figure 3.4, and the design is very similar to the standard view-sequential optics, with a few folding mirrors included to optimize image, lens separation, and viewing distances. A diffuser is included to eliminate the visibility of the projector bulb filament since viewers are essentially looking down the optics of the system. A variety of diffusers was chosen, and one was eventually used which diffuses at an angle of about 20^o. Less powerful diffusers did not eliminate the filament image and caused the viewzone to be unevenly illuminated.The relevant quantities are:

- $W_{diffuser} = 5cm$
- $W_{viewzone} = 30cm$
- $W_{display} = 30cm$
- $W_{pupil} = 20cm$
- $D_{diffuser} = 10cm$
- $D_{sep} = 40cm$
- $D_{view} = 125cm$
- $F_{eff,objective} = 75cm$
- $F_{eff,field} = 75cm$

The magnification of the diffuser image is 6 and the distances are given relative to the principal planes of the lens that is creating the image.

3.4.2 Hardware Implementation

The hardware implementation for the Cambridge/MIT display is an original design that was developed by Cambridge university, Vizta3D, Christian Moller, and the author. All of the projector electronics was developed by Vizta3D, and the FLCD dirver electronics was developed by John Moore at Cambridge University. The author and Christian Moller designed and built electronics to interface the FLCD shutter to the DMD^{TM}projector.

There are four electronics components required to drive the proposed display: DMD driver, FLCD driver. FPGA synchronization device, and desktop CPU for image generation and data transfer. The DMD and FLCD driver were provided for us. The DMD driver provides a sync output signal that provides a rising edge every time a new image is loaded on the DMDs. This signal was used to activate the successive LCD panels in Vizta3D's DepthCube^{TM} display. The FLCD driver was designed to receive two input signals: a horizontal sync and a Z-sync. The horizontal sync is used to sequence to successive shutter columns, and the Z-sync is used to reset the sequence back to the first column. An FPGA was built to generate the Z-sync signal from the signal generated by the projection engine. The FPGA feeds both the h-sync and Z-sync signals to the FLCD driver. The FLCD driver requires a 5volt input to control logic as well as +/-40volts to switch the cells between states.

A NIDAQ PCI card was used to transfer data to the frame buffers that feeds the DMDs. The card outputs data on a 32bit wide bus, which runs at 20MHz. Data transfer is designed so that a single pixel is sent at a time over the bus, and half of the lines are consequently wasted. The result is that an entire set of views can be updated in 16 views x 800x600pixel / 20Mpixels/sec = .38sec. In practice, loading times closer to .5sec were demonstrated. There are two frame buffers so that a new set of images can be loaded while still displaying the previous set. The framebuffer was designed to receive an entire set of images at once, and was not capable of loading partial frames. Both of these factors severely limited the maximum achievable interactivity of the display.

3.4.3 Imaging

All of the graphics routines to generate imagery for the display were created by the author. The imaging pipeline was developed in two stages, and computer generated imagery was used exclusively. . The first technique was to use 3D modeling software to generate a sequence of bitmaps that could be loaded onto the display all at once. This was a very tedious process, and involved transferring data from one machine to another. Achieving desired images was an iterative process, and very time consuming. The second technique was to utilize a custom built software application which provided the following:

- Load and manipulate 3D files (e.g. .obj, .3DS, etc.).
- Provide custom rendering based on chosen viewing position.
- Reformat images to load into the frame buffer correctly.
- Call the NIDAQ API to transfer data to the frame buffer.

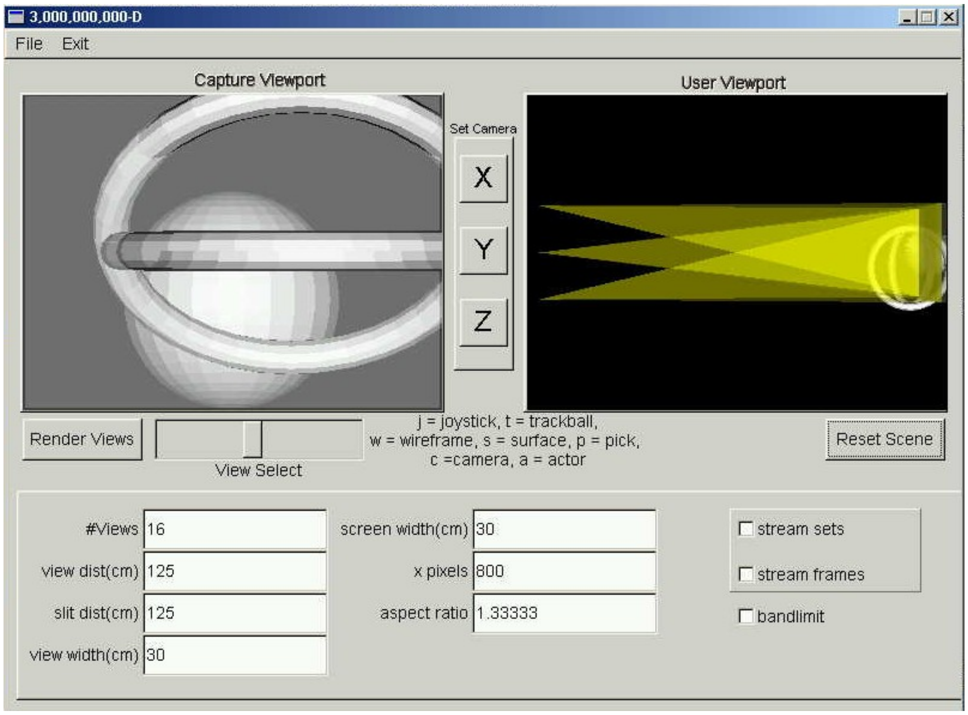
Though it required building new software, the second technique dramatically reduced the iterations between loading images onto the display.

The new software also set a framework to achieve the maximum interactivity that the display is capable of. It was the hope of the author to be able to load a set of images close to 1sec, thus providing reasonable interaction with 3D content. All the rendering was done on a P3 desktop without a fast graphics card, and generating and formatting the data typically took 5-10sec.

Generating the Images

Correct rendering of images requires matching the view distance/display width ratio to the camera distance/object size in a virtual 3D scene. This ratio defines the field of view that the virtual cameras that render the views should have. A virtual camera should render a sequence of images by translating along a track which represents the viewing zone. The ratio of viewing width to distance and virtual camera distance to track width should be equivalent (see Figure 3.6).

As the camera translates across its track in the virtual world, the position of the object also translates in the image plane. It is necessary to implement a shearing transformation to the camera projection in order to stabilize the position of the object in the rendered views. The shearing operation also eliminates the rendering of unwanted information (see Figure 3.7).



j = joystick, t = trackball,
w = wireframe, s = surface, p = pick,
c =camera, a = actor

#Views: 16
view dist(cm): 125
slit dist(cm): 125
view width(cm): 30
screen width(cm): 30
x pixels: 800
aspect ratio: 1.33333
☐ stream sets
☐ stream frames
☐ bandlimit

Figure 3.5: *The custom built GUI allows loading and manipulation of 3D models. The left view shows the imagery that will be captured by the cameras. and the right view shows the position and proportion of the object relative to the cameras. Text boxes in the bottom of the window allow the dimensions of the real display to be entered so that any changes in the optics can easily be taken into account. A button begins the rendering process, and loads the views into the display.*

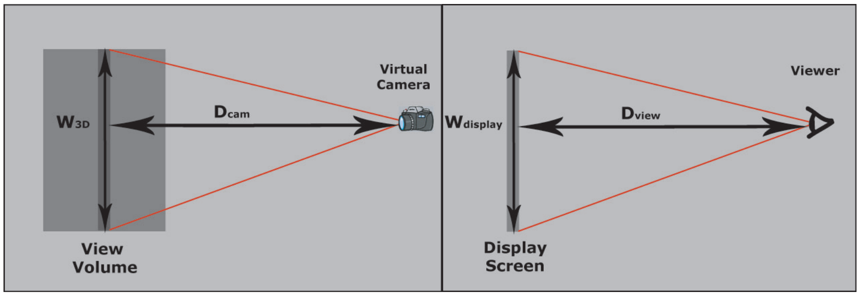


Figure 3.6: *The figure demonstrates how the virtual scene that is used to capture 2D views from a 3D model should match the geometry of viewing the display.*

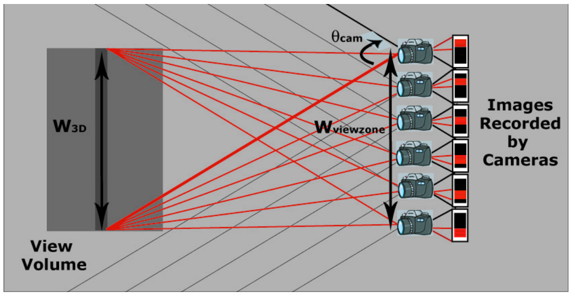


Figure 3.7: *The red portion of the images captured by the cameras is the region that contains information about the intended view volume. The rest of the information can be discarded.*

Chapter 4

Theoretical Analysis and Experimental Results

A view-sequential display was built based with the characteristics described in chapter 5. The image quality is compared with the predicted quality according to theoretical calculations. Particularly, the dimensionality of the image is discussed, as well as distortion and 2D imaging quality.

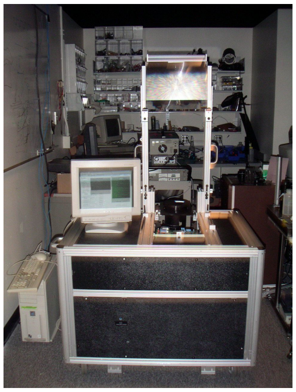


Figure 4.1: *The picture shows the front of the display). A Fresnel lens is hoisted above the projection engine, and sits at about 5' above the ground.*

4.0.4 Horizontal Parallax Only (HPO) Systems

There are several restrictions on Horizontal Parallax Only systems, which are particularly important to the field of 3D displays. HPO systems only present the correct perspective information when observed from a fixed viewer distance. If a viewer migrates from this distance, then a distorted image will be observed. It is possible to distort images to provide the correct perspective information, but the viewing distance must be predetermined [17]. This is a requirement that is not placed on full-parallax systems, but it is the price that is paid for a dramatic reduction in bandwidth.

The horizontal perspective information is accurate regardless of viewer position, but the vertical perspective of the image is fixed to the display surface. This can be considered to cause an astigmatism in points which are converging in front of or behind the display. There is a maximum tolerable convergence which humans generally considerable acceptable, limiting the maximum depth in HPO systems for comfortable viewing conditions. In practice, the exact amount that is tolerable is much more generous than expected, as evidence in the abundance of deep rainbow holograms viewable in museums around the world. It is good practice though, for 3D display designers to consider upgradability between HPO and full-parallax systems if they choose to develop for interactive applications where smaller bandwidths are required.

4.1 3-D Imaging Analysis

The analysis of multi-view displays can be approached in two different ways. Both methods attempt to characterize the maximum depth that a display can accurately show based on the number of pixels and views, and the size of display, pixel, and viewzones.

4.1.1 Geometric Analysis

The maximum depth Z_{max} that a multi-view display is capable of conveying can be derived by considering a point that the display is attempting to focus in front of the screen. Z_{max} should be such that a shift in viewing position will cause a transition to neighboring pixels from the projection of that point from the eye onto the display screen (see Figure 4.2).

If points which are deeper than Z_{max} are shown on the display, there is a discontinuity between the images that viewers observe when transitioning from one viewzone to the next (see Figure 4.3). This determines that the maximum depth a multi-view can display will be:

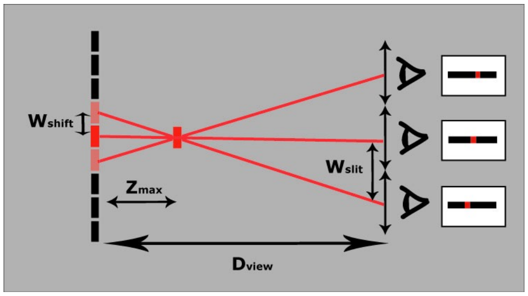


Figure 4.2: *By similar triangles, the maximum depth Z_{max} that a multi-view display can convey is determined by constraining that projections of a point onto the display screen abut each other when viewed from adjacent viewzones.*

$$Z_{max} \leq \frac{D_{view}}{1 + \frac{W_{pix}}{W_{slit}}} \tag{4.1}$$

or

$$Z_{max} \leq \frac{D_{view}}{1 + \frac{N_{pix}}{N_{slits}}} \tag{4.2}$$

because we chose to have our viewzone width $W_{viewzone}$ the same size as our display width $W_{display}$. For the view-sequential display in discussion, the following relevant quantities are:

- $D_{view} = 125cm$
- $W_{viewzone} = 30cm$
- $N_{slits} = 16$
- $W_{slit} = 1.9cm$
- $W_{display} = 30cm$
- $N_{pix,x} = 800$
- $W_{pix} = 375\mu m$
- $Z_{max} = 2.45cm$

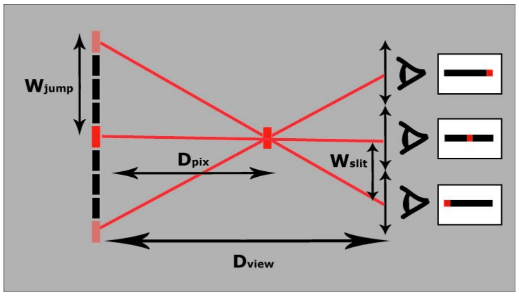


Figure 4.3: *If the constraints of Figure 4.2 are not met, then the position of the point appears to jump from one position to the next when viewers transition between adjacent viewzones.*

4.1.2 Fourier Analysis

This section attempts to expand upon the results of the previous section by introducing the analysis of spatial frequency content within images.

Diffraction Limited Pupil Size

As pixel sizes approach the wavelength of light, they no longer merely modulate the amplitude of light. The edges of extremely small pixels cause diffraction which limits the minimum divergence of rays emanating from an SLM[22]. This limits the angular separation between adjacent views according to the relation:

$$\delta\theta \geq \frac{\lambda}{W_{pix}} \tag{4.3}$$

If this condition is not achieved, then adjacent views will be visible regardless of viewing position. This is not of particular importance for prototype displays, where the number of views must be compromised in order to achieve a sufficient viewing angle.

Filtering, Bandlimiting, and Antialiasing

At first glance, the quantity Z_{max} would appear hopelessly small, and indeed, it indicates that in order for our display to accurately fill a cube with depth equal to its width, then the number of viewzones present must equal the number of pixels. This is typically considered the optimum sampling amount, though in some cases there are further restrictions on this amount. It is intuitive to consider that the same amount of information must be used to generate a volumetric display of the same size, and indeed, this number can be represented by the Space-Bandwidth Product (SBP) of the system.

What is not generally considered is that it is possible to represent images of greater depth than Z_{max} by applying a filtering technique to the 2D images that are displayed on a multi-view system [9]. The previous statement qualifies the property that the quantity Z_{max} is not solely dependent on the information content of a display. A technique of *filtering* or *bandlimiting* the parallax information can result in deeper images, at the sacrifice of reduced resolution. Figure 4.4 demonstrates how this process is carried out. The spatial frequency (it size W_{3D}) of the 3D pixel is increased to match the appropriate sampling of parallax information.

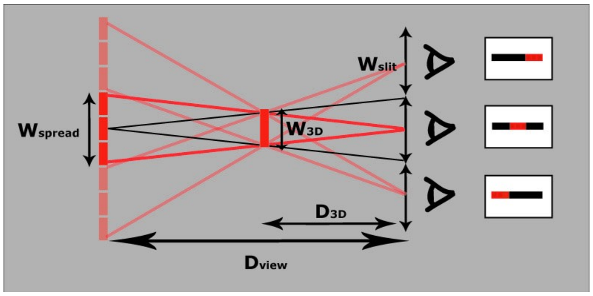


Figure 4.4: *If proper bandlimiting conditions are met, than the spatial extent of images shall increase as points are represented further away from the display screen. The effect is similar to depth of field in cameras. The filtering assures that images from neighboring views abut each other on the display screen.*

It is possible to analyze the optimum slit width by considering the slit as a limiting aperture, and finding the Optical Transfer Function (OTF) of the image that the eye projects onto the retina. There is a phase error in the image that a viewer observes that represents a difference in the radius of curvature of a point on the display and an intended 3D point. This error is analogous to a focusing error, and can be analyzed in terms of the quantities D_{view}, D_{3D}, and W_{pupil}. St. Hilaire derives the OTF in reference[10], with the result:

$$H(f_x) = \Lambda(\frac{f_x}{f_0}) sinc[(\frac{D_{view} - D_{3D}}{D_{view} D_{3D}})(W_{pupil} D_{view} f_x)(1 - \frac{f_x}{f_0})] \tag{4.4}$$

Where:

$$f_0 = \frac{\lambda D_{view}}{W_{pupil}} \tag{4.5}$$

For this equation, Λ(x) is the triangle function, and sinc is defined by sinc(x) = sin(πx)/πx. The equation reduces to a triangle function with bandwidth $\Delta f_x = 2f_0$ when $D_{view} = D_{3D}$ and the 3D point is on the display screen. The OTF is degraded as the "focusing error" increases and points in front of or behind the screen are shown. The effect is similar to a low pass filter though it shows more complicated behavior for large errors.

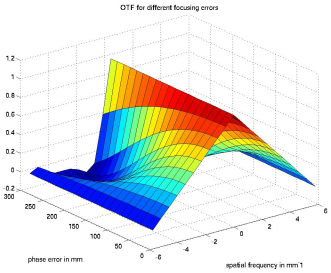


Figure 4.5: *The image demonstrates the degrading of the OTF for images of different depth where the pupil of the eye (3mm) is taken to be the limiting aperture.*

This analysis merely states that each point from each 2D view should be made to look as though they are "out of focus" at the screen surface if they are not meant to appear there by the 3D display. In order to correctly display points at different depths, all points should be filtered by equation 4.4 for the appropriate values. This filtering has the same effect as the geometrical interpretation prescribed: points displayed further away from the display surface should occupy a larger portion of that surface. Figure 4.9 shows a comparison of the PSF with different focusing errors for the eye, 5 unfiltered pinhole cameras, and 5 filtered pinhole cameras. From this graph it is apparent that the filtering action matches the intended effect much more closely than without.

This analysis is particularly important to the display of CG images which are rendered through a pinhole camera, and thus have an infinitely narrow PSF. When the image of the shutter slits are smaller than the pupil of the eye, the slits perform the proper bandlimiting of the 2D image via equation 4.4. This means that the Point-Spread function for each slit will be wide enough to ensure that images from adjacent views abut each other on the display screen. When the slits are larger than the pupil of the eye, the eye becomes the limiting aperture for the system. The consequence of this is that each 2D image is not projected through the appropriate bandlimiting aperture. The images displayed on the screen should therefore be prefiltered with equation 4.4 for $W_{pupil} = W_{slit}$ in order to correctly match the display properties of the screen. This is demonstrated in Figures 4.6-8, where a single vertical line is recorded by pinhole cameras, then filtered with the appropriate OTF.

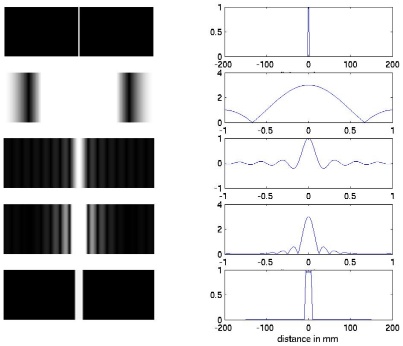


Figure 4.6: *The top image is a snapshot from a pinhole camera of a vertical line intended to project in front of the display screen. The next images show in sequence the frequency content of this image, the appropriate OTF for a system, and the product of these two quantities. The last image shows the appropriate spread that will occur.*

Two techniques for filtering computer graphics imagery were explored which are shown in Figure 4.10. In the first technique, the maximum depth Z_{max} of a 3D object is used to calculate the OTF from equation 4.4, then

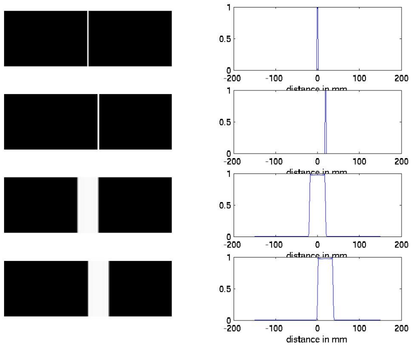


Figure 4.7: *The top two images demonstrate the displacement from two perspective views of a vertical line intended to float in front of the display. The following two images show how the images abut one another after they have been filtered appropriately with equation 4.4*

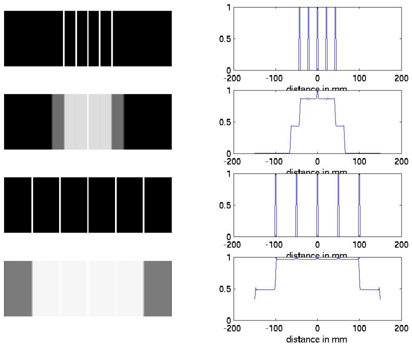


Figure 4.8: *The top image shows five images superimposed which are taken from pinhole camera projections of a vertical line intended to appear 100mm in front of the display. This image demonstrates that the net PSF from all views is not sampled correctly. The next two images show the frequency content of each image, and the appropriate OTF for the system. The last image demonstrates that after filtering, the PSF is correctly sampled and fills the display screen without discontinuities.*

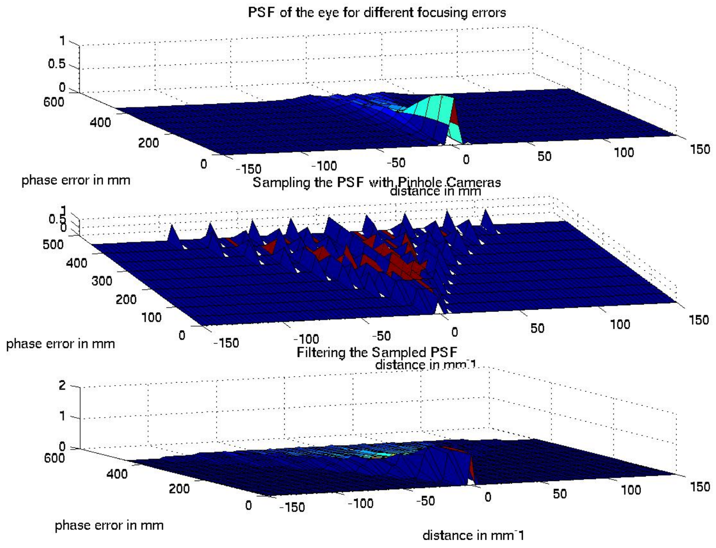


Figure 4.9: *The top image shows the PSF for different depths where the eye is the limiting aperture. The next image shows the sampling of MTF from a 5 view system with the same parameters as the one in this thesis. The next example shows the previous image filtered by the appropriate OTF given by equation 4.4 and shows much greater similarity to the first image.*

each image is filtered with that equation. The result is that the maximum horizontal dimension for points at all depths is constrained to the width of points at Z_{max}. The alternative method is more computationally expensive, but allows points of smaller width to be displayed at points closer to the display screen. In this method, a sequence of images across each slit are averaged to create the single image which will be projected. The number of images for each slit is chosen so that points at Z_{max} are projected onto the display screen without discontinuities. A third possible method of filtering is to associate the appropriate MTF with each vertex of the 3D object while rendering images so that it projects to the appropriate size on the displays screen. This method was not implemented, but it should provide dramatically quicker times to generate images.

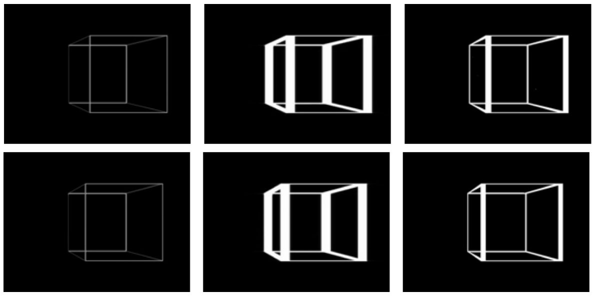


Figure 4.10: *The left images are computer graphics renderings of a box whose base touches the display screen, and with depth equal to 350mm. The middle images show the filtering technique and the rightmost images show the averaging technique.*

4.2 Demonstrated Image Quality

The display produces very satisfactory 3D images, with exceptional 2D image quality. The color depth of the images is remarkable, and the images are quite bright. The display has a brightness of $40 cd/m^2$ and a contrast of about 60:1. Figure 4.11 demonstrates the details and colors that the display is capable of. Images as deep as 350cm have been demonstrated using the filtering technique described earlier in this chapter.

Figure 4.11: *Two images which have not been filtered. The image were captured by a camera whose pupil was placed within a single slit.*

4.3 Experimental Geometries

The main drawback of the display is the limited viewing angle that it allows, which is particularly frustrating for a group of visitors who must wait their turn for a glance at the display. Often this wanes enthusiasm, and the display is perceived as less impressive. It certainly detracts from one of the main features of multi-view displays, which is to allow multiple viewers to view the display simultaneously.

For this reason, the *Ultragram* [17] approach to 3D imaging was experimented with. This technique was designed to allow greater viewing zones for holographic stereograms by allowing viewers to observe images from distances other than the projected slit distance. The appropriate distortions are computed based on the relative slit and observer distances. This technique is relatively easy to implement, and we utilized it to allow viewers to observe our display from further back than the image of the shutters in our display. This allowed the correct viewing of images with a greater viewing width, though the viewing angle remained the same.

Chapter 5

Conclusions

View-sequential 3D display systems show promising features that may soon be commercially viable. The fact that other 3D displays are already on the market is of great importance, for the greatest challenge that a successful display will face is customer awareness. The world at large is not greatly aware of the state of 3D display technology, and all marketing and advertising of 3D displays assists the entire 3D display community.

5.1 Discussion on Size Constraints of the System

The display that has been presented has the great disadvantage of being physically bulky. At a time where flat-panel displays are extremely popular, the question naturally arises: *How do you make it smaller?* The system that is presented certainly can be reduced in size by several factors, possibly even an order of magnitude, but there are certain constraints on the system that cannot be overcome.

The size of the projection lens is of critical importance to the system, for it determines the maximum viewing angle that can be achieved. The lens must have a large Numerical Aperture (NA) if it is to be placed close enough to the field lens as to provide a reasonable viewing angle. Generally, such a lens will be expensive, and will probably consist of a compound of several lenses, thus increasing the size of the system. Typically the magnification of the objective lens must be greater than 4 because SLMs are currently built so small. In order for the image to be magnified, it must be placed greater than twice the focal length from the objective lens.

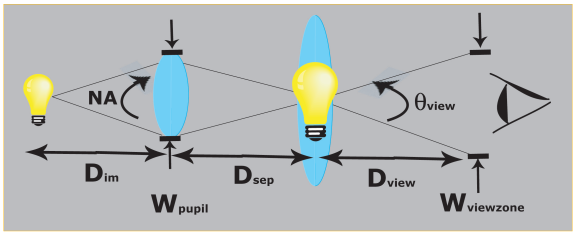


Figure 5.1: *There is a fundamental limit set on the viewing angle θ_{view} that is set by the numerical aperture of the objective lens. The focal length of the objective lens is designated f_1, and the focal length of the field lens is designated f_2.*

There are two constraints which limit the viewing angle of this type of system:

- The Numerical Aperture of the objective lens typically must be greater than one.
- The objective lens must magnify the image of the SLM.

By the paraxial approximation the magnification for the field lens is:

$$M_2 = \frac{W_{viewzone}}{W_{pupil}} = \frac{D_{view}}{D_{sep}} \tag{5.1}$$

The angle of the cone of rays converging from the objective lens which focus the SLM image sets the viewing angle of the display. In the paraxial regime, this means that the ratio $\frac{W_{pupil}}{D_{sep}}$ defines the viewing angle. The magnification for the objective lens is:

$$M_1 = \frac{D_{sep}}{D_{im}} = \frac{NA_1 \times W_{pupil}}{D_{sep}} \tag{5.2}$$

Where $NA_1 = \frac{D_{im}}{W_{pupil}}$. Rearranging gives:

$$\frac{W_{pupil}}{D_{sep}} = \frac{NA_1}{M_1} \tag{5.3}$$

This together with the inequalities $M_1 \geq 4$ and $NA_1 \geq 1$ and the relation $\theta_{view} = 2 \times \frac{W_{pupil}}{2 \times D_{sep}}$ gives the result:

$$\theta view \lessapprox 15^o \tag{5.4}$$

The results of this analysis demonstrate that it is hopeless to use microdisplays with this technique because magnifying the SLM image places such a large restriction on the viewing angle of the display. Indeed our display only had a viewing angle of $\theta_{view} = 13.8^o$. The initial systems built by Cambridge University did not need magnification because the CRTs used were very large, and thus wider viewzones could be achieved. In any case, if a viewzone is too large, it will violate the paraxial regime and cause distortions to viewing positions that are too far off-axis. Other techniques for projection-based multi-view displays such as Travis describes in [5] will have to be developed if these displays are to achieve success.

Initially, it was considered that one of the advantage of using DMDs over CRTs was to ensure that the minimum resolution of the eye be met. Our view-sequential display had a pixel size of .375mm, which matches the minimum resolution of the eye at a viewing distance of 430mm. Viewers are ensured to have this condition met, and discontinuities in the observed image are avoided. In retrospect, I consider this to be a waste of valuable bandwidth. I believe that it would be a much greater advantage to have fewer and larger pixels, and use the extra bandwidth for more parallax sampling so that deeper images can be displayed.

In this respect, this display does not make improvements on the size constraints of view-sequential displays. Smaller pixel generating devices have been implemented successfully, but the size constraints inherent in the system have not been overcome. Indeed, an improved version of the system could utilize a larger diffuser image, reducing the magnification needed by the objective lens. Such a system could have a much larger viewing angle, but would *increase* the size to greater than previous versions. The success of the display is in demonstrating a greater bandwidth system with much greater brightness. These attributes are important enough to warrant further investigation, and be featured in following systems which successfully reduce the size of the system.

5.2 Suggestions for Further Work

The results of this thesis are remarkably impressive, but several improvements on image quality could be made. The color depth of the display should remain the same, for the richness in colors greatly attributes to a pleasurable viewing experience. The resolution of the display screen should be sacrificed to achieve greater sampling of parallax information, but this is not possible with the current projection engine. The next generation DMD projection system is capable of displaying images at twice the speed of the engine used for the current system, which could allow twice as many views to be displayed. The newer DMDs also allow partial frames to be updated at increased frame rates. This would allow even more views to be included in exchange for lower resolution images sent to the display screen. A display with $1/16^{th}$ the maximum resolution would still provide 256 samples on the display screen, and allow an equivalent amount of parallax sampling to be achieved. Such a display would be capable of conveying as much depth as current commercially available volumetric displays, but with the added feature of occlusion.

5.2.1 Light Loss

Luckily, light loss was not a concern for this display. However, if a new system were to be built, it would necessarily have more views, and thus an increased light loss as well. One possible improvement might be to use two one-dimensional diffusers: one in the place of the narrow-angle diffuser that we used, and another at the location of the field lens. The former would allow considerably more light to enter the optical relay system, and the latter would allow the vertical viewing angle to be increased.

5.2.2 Optics

It is the authors opinion that much stronger displays could be built by eliminating the relay optics that are used for the current system. If fast enough transmissive modulators were available, the system described in Figure 5.2 could be built. However, a variation of this theme could be built which utilizes an OASLM and a Wedge$^{\mathrm{TM}}$light valve. Such a system would relieve the need for a large projection lens, introducing the flexibility of scaling and increased viewing zone. It could also build a system which does not take up much more space than a flat panel display because a small projection engine could be used to write onto the OASLM and overhead illumination could be used to read from it. The system relies on suitable strong sources so that each view zone will have the appropriate brightness, but advances in bright LEDs seem to point in the favor of such developments happening soon. It would also be possible to use a diffusely backed FLCD shutters as fast switching sources, but if LEDs can be adapted, then displays could be manufactured at a greatly reduced price. The display could also be easily adapted for quasi multi-view/holovideo approaches that are described in detail in Reference [22]. Cambridge University has demonstrated the principle of such a display, and the main challenge that the system faces is the manufacture of sufficiently large OASLMs to produce reasonably large display screens.

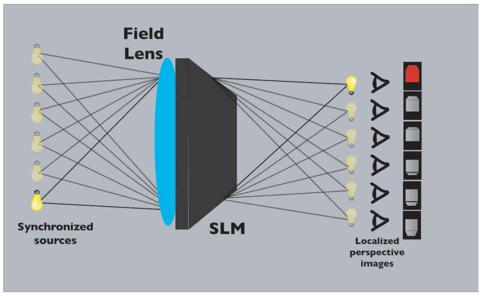


Figure 5.2: *An ideal view-sequential system uses only a fast SLM, a lens, and several light sources which are synchronized with the SLM. The display shows a different perspective image for each light source, and the entire sequences of light source/image pairs is updated within* $1/50^{th}$ *sec.*

5.2.3 Rasterization Engine

The most interesting challenge that view-sequential displays face is the development of a sufficiently fast rasterization and graphics processing engine to handle the simple view-transformations that are necessary to generate the sequence of images that need to be displayed. The task is hardware intensive, but could very feasibly result in the real-time generation of 3D computer models on the display. Current graphics processors are as fast as CPUs, and if one can be adapted to feed directly to the frame buffers which a projection engine reads from, then images can be updated in real

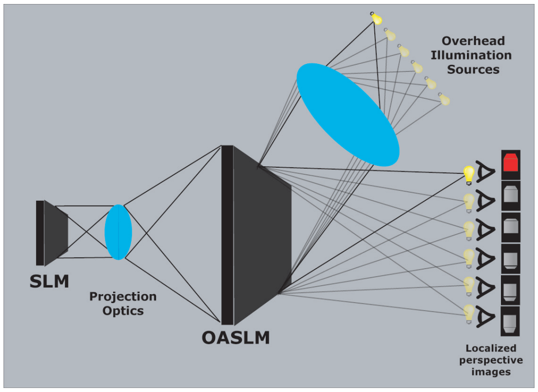


Figure 5.3: *A new type of view-sequential system could use a microdisplay as the SLM, an OASLM, and several light sources which are synchronized with the SLM. This design has been demonstrated at Cambridge University and could prove to be the most desirable technique for view-sequential displays. The design offers the potential for larger displays and small projection devices, but large enough OASLMs are still not available.*

time. As an example consider the rasterization engine developed by Actuality Systems, which feeds to frame buffers via a 3Gb/sec bus. The display contains 100Million voxels at 3bits/voxel, which means that a new set of data can be sent to the display within $1/10^t h$ second. The system. Such a system could be developed for a view-sequential display and allow real-time interactivity with 3D models, greatly enhancing the display's functionality.

5.3 Final Remarks

The display described in this thesis will be deconstructed by September 2003. The work will be continued at Cambridge University, where many of the ideas suggested in this chapter will be explored. There is limited interest from some of the 3D display companies mentioned in this thesis, but it is unlikely that the display shall exist again in its current form. The research has provided its authors with valuable insight into developing 3D technology, and most importantly, provided the author with experience in demonstrating 3D technology to a wide variety of viewers. It is apparent from the response received that the current display is not sophisticated enough to lend itself to applications that would require its continued existence. It has managed to capture the imagination of those who have had the opportunity to see it, especially the author of this thesis, who hopes to be able to continue this work in whatever capacity is available to him in the future.

The display that was built did revive an interest in view-sequential 3D and demonstrated that the progress of fast switching binary devices has opened up new possibilities for this technique. The time-multiplexing principle was at the heart of the success of this display, as with the Perspecta $^{\text{TM}}$, Mirage, and DepthCube$^{\text{TM}}$ displays that Actuality, Stereographics/Christie Digital, and Vizta3D currently sell. The *time* and *spatial-multiplexing* techniques compete for commercial viability, and until recently, *spatial-multiplexing* had predominantly found itself in the lead. The success of the Cambridge/MIT view-sequential 3D display marks the strength of the *time-multiplexing* technique, and this thesis has aimed to outline its commercial potential. This this thesis has successfully demonstrated that binary pixel generating devices can easily and effectively be integrated into 3D display technology. This thesis has also outlined the advantage that they have brought to a variety of 3D displays.

Binary SLMs and *time-multiplexing* techniques will not bring 3D displays commercial success and expansive popularity on their own. Currently, displays which use *spatial-multiplexing* techniques such as the 4D-50$^{\text{TM}}$ by 4DVision, Synthagram$^{\text{TM}}$222 by StereoGraphics, and 2D/3D by Sharp are quite elegant solutions because they are so similar to the 2-D displays which the public is largely familiar with. The role that these displays play in the market is a necessary one. Once there is a larger awareness of the existence of 3D displays, there will be room for more expensive displays that are capable of much more dimensionality. This time may be soon, and if it is, view-sequential 3D displays could make an enormous impact on the quality of commercially available 3D displays. Though several important problems need to be overcome before view-sequential 3D is an ideal solution, it has provided strong enough results already to compete in the commercial market. If it is adopted by a company that is interested in developing the technology further, the technique will likely receive lasting commercial success.

The 3D boom may be just around the corner, or it may always be just out of reach. Though the outcome of the 3D display industry is uncertain, and its reputation questionable, stunning 3D images continue to be displayed with greater sophistication than ever before. Scientists, artists, businessmen, and laymen alike are attracted to the stunning realism that is offered from 3D displays. Even if they forever disappear, 3D displays have made their mark on humanity by capturing the imagination of people from all walks of life. With any luck, they will continue to do so for a long time.

Bibliography

[1] Shigeyuki Baba Akira Shirakura, Nobuhiro Kihara. Instant holographic portrait printing system. *Proc. SPIE*, pages 3293–35, 1998.

[2] A.R.L. Travis, John R. Moore, Neil A. Dodgeson, Stewart R. Lang. Time-multiplexed color autostereoscopic display. *SPIE Proceedings*, 2653, February 1996.

[3] A.R.L. Travis, Nathan Marston. 3D displays with wide fields of view.

[4] Stephen A. Benton. Experiments in holographic video imaging. *SPIE Holography*, IS 8:247–267, 1991.

[5] C.H. Chen, A.R.L Travis. Wide Field of View Optics for Flat Panel 3D displays. *Society for Information Displays*, 1999.

[6] Chris Slinger, Bob Bannister, Colin Cameron, Stuart Coomber, Ian Cresswell, Peter Hallett, John Hughes, Victor Hui, Cliff Jones, Richard Miller, Vikki Minter, Doug Pain, Dave Scattergood, David Sheerin, Mark Smith, Maurice Stanley. Progress and prospects for practical electro-holographic display systems. *SPIE Conference on Pracrical Holography XV*, (4296A-03), January 2001.

[7] Evans and Sutherland. http://www.es.com.

[8] G. Favalora, J. Napoli, D.M. Hall, R.K. Dorvalm M.G. Giovinco, M.J. Richmond, W.S. Chun. 100 million-voxel volumetric display. *AeroSense 2002 -for Cockpit Displays IX: Displays for Defense Applications*, 2002.

[9] Michael W. Halle. Holographic stereograms as discreet imaging systems. *SPIE Proceedings 2176 "Practical Holography VIII"*, February 1994.

[10] Pierre St. Hilaire. Modulation transfer function and optimum sampling of holographic stereograms. *Applied Optics*, 33(5), February 1994.

[11] Pierre St. Hilaire. *Scalable Optical Architectures for Electronic Holography.* PhD thesis, MIT Media Lab, 1994.

[12] Stanley Seokjong Hong. Surface acoustic wave optical modualtion. Master's thesis, MIT, 2001.

[13] Zebra Imaging. http://www.zebraimaging.com.

[14] K. Perlin, S. Paxia, J. Kollin, D. Kristjansson, C. Poultney. Recent Advances in the NYU Autostereoscopic Display. *SIGGRAPH 2000 Conference Proceedings*, 33(3), 2000.

[15] K. Perlin, S. Paxia, J. Kollin, D. Kristjansson, C. Poultney. An Autostereoscopic Display. *SPIE Proceedings*, 4297, January 2001.

[16] Liti3D. http://www.Liti3D.com.

[17] Michael W. Halle, Stephen A. Benton, Michael A. Klug, John S. Underkoffer. The Ultragram: A Generalized Holographic Stereogram. *SPIE Proceedings 1461 "Practical Holography V"*, February 1991.

[18] N. A. Dodgson, J. R. Moore, S. R. Lang, G. Martin, P. Canepa. A 50" time-multiplexed autostereoscopic display. *SPIE Proceedings, "Stereoscopic Displays and Applications XI"*, 3957, January 2001.

[19] StereoGraphics Corporation White Papers. Synthagram ^{TM}handbook, February 2003.

[20] Elroy Pearson. Mems spatial light modulator for holographic displays. Master's thesis, MIT Media Lab, 2001.

[21] Sharp Press Release, March 2003. http://sharp-world.com/corporate/news/030304.html.

[22] A.R.L. Travis. The Three-Dimensional Display of Video Images. *Proceedings of the IEEE*, 85(11), November 1997.