

Holographic Television: Measuring Visual Performance with Holographic and Other 3D Television Technologies

by

James Barabas

M.S., Massachusetts Institute of Technology (2006)

Submitted to the Program in Media Arts and Sciences,

School of Architecture and Planning,

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Media Arts and Sciences

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2014

© Massachusetts Institute of Technology 2014. All rights reserved.

Author

Program in Media Arts and Sciences
Jan 10, 2014

Certified by

V. Michael Bove, Jr
Principal Research Scientist
Program in Media Arts and Sciences
Thesis Supervisor

Accepted by

Pattie Maes
Associate Program Head
Program in Media Arts and Sciences

**Holographic Television: Measuring Visual Performance with
Holographic and Other 3D Television Technologies**

by

James Barabas

The following people served as readers for this thesis:

Thesis Reader _____

George Barbastathis
Professor
MIT Department of Mechanical Engineering

**Holographic Television: Measuring Visual Performance with
Holographic and Other 3D Television Technologies**

by

James Barabas

The following people served as readers for this thesis:

Thesis Reader _____

Ramesh Raskar
Associate Professor
Program in Media Arts and Sciences

Acknowledgements

I like to thank those that contributed to this dissertation, through feedback, collaboration and critique, and through financial and moral support.

V. Michael Bove Jr, George Barbastathis, Ramesh Raskar

Quinn Smithwick, Dan Smalley, Sunny Jolly, Wendy Plesniak, Michael Halle

With generous support from:

My family: Sarah, Emily, Kathy, David, Pat.

MIT: Kristin, Linda, Santiago, Edwina, Jinjoo, Vram, Ermal, Dan, Laura

Equipment Donations: zSpace Inc., NVIDIA University Relations

Funding: MIT Media Lab Consortia

Contents

1	Introduction	11
1.1	Motivation	12
1.2	Related Work Overview	14
1.2.1	Holograms – computation and display	15
1.2.2	Taking Pictures	16
1.2.3	Evaluating holographic television	17
1.3	Summary of Results and Contributions	18
1.4	Overview of this document	18
2	Human Vision Considerations for Holographic Display	19
2.1	Introduction	19
2.2	Holography	20
2.3	Computer-generated holography	20
2.4	Displaying Holograms	21
2.5	Spatial frequency and view zone	22
2.5.1	Depth Perception	23
2.5.2	Image appearance	30
2.6	Conclusions	31
3	Computation for Holographic Television	33
3.1	Overview	33
3.2	Optical Holography	34
3.2.1	Mark II Holographic Display	35
3.2.2	Reconfigurable Image Projection (RIP) Holograms	41
3.3	Initial work	43
3.3.1	Modifications using Framebuffer Objects	44
3.3.2	Diffraction-Specific Coherent Panaramograms	45
3.3.3	Interference-modelled Holograms	48
3.3.4	Two-view Holographic Auto-stereograms	49
3.4	Software architecture	50
3.4.1	Software Programs	52
3.4.2	Internal representations	54

3.4.3	Future work	55
4	Vision Experiments	57
4.0.4	Introduction	58
4.1	Pilot Experiments	58
4.1.1	Motivation and Related Work	59
4.2	Stimulus design	60
4.3	Displays	61
4.4	Data Collection	62
4.5	Data Analysis	63
4.6	Time-To-Perceive Experiment	64
4.6.1	Methods for line-plane time-to-perceive experiment	64
4.6.2	Results comparing performance viewing holographic and autostereoscopic images on the Mark II display:	68
4.6.3	Results comparing performance viewing on holographic and LCD displays:	69
4.7	Interactive pointing experiment	71
4.7.1	Task	72
4.7.2	Discussion	73
4.8	Layer Discrimination Experiment	74
4.8.1	Methods	75
4.8.2	Discussion	78
5	Conclusion	81
5.1	Contributions	81
5.2	Closing Thoughts	83
A	Illustration of Distortion from Mark II Horizontal Optical Design	85

Chapter 1

Introduction

Abstract

We are surrounded by visual reproductions: computer screens, photographs, televisions, and countless other technologies allow us to perceive objects and scenes that are not physically in-front of us. All existing technologies that reproduce images perform engineering tradeoffs that provide the viewer with some subset of the visual information that would be available in person, in exchange for cost, convenience, or practicality. For many viewing tasks, incomplete reproductions go unnoticed. This dissertation provides a set of findings that illuminate the value of binocular disparity, and ocular focus information that can be provided by some three-dimensional display technologies. These findings include new experimental methods, as well as results, for conducting evaluations of current and future display technologies.

Methodologies were validated on an implementation of digital holographic television, an image capture and reproduction system for visual telepresence. The holographic television system, allows viewers to observe, in real-time, a remote 3D scene, through a display that preserves focus (individual objects can be brought into optical focus at the expense of others), and horizontal motion parallax (depth and other geometry of objects appears natural over a range of head movement). Holographic television can also emulate other precursor 2D and 3D display technologies. This capability was used to validate the evaluation methodologies (meta-evaluation) by comparing visual performance on simulations of conventional displays to results of past studies by other researchers.

1.1 Motivation

Holography is a versatile tool for capture and display of 3D scenes. It allows us to create visual reproductions of computer graphics or real objects that reproduce more of the visual, structure-revealing information present in natural scenes. Over the past few years, I have been working to enable the creation of video-rate animated holographic images, that are computed as they are displayed. Furthermore, a gaming-driven interest in live capture of live three-dimensional human movement has lead to the availability of new consumer cameras that record not only color of objects in front of them, but also distances to points visible to the cameras[Freedman et al. 2008]. The combination of these developments has allowed me to develop a platform for generation and display of holograms that reflect, in real-time the appearance of real moving objects [Barabas et al. 2011]. Some previous demonstrations of holographic television have been literal transmissions of raw holographic data [Hashimoto et al. 1992],[Kim et al. 2002]. These coherent capture approaches required impractical laser-illumination, and images were very high bandwidth and could not be digitally stored or manipulated. Other demonstrations have used incoherent [Shaked August 2009] photography, but were too slow to provide live interaction [Blanche et al. 2010]. One promising demonstration did achieve real-time incoherent holographic television [Yamamoto et al. 2010], but the display was limited by a small image and viewing zone. My current holographic television platform overcomes these limitations, providing a real-time holographic image that is viewable on a phone-sized display. The system serves as both an intuitive proof-of-concept demonstration of live digital holographic television, as well as a platform for vision experiments where the experimenter has unprecedented flexibility and control of how live three-dimensional images are presented to the viewer.

Tools for remotely conveying the appearance of objects are in wide use today, but a new tool, digital holographic television, has the potential to convey substantially richer information than methods currently in use. Holographic images contain additional information about the shape and appearance of objects that are not available in traditional imaging.

A live television system employing holographic display enables both new applications of remote viewing, as well as improvement of efficiency or cogency in existing uses. This dissertation describes an implementation of a holographic remote-viewing system, and describes potential benefits of this technology. Additionally, the dissertation describes designs for and results of experiments evaluating this system. Experiments focus on viewers' ability to discern spatial relationships in scenes, and explore how the ability to discern these details changes across different methods for conveying depth.

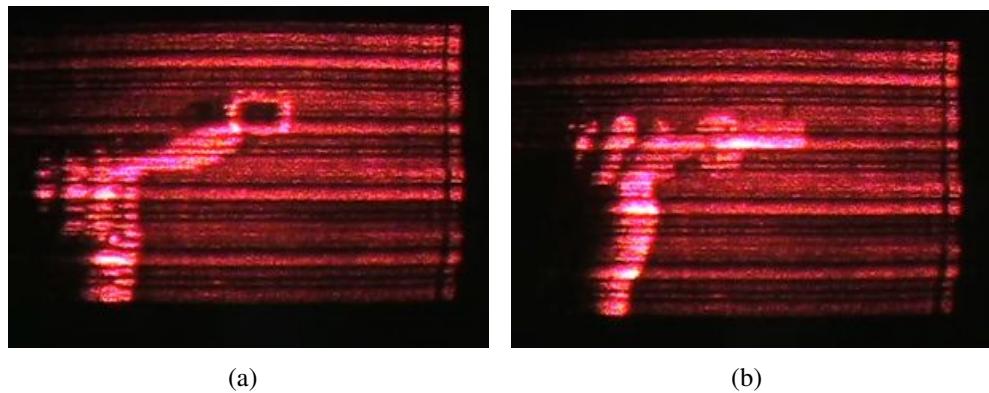


Figure 1-1: Example frames from real-time holographic television display of hands performing a ring capture task

Live images are viewed remotely for a variety of reasons. Live television is a ubiquitous example of this – images are captured by a traditional two-dimensional camera, and are displayed nearly instantaneously on a distant two-dimensional screen. Uses for live television and related live-viewing technologies include entertainment [Bourdon 2000] (such as the viewing of sporting events), mass communication of faraway events, and securing or monitoring facilities. Uses also include face-to-face communication over distance [Raskar et al. 1998], laparoscopic or other remote surgery [Ballantyne 2002], and operation of equipment in inaccessible locations such as underwater [Bleicher 2010], or in space[Kim and Bejczy 1993]. In nearly all of these uses, remote viewers are using the images they see on a screen to build a mental model of a space or activity in a remote scene. Frequently, the understanding of these spatial relationships is critical to the task at hand. For sports

spectators, being able to discern where players and equipment are on a field is primary to the viewing experience. For remote operation, discerning position of tools relative to other equipment is essential. Two-dimensional image display technologies currently in widespread use are able to convey the three-dimensional configuration of distant spaces, but viewers have fewer visual cues to the true geometry of the subject than they would have in person [Murphy 2004]. Viewers using two-dimensional screens can rely on cues like shading, foreshortening, color tint, texture and movement for help in discerning 3D scene structure . But unlike in person, natural binocular parallax (seeing different images in the two eyes), ego-motion parallax (being able to move around to see objects from a different perspective) and defocus (from viewing objects that are in front of, or behind the viewer's current plane of ocular focus) are not available. Holographic displays can provide these, or approximations to these missing pieces of information, giving viewers the potential to view and interpret scenes more naturally. Furthermore, more natural viewing has the potential to improve performance for spatial tasks that are mediated by display technology.

Evaluation of how reproduced images are perceived can inform and be informed by other work studying spatial perception. While display technology in a complete system can be evaluated, aspects of the viewing and interaction can also be isolated for careful study [Hinckley et al. 1994].

1.2 Related Work Overview

Related work relevant to this dissertation comes primarily from the fields of Optical Engineering, Computational Photography, Computer Graphics, and Human Perception. The physical manifestation of the holographic television builds on equipment from previous work done at the Media Lab, as well as mathematical models of camera geometry used in computational photography, and computer graphics. Methodology for evaluation of the holographic television system draws on other work in basic research on human visual per-

formance, and on work exploring the impact of television and other technologies on visual experience.

1.2.1 Holograms – computation and display

Holographic television uses wave properties of light to re-create the appearance of 3D scenes. The first description of optical reproduction of objects using interference was Gabor's note[1948] describing electron micrographs captured through electron-interference and reproduced by optical reconstruction. Following this work, Leith and Upatnieks [1962] [1965], developed new techniques for improving signal-to-noise ratio by separating reproduced scenes from undiffracted light. These holograms are the basis for most holographic capture and reconstruction techniques used today.

Holograms can be optically recorded directly to a photographic plate, much as patterns of light are recorded in a traditional film photograph. Interference between light striking the photographic plate directly (the “reference beam”), and light reaching the plate after bouncing off a scene creates tightly spaced patterns of light and dark that contain detailed information about how light emanated from the scene. When displaying these holograms, the photographic plate can be lit by a light source shining along the path (but in the reverse direction) of the original beam illuminating the plate. When illuminated in this way, a viewer can see a reproduction of the original scene in the light diffracting through the recorded features in the plate.

The holograms computed for my work draw from the above approach, but deviate from it in that they are computed, not recorded optically. These Diffraction-Specific Holograms [Lucente 1994] are holographic patterns computed from a mathematical model of a scene such that when they are transferred onto a physical plate and illuminated, they reproduce an image in the same way as a true hologram, but with some advantages. In a true hologram, the recorded holographic pattern results from the interference between light in the reference

beam, and light reflecting off the scene, but also has an additional component that results from interference between paths where light scatters off multiple, separate parts of the scene. This additional undesired component results in extra light appearing in a “haze” over the scene. Diffraction-Specific holograms eliminate this extra component, and therefore have the potential to produce sharper images.

Digital holographic television uses the Mark II holographic display to present images to the viewer. This display was developed at the Media Lab [St-Hilaire 1994] improving upon the design of a previous electro-holographic display [Jepsen 1989]. The Mark II display was subsequently modified by myself and others[Quentmeyer 2004] to allow for realtime image generation and display.

1.2.2 Taking Pictures

Digital holographic television also requires 3D scene representations that can be processed into holograms. My work on capturing scene appearance draws from other work on holography, as well as work performed for generic 3D scene capture. Work in this area includes that of Michael Halle, who developed a software framework for simultaneously rendering a single scene from multiple viewpoints, exploiting redundancies to compute the images quickly [Halle 1997]. His work is based on the representation of the range of appearances of an object as a camera moves across it. This higher-dimensional description of object appearance has also been described as the plenoptic function [Adelson and Bergen 1991], or the light-field [Gershun 1936]. First mentions of the concept were described by Faraday [1846], and have recently become a major direction of research in the computational photography community [Levoy 2006].

In a recent review of techniques for capture and computation of holograms using standard cameras [Shaked et al. 2009], the authors describe an opportunity these new techniques present for holographic video conferencing, although as of their writing, none had been

demonstrated. The television system implemented for this dissertation represents a realization of this opportunity.

1.2.3 Evaluating holographic television

Although holographic television could be evaluated in many ways, I have concentrated on the aspect I see as the most beneficial difference between this technology and its predecessors – the ability of the system to rapidly communicate spatial relationships.

One publication suggesting opportunities for holographic television studied operator's ability to judge whether a mobile robot could fit through an aperture [Moore et al. 2009]. In this study, Moore and her co-authors compared operator performance while remotely driving robots in the operators' direct view, compared with teleoperation of the robots using a robot-mounted camera. Compared with direct viewing, operators tended to overestimate the size of openings when viewed through a television camera. Although standard television is widely used for remote operation, holographic television has the potential to bring performance closer to that which is achieved when viewing directly.

Holographic displays provide for natural accommodation and motion parallax. Studies examining conflicts between accommodation and space-varying scene appearance [Hoffman et al. 2008] have reliably shown that these attributes of standard stereo displays cause both fatigue and misjudgments of scene geometry. Studies examining simulated binocular stereo and motion parallax in a visually-guided tapping task[Arsenault and Ware 2004] showed benefits of both of these added cues that increased as the task became more difficult.

1.3 Summary of Results and Contributions

This dissertation contributes documentation of apparatus and algorithms for real-time rendering on a holographic display. In addition, the dissertation contributes methods for and results of experiments for evaluating viewer performance with holographic displays. The experiments explore subjects ability to perceive spatial relationships represented in scenes on a holographic display, concentrating on experiments where subjects must discriminate the depth layering of multiple objects. The results of these experiments strongly suggest a performance benefit for viewers observing objects on holographic displays. Specifically, in comparison with stereoscopic and simulated stereoscopic displays, subjects needed less time to correctly make spatial assessments of relative position of pairs of distant objects when portrayed as holograms.

1.4 Overview of this document

This dissertation is organized into five chapters. This first chapter presents an introduction and provides context and motivation for the work presented. The second chapter provides an overview of human vision considerations for holographic display systems. The third chapter describes my methodology for computing and displaying holograms. The fourth chapter presents experimental methods and results for testing how human vision interacts with holographic displays.

Chapter 2

Human Vision Considerations for Holographic Display

2.1 Introduction

Holographic displays have the potential to reproduce the natural parallax and focusing affordances of real scenes. Although holographic displays are still far from maturity, no other display technologies have the potential to reproduce these affordances as accurately. This chapter reviews visual human-factors considerations for current and future holographic displays.

Holograms can produce highly detailed images that can have an uncanny visual realism. Image appearance can change instantly and naturally with the viewer's head motion, allowing viewers to experience direction-sensitive effects, like a shifting reflection on a glossy surface, or the discovery of objects hidden behind others. In addition, viewers can selectively focus their eyes on different parts of the image, allowing natural exploration of both near and far subjects, and producing natural defocus for other objects in the scene. Despite this potential, electro-holographic display systems are still in their infancy. Analog

holograms that fully produce the effects above have extremely high information content, requiring digital displays to move orders of magnitude more data than traditional 2D displays in order to achieve the same visual fidelity. Current holographic displays make computation and transmission practical by reducing information content in places that have the lowest perceptual impact. Some of these techniques are adapted from traditional display holography, while others were developed solely for digital holograms. This chapter reviews the capabilities of the human visual system that are of interest to designers of holographic display systems, highlighting the differences of holographic displays from other, more widely available display technologies. Although very little vision research has been conducted with holographic images, work in related areas can also provide insight into constraints on the perception of holograms.

2.2 Holography

For the purposes of this dissertation, I define, holography as the re-creation of appearance of a visual scene by exploiting diffraction to reconstruct light wavefronts. Wavefronts reconstructed by holography are then available to be viewed and focused by the eyes of an observer, in the same way a viewer can interact with the light emanating from a natural scene.

2.3 Computer-generated holography

The first holograms used interference of light to capture a representation of the wavelength-scale pattern of the amplitude and phase (often called “fringe pattern”) of light striking a photographic plate. Computer-generated holography uses a simulation of this process to compute such fringe patterns from digital representations of scenes. These computed

fringe patterns can be printed optically in photosensitive materials for viewing, or can be reproduced electronically (electro-holography) by re-creating the fringes in an electronically updatable light modulator. Electro-holography requires light modulators that control the amplitude and/or phase of light at a scale small enough to produce diffraction. Devices commonly used include liquid-crystal and MEMS light modulators, and acousto-optic devices.

2.4 Displaying Holograms

To reconstruct wavefronts from a hologram (either traditional photographically captured or computer-graphics holograms), the material containing fringe patterns is illuminated with some light source. Light that is modulated by the fringes contains the wavefronts that can then be viewed by an observer. Depending on the type of light modulator used, light illuminating the hologram may be passed through or reflected back from the fringe pattern, and may be viewed directly, or through some (for example magnifying) optical system. The type of illumination required depends on the properties and geometry of the light modulator used to display the holographic fringe pattern. Although only a two-dimensional fringe pattern is needed to reconstruct a hologram, these “thin” holograms contain no color information. Since the angle of diffraction by fringes is wavelength-dependent, thin holograms produce differently scaled images for each wavelength of illuminating light. These holograms appear most clear when illuminated by monochromatic or nearly monochromatic light. Holographic fringes can also be recorded in the volume of the light modulating material. In such a “thick” hologram, illuminating light along a straight path can encounter many reflecting surfaces, which can produce wavelength-selecting Bragg reflections. Thick holograms that can produce Bragg-reflections can contain color information, and can be illuminated under white light. Although electronically controlled volumetric (thick) light modulators can in principle be developed, most work on electro-holographic displays uses

light modulators that control the properties of a medium over only one or two dimensions.

2.5 Spatial frequency and view zone

Holograms require modulation of light on a scale much smaller than is needed with conventional displays. In general, an electro-holographic display that can correctly illuminate a view-zone with light of wavelength λ over a viewing angle θ requires a diffraction pattern with a maximum spatial frequency of $\sin(\theta)/\lambda$. For green light, a display that can be viewed over a range of 30 degrees requires fringes spaced about one micron apart, and thus pixels spaced a half-micron apart. For a meter-wide display, each row of the display would require two million pixels. This is about one thousand times the horizontal resolution of current High-Definition TV. Holographic displays that most completely reproduce natural scene appearance require a light modulator with similarly high resolution in both horizontal and vertical directions, and thus a square meter-sized display with a 30-degree horizontal and vertical view zone would have a pixel count on the order of 10^{12} pixels. Current prototype displays larger than a few centimeters generally create holograms by tiling or otherwise multiplexing smaller light modulators [Benton and Bove 2008, Chapter 21]. Some holographic displays reduce the information requirements by creating images that are holographic only in the horizontal direction. These Horizontal-parallax-only (HPO) displays contain one-dimensional holographic fringes on each display row, but have vertical resolution of only a few hundred or thousand lines. Taking advantage of the horizontal separation of viewer's eyes, these holograms can display correct (or nearly correct) images to both eyes simultaneously. HPO holograms can also maintain correct parallax under horizontal head motion. In exchange for this savings in information content, HPO displays sacrifice both vertical motion parallax and vertical defocus blur. One notable approach to lowering the information requirements for driving electro-holographic displays is to dramatically reduce the view zone to less than a degree (covering little more than the viewer's

eye). Such a display can use a light modulator of much lower density, but shows images that are only visible from a highly constrained portion of the space in front of the display. To compensate for all but trivial viewer head movement, displays of this kind have been designed that track the viewer's eye locations, and steer the output of the display to follow a viewer's two eyes [Häussler et al. 2008].

2.5.1 Depth Perception

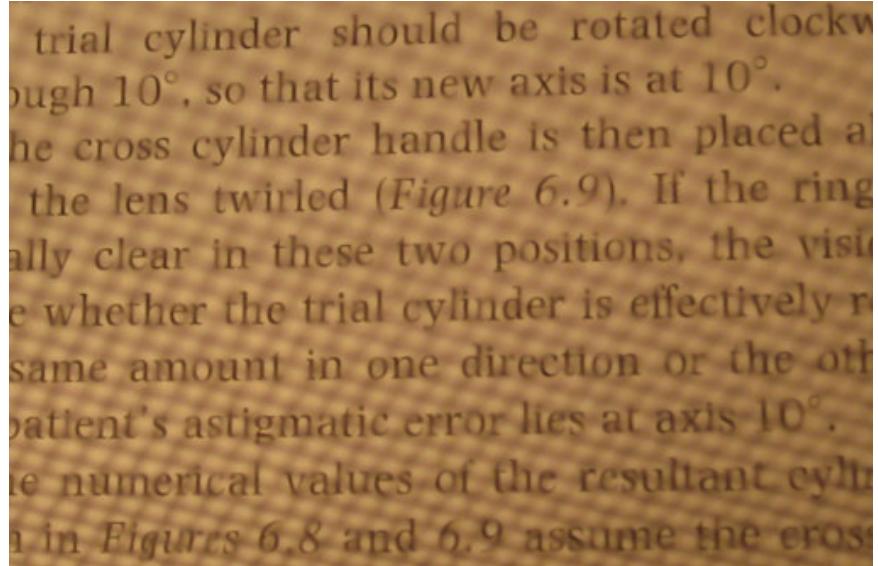
Human viewers estimate their distance from objects by combining information from visual and other sources. Sources of visual information that can reveal distance or relative distance are referred to as “depth cues.”

Pictorial information

Pictorial cues are aspects of a scene that allow a viewer to discern depth using only information that appears on the retina of a single, nonmoving eye. Pictorial cues are those used when we discern three-dimensional form depicted in two-dimensional photographs or illustrations. These cues include geometric cues, like the relative size and position of objects, blur gradients indicating deviations of objects from the plane of focus, color shifts due to atmospheric effects, and others [O’Shea and Govan 1997]. Holographic images can provide all pictorial cues displayed by current television and film display equipment, although some prototype holographic displays currently provide limited color fidelity, contrast, or detail.

Motion Information

Relative movement of objects within a scene can also be used to inform the viewer of scene geometry. In natural viewing, this motion can contain contributions from motion of



trial cylinder should be rotated clockwise through 10° , so that its new axis is at 10° . The cross cylinder handle is then placed and the lens twirled (*Figure 6.9*). If the ring is equally clear in these two positions, the viewer can determine whether the trial cylinder is effectively reading the same amount in one direction or the other. The patient's astigmatic error lies at axis 10° . The numerical values of the resultant cylinder shown in *Figures 6.8* and *6.9* assume the cross cylinder handle is at 0° .

Figure 2-1: Photograph of text through window screen. Defocus allows reading through the screen.

the viewer relative to the scene, as well as motion of objects in the scene relative to each other.

Relative object motion within a reproduced scene (resulting from either moving objects, or moving camera) can convey information of relative positions and sizes of parts of a scene. These relative motions reveal spatial relationships even when viewed on a traditional two-dimensional television screen. Holographic displays are equally capable of displaying this kind of relative motion.

In natural viewing, movement of the viewer's head can also provide information about scene geometry. Translational head movement allows viewers to discern distance to objects by observing the resulting induced relative motion of the objects viewed. Traditional two-dimensional and stereoscopic three-dimensional displays do not provide this cue to depth. When viewing a two-dimensional display, head motion can produce a distortion of the apparent geometry of the image of the entire scene [Arsenault and Ware 2004]. Motion also produces the appearance of rigid uniform motion of the entire scene, which provides information about the location of the screen surface, but not about the geometry

of the scene. Virtual Environment researchers have sought to overcome this limitation by electronically tracking motion of the head, and adjusting the image displayed accordingly. The addition of correct head-motion parallax to virtual environments has been found, for example, to improve time performance at a tapping task [Arsenault and Ware 2004] by 12%.

But display changes based on head motion are generally suitable for only a single viewer, and still introduce latency between head motion and screen update [Mania et al. 2004]. This latency has been shown to degrade performance [Watson et al. 2003], and research suggests that latency of as little as 8 to 17 milliseconds can be noticed by observers [Adelstein et al. 2003]. Full-parallax holographic displays can however present correct pictorial information simultaneously for all eye positions within the viewing zone. Holographic displays are for this reason a natural candidate for tasks that require frequent or rapid viewer motion, and also for tasks that benefit from simultaneous inspection by multiple viewers.

Holographic displays that restrict images to HPO have appearance that can change correctly and instantly for horizontal head motion, but sacrifice correct reproduction of vertical motion parallax. In addition, HPO holograms require assumptions about the location relative to the hologram of the line connecting the viewer's eyes [Benton and Bove 2008]. When viewed from locations increasingly distant from this line, objects in HPO holograms experience increasing amounts of distortion. Head-tracking approaches can adjust the geometry of computer-generated HPO holograms to match viewer head position, but will still introduce delay before correct images are shown under non-horizontal head motion.

Ocular information

Viewers also discern depth information from systems that control muscles in the eyes themselves.

Vergence

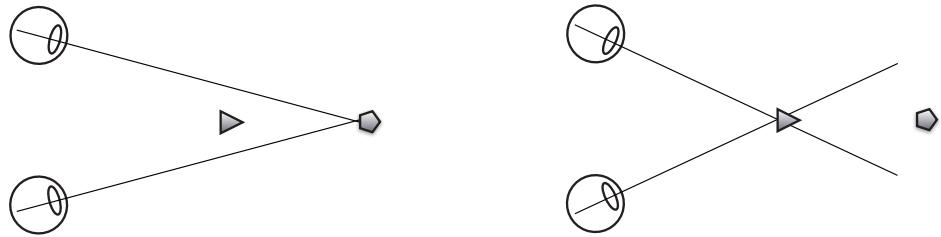


Figure 2-2: Vergence eye movements – eyes rotate inward to fuse near objects.

When viewing objects or displays, viewers make vergence eye movements that rotate the eyes to bring the images of points in the environment onto corresponding points on the two retinas [Collewijn and Erkelens 1990]. Distant objects are best viewed with the eyes both pointing nearly parallel to each other, while viewing close objects binocularly requires crossing the eyes inward. The vergence angle between the two eyes can be used to triangulate the absolute distance to an object. Binocular convergence is most useful as a depth cue for object distances of less than one meter [Palmer 1999]. For objects beyond a few meters, differences in the angle between the axes of the eyes becomes negligible.

Images on traditional two-dimensional displays produce matching images on the two retinas when the lines of sight of the eyes cross at the screen surface. Stereoscopic displays can present different images to each of the eyes, allowing the creation of image pairs that match when the eyes are verged at distances closer or farther than the screen. Vergence angle when fusing a pair of stereo images provides a cue to the absolute distance to the fused objects.

When viewing a holographic display, each of the two eyes captures a different view of the wavefronts reconstructed by the hologram. This allows each eye to receive different images as with a stereoscopic display.

Focus

Viewers actively control the refractive power of their eyes (accommodation) to keep retinal images in focus. The finite size of the pupil produces blur for objects that are not in the plane of focus. This blur can be undesirable when the eye is slow or unable to focus on an object of interest, but may also help with removing visual clutter for some tasks (see Figures 2-1 and 2-4). In much of natural viewing, the visual system adjusts focus quickly as we direct our gaze, and blur is not noticed. Although viewers tend to be poor at using the accommodative state of their eyes alone to accurately gauge distances [Mon-Williams and Tresilian 2000], accommodation can provide information when used in combination with other cues [Ellis and Menges 1997]; conversely, when accommodation is not correctly provided by a display the perceptual accuracy of other depth cues can be reduced [Hoffman et al. 2008].

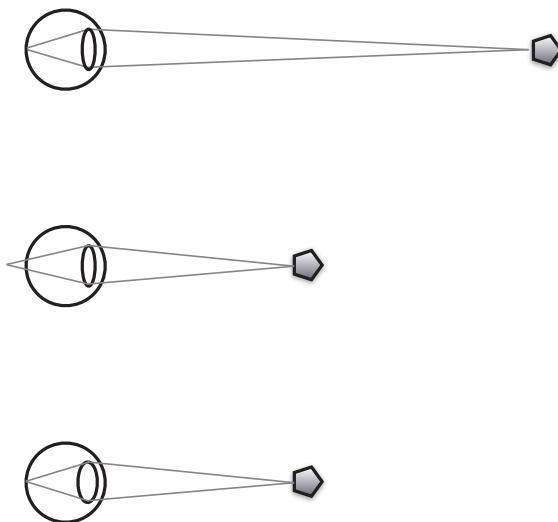


Figure 2-3: Accommodative eye movements – eyes adjust optical power to allow objects at different distances to appear in sharp focus.

Holographic displays, unlike other stereoscopic displays, can present wavefronts that ap-

pear in focus at distances in front of, as well as behind the display surface. Designs for HPO holographic displays can use a vertical diffusing screen to allow images to be viewed from a range of vertical positions [St-Hilaire 1994]. These vertical diffusers have a side effect of creating a single plane of vertical focus for light exiting the display. HPO displays can create image points with controllable horizontal focus, but vertical focus remains fixed at the diffuser. For these displays, scene points in the plane of the diffuser appear to focus and defocus naturally, but points increasingly in front of or behind the diffuser have increasingly astigmatic focus. Viewers have demonstrated the ability to mitigate some of the effects of astigmatic images, adjusting accommodation to compensate for differently oriented targets [Charman and Whitefoot 1978]. Studies evaluating the accommodative state of the eye while viewing images on an HPO holographic display show that viewers can use these images to drive an accommodative response [Takaki and Yokouchi 2012]. When viewing HPO holograms from 1100 mm, viewers were able to accommodate equivalently to real-object controls when presented with targets placed up to 450 mm (1.43 Diopters) in front of the hologram plane. Despite these results, the appearance of retinal blur for out-of-focus image points on an HPO display is different than the blur seen in natural scenes, and full-parallax displays. More research is needed to explore perceptual consequences of the scene-dependent astigmatism present in HPO displays.

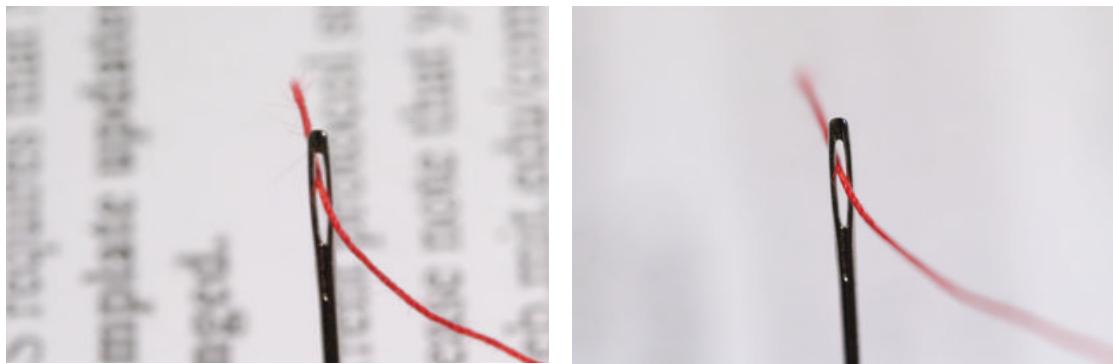


Figure 2-4: Eye of a needle with deep and shallow depth of field. Shallow depth of field may remove distractions when performing some tasks.

Holograms are most practically illuminated using one or more monochromatic light sources. This monochromatic illumination may have an impact on viewer's performance at focusing an image. Some research suggests that viewers use chromatic aberration of the eye to inform changes in accommodation, and that viewers accommodate more slowly when viewing objects under monochromatic illumination[Lee et al. 1999]. Depending on the choice of wavelengths however, a holographic display illuminated by multiple different monochromatic sources would likely allow viewers to once again exploit these aberrations to some extent. Existing displays that use fluorescent, laser, or LED illumination also frequently rely on combinations of narrow-band primary colors. Viewers' general ease at focusing images on these displays suggests that the narrow spectrum of holographic displays will not impose serious constraints.

Ocular Cue Interaction

Vergence and accommodation systems perform best when the cues they provide are in agreement. Stereoscopic displays that can only provide sharp focus at a single plane introduce conflicts between these cues. Since both vergence and accommodation are most active when viewing near scenes, correct presentation of these cues is especially important for images that appear close to the viewer. Although stereoscopic 3D is well suited for theatrical presentation, where images appear at long distances from the audience, home or mobile displays have more strict requirements for conflict-free viewing. Both the relatively small size of current holographic displays, and their ability to correctly reproduce both vergence and accommodation cues suggest that holographic displays will first find adoption for near viewing applications.

Stereopsis

The difference in scene appearance on the retinas of each of the two eyes also plays a role in the perception of depth. This phenomenon is exploited by current stereoscopic 3D dis-

plays, showing images captured from two laterally displaced viewpoints, one to each of the viewer's two eyes. Stereopsis takes advantage of the lateral displacements of corresponding points in retinal images of objects. Since eye movements can shift the retinas of the eyes relative to each other, retinal position (and stereopsis) can provide information about relative depth of objects, but not absolute depth [Palmer 1999]. The relative displacements between points in retinal images allow us to triangulate relative distances.

For stereopsis to function, the viewer's eyes must first move to place the image of some object at corresponding points on the two retinas. Parts of objects at similar depths to the fixated point can also be seen clearly, but appearance degrades for objects at distance that produce large disparity in retinal image. Producers of stereoscopic 3D content use knowledge of this phenomenon to predict which parts of a scene will be clearly visible when a viewer is verged on a particular subject. Careful planning of the position of the intended subject relative to the screen can provide more clear and comfortable viewing. Developers of content for HPO holographic displays may need to make similar considerations for managing the depth of the intended subject relative to the screen surface, as HPO displays impose these stereoscopic as well as additional astigmatism-related constraints on clear viewing.

In some wide-angle binocular stereo display configurations, low horizontal resolution of displays has restricted the presentation of stereo images to only two or three distinct planes in depth [Arsenault and Ware 2004]. Holographic displays, even HPO displays, require high spatial resolution in the horizontal direction, and therefore are unlikely to impose quantization on the stereopsis they can convey.

2.5.2 Image appearance

When designing holographic displays, aspects of vision relating to laser light and color vision may impose some constraints.

Although the use of coherent light allows the sharpest reproduction of holograms, the images created with coherent light are extremely sensitive to optical imperfections (in any part of the optical system, including the eye). In the presence of these imperfections, laser images appear to be modulated by a sharp, high-frequency speckled texture (“laser speckle”) [Goodman 1976]. The appearance of this speckle changes with movement of the eyes, and with constriction and dilation of the pupil. Conventional displays such as projectors that employ laser illumination can also suffer from similar speckle. Techniques for reducing speckle have been employed for these devices [Rawson et al. 1976] and are applicable to holographic displays as well. Depending on various characteristics of the holographic display, it may be possible to use incoherent light sources like light-emitting diodes instead of lasers.

2.6 Conclusions

Although holographic displays have potential for providing visual experiences that are far more natural than current “3D,” much work must be done in reducing the cost of creating and transmitting content to these displays. Careful consideration of human perceptual capabilities will allow holographic displays to come to market more quickly, allowing for immersive viewing experiences that provide both comfort and realism. Despite active research on perception relating to displays in general, very little work has been done to explore the viewing of holographic images [Barabas et al. 2010]. Holographic displays have unprecedented potential for reproducing many aspects of natural viewing, but also allow for creation of holograms simulating the constraints of other display systems. As such, holographic displays could be an invaluable tool for basic research in vision science. As more prototype displays become available, there are increasing opportunities for exploration of applications that take advantage of natural focus, occlusion and motion parallax. Telepresence and teleoperation research will likely benefit from experimentation with holographic

displays, as these applications are particularly well suited to the features of holographic images [Barabas et al. 2011].

Chapter 3

Computation for Holographic Television

3.1 Overview

One part of creating convincing imagery for television is reproduction of images that are interpreted by the viewer in a way that takes advantages of the viewer’s existing capabilities for comprehending visual images. All display technologies attempt to exploit some of the characteristics of natural images for conveying visual information. Of existing display technologies, holographic displays offer one of the most attractive sets of such characteristics. Holographic displays can reproduce scenes that respond more naturally to interrogation by a head and eye movement (including focus). But in order to take advantage of the features of holographic displays, images must be generated for display. The earliest, and many existing holograms were generated physically, using the interference of light from real physical objects to “optically compute” the fringe patterns needed for reconstructing scene images. Digital holography provides an opportunity for displaying holographic fringe patterns that originate not only from physical light interference, but also from computational simulations of light interference. Much as how computer graphics have enabled uncanny visual simulations of reality for applications like architecture, film and video games, syn-

thetic digital holography can create scenes that appear convincingly “real” over a range of viewpoints and viewing conditions.

This chapter describes methods for creating synthetic digital holographic fringe patterns that can be illuminated to reconstruct holographic images of scenes. Many methods for computing fringe patterns have been documented. Methods for computing the most accurate reproductions of physical reality tend to be computationally expensive. For this reason, much work in this area has focused on approximations that preserve many of the features of complete simulations while providing for faster computation. For visual telepresence applications, fringe patterns must be computed in real-time. This presents a substantial constraint on techniques that may be employed. This chapter seeks to survey existing methods that are applicable to realtime computation, and to describe scientific and engineering approaches to achieving hologram generation that is both sufficiently fast and sufficiently accurate for visual telepresence.

3.2 Optical Holography

When a traditional thin-film holographic plate is exposed, each 2D location on the plane of the hologram is exposed to light radiating simultaneously from many points in the scene. The light from these points can in turn be the result of interactions of light inter-reflected by other parts of the scene. To completely simulate this for up to n bounces off scene points before reaching the hologram, each location on the hologram must consider and aggregate P^n sources of radiation. For a square hologram of size l , computation would require $O(l^2 P^n)$ operations. By comparison, a ray-traced 2D computer graphics rendering requires $O(P^n)$ operations, and a simple scanline rendering approach requires only $O(P)$ operations. Most computer systems producing realtime graphics use approaches of roughly $O(P)$ complexity. Currently, complete realtime simulation of all but the simplest holograms is prohibitively costly.

One simplification that can be employed is to ignore light transport within the scene, and simulate only the light traveling from the scene to the hologram plane. For this type of simulation, we are computing the equivalent of a scene that was self-lit by coherent light, instead of one lit externally. Each point in the scene is treated as an emitter of light. The properties (direction, power density) of that light can be pre-computed from a simulated light source, or alternatively, simple holograms can be made by assuming that scene points emit uniformly.

3.2.1 Mark II Holographic Display

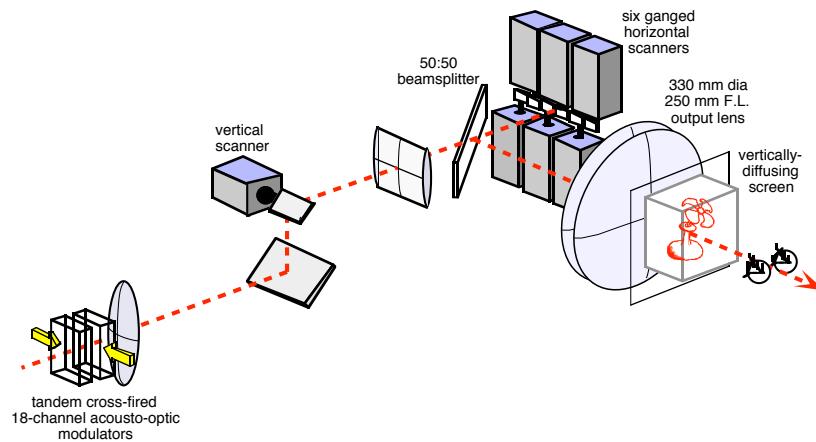


Figure 3-1: Simplified diagram of original configuration of Mark II Holographic Display [St-Hilaire 1994]

Optical Design

The Mark II holographic display [St-Hilaire 1994] was designed to be a reconfigurable platform for experiments in horizontal-parallax computational holography. The display

uses two cross-fired 18-channel tellurium dioxide acousto-optic modulators (AOMs), re-layed and de-scanned to project stacks of 144 fine-pitched nearly arbitrary 1-D patterns. In its current configuration, it creates 144 horizontal scanlines of up to 150 mm long, each with 218 samples (about 1747 samples per mm, with a window of about 1.2 mm coherently illuminated at a time using a HeNe laser). The image of the acousto-optic modulators is viewed on a vertical diffusing screen, allowing the 30-degree-wide image volume to be seen from a range of vertical positions. The horizontal optics of the display are designed to demagnify the image of the AOM by a factor of 10.3 to create diffraction angles wide enough to cover the 30-degree view zone. The display employs a horizontal Fourier-plane galvanometric scanner system to stabilize the moving hologram pattern traveling through the AOM. The horizontal optical system is configured to stabilize the AOM image while painting it across the width of each raster line. The vertical optical system performs modest demagnification, and scans the block of 18 optical channels in groups to paint out the 144 total raster lines. The Mark II display employs a vertical-only diffuser at the image plane of the AOM to allow viewing from vertical positions above or below the horizontal plane of the display center. Since the output of the display is vertically nearly collimated, without the diffuser, all scan lines would not be visible simultaneously. This diffuser sets the vertical plane of focus for points shown on the display, while the horizontal focus can be controlled by the holographic waveform imaged on the AOM.

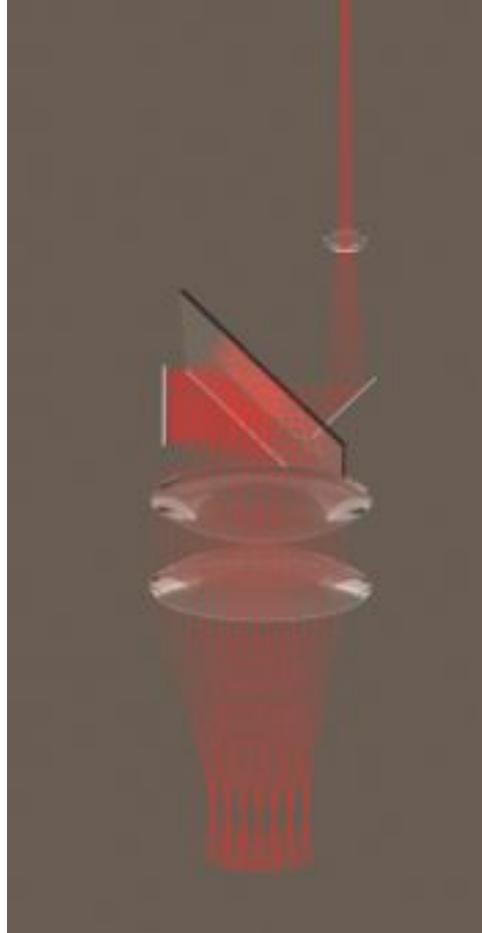


Figure 3-2: Rendering of horizontal optics for folded configuration of Mark II Display

Analog Electronics

The Mark II display has two sets of 18 acousto-optic channels for modulating light with holographic patterns. The acousto-optic modulators are driven by custom radio-frequency electronics built as a set of nine rack-mounted circuit boards. Each board contains inputs from two analog video signals, and has outputs connected to the transducers on four AOM channels. Only one block of 18 AOM channels is used at a time, on alternating horizontal scan passes (one block for each scan direction), so RF electronics are multiplexed so that each single input channel is used to drive two AOM channels on alternating scans. The RF electronics take the 50 MHz baseband RF signal generated as a VGA video signal,

modulate it with a 100 MHz carrier, and then filter it to select the lower side-band. The electronics then provide power amplification for the signals to drive the AOM transducers.

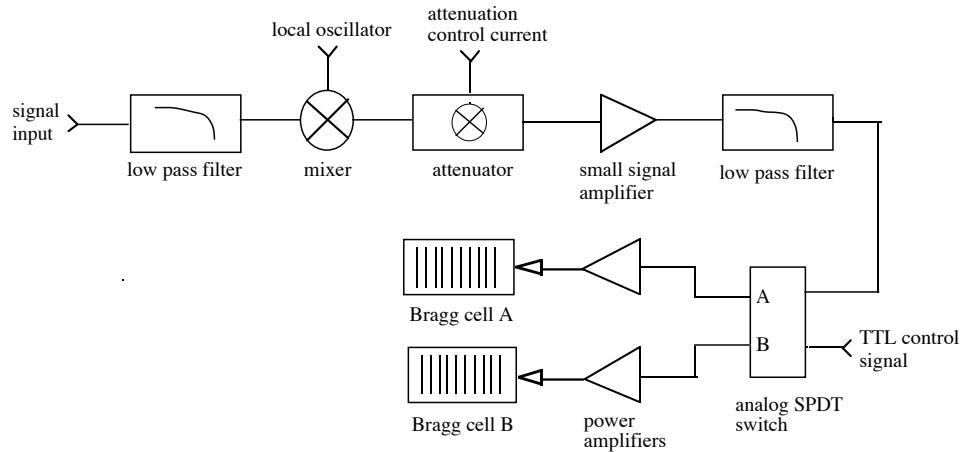


Figure 3-3: Radio Frequency Electronics for Mark II Holographic Display [St-Hilaire 1994]. Display contains 18 RF channels, each with the above design.

Digital Hologram Waveform Generation

Computed holographic fringe patterns are output to analog electronics by formatting them as a large image in the framebuffer of a PC graphics card, and then using the VGA video outputs of the card, converted to an analog signal [Quentmeyer 2004]. The final configuration of the Mark II display used three NVIDIA Quadro K5000 graphics cards, each connected to two DisplayPort to VGA adapters. Since most video cards are designed for analog displays with no more than a few thousand pixels per line, each hologram line is wrapped onto 128 individual video lines.



Figure 3-4: Eight holographic scanlines as packed into a single 2048 x 1760 video frame. Each green bar includes the samples for a single 2^{18} pixel hololine, wrapped onto 512 lines of 2048 pixels. Red boundaries are portions of the video signal that are forced to zero due to mandatory blanking intervals. Blue areas are horizontal and vertical blanking periods for the Mark II display that allow scanning mirror systems to prepare for the next line/frame.

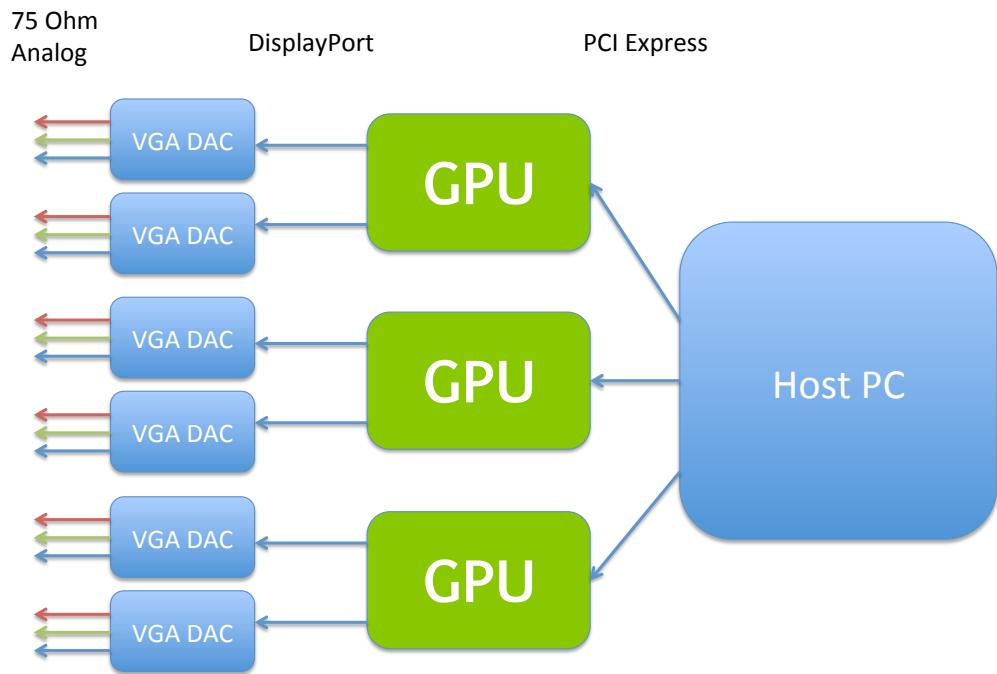


Figure 3-5: Hologram computation and digital synthesis hardware. Computation performed in part on Central Processing Unit of host PC, and part on Graphics Processing Units (GPUs).

3.2.2 Reconfigurable Image Projection (RIP) Holograms

My first contact with interactive digital holography was with the work from Steven Benton's group at the Media Lab. Michael Halle, Wendy Plesniak and others [2006] documented their Reconfigurable Image Projection approach in 2006. Referred to here as the RIP algorithm, this method was developed for rendering holograms in realtime, or near realtime on the Mark II holographic Display. The RIP algorithm takes as input a 3D model of a scene, renders a series of 2D projections of the radiance of the scene, and then constructs a hologram that projects a light field with directions and irradiances that could produce those images. Early implementations collected on the order of 100 separate parallax views of the scene, allowing for reconstruction of smooth parallax over the 30 degree view zone.

RIP Algorithm Summary

Computation of fringe patterns by the RIP algorithm starts with a simple 3D model of the scene to be rendered. Models for the implementations in current use comprise of a series of polygons defined in three dimensional space, with information about surface color (reflectance) stored per-polygon, or in 2D texture maps. Models also contain simple descriptions of lighting including location and intensity of omnidirectional scene lights. Models are first processed by using a standard OpenGL pipeline to compute scanline-rendered shaded 2D views of the scene. A single plane (usually at the center of the scene) is selected to be the reference plane (or "emitter plane") within the model. Views are carefully constructed using skewed frustum geometry such that each view describes light directed from the same 2D reference plane, but towards a different camera position. For holograms possessing parallax in the horizontal direction, virtual camera positions are chosen to lie in a single horizontal plane that runs through the center of the scene. This process is analogous to that used to optically photograph a set of parallax views of a scene for exposing an optical holographic stereogram[Benton and Bove 2008]. For the RIP algorithm, this set of

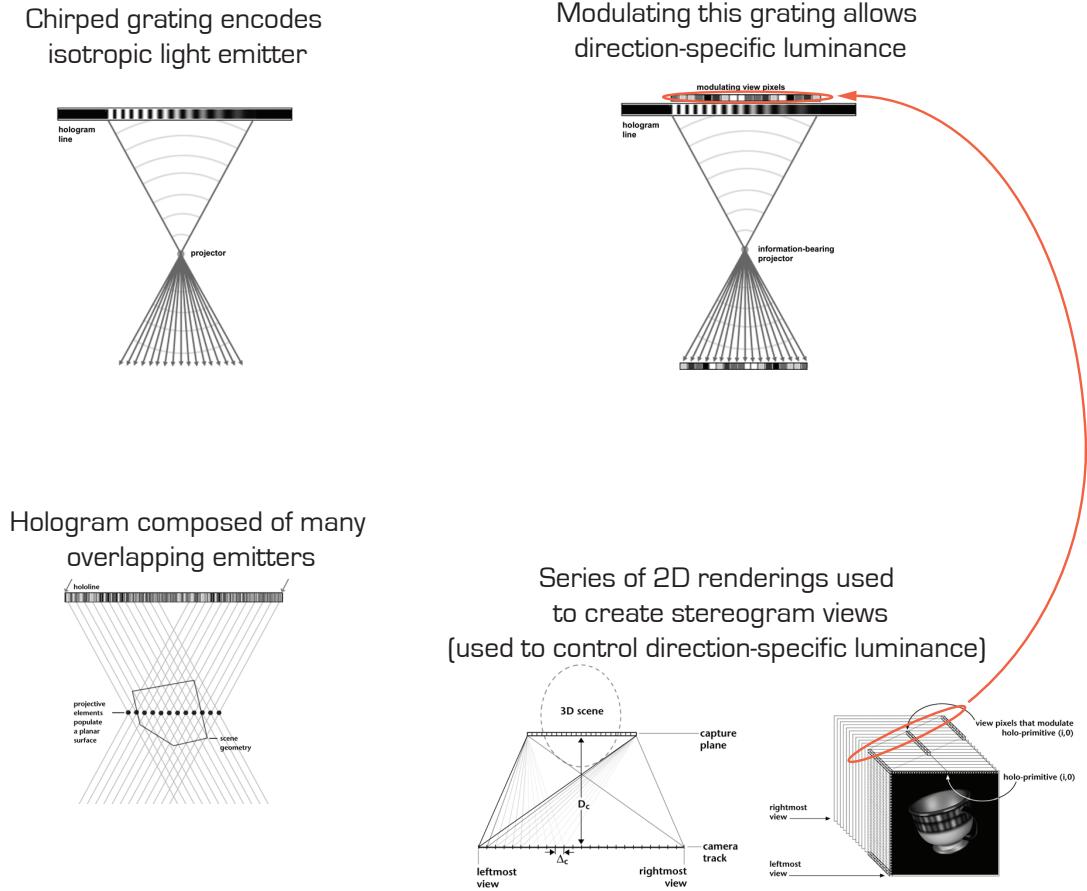


Figure 3-6: Geometry for computation of Reconfigurable Image Projection (RIP) stereograms. (After [Plesniak et al. 2006])

2D views can be used as a look-up table to find for any point on the reference plane the approximate scene radiance appearing to leave that point towards a given external viewpoint. (Alternatively, the radiance for any horizontal ray originating from that point on the reference plane.) Next, a holographic fringe pattern is constructed to approximate the pattern of light described in the lookup table. In the RIP algorithm, the emitter plane is populated with a grid discrete locations from which light will appear to radiate. A second parallel "hologram plane" is chosen to represent the light modulator used by the display. For a single point on the emitter plane, an interference simulation is performed for light falling on the hologram plane. In this simulation, the amplitude and phase of wavefronts originating

from the emitter point are superposed with parallel wavefronts striking the hologram plane from an oblique angle. This simulation provides the fringe pattern that would reconstruct the emitter point if illuminated from behind by an oblique plane wave.

Algorithm 1: Reconfigurable Image Projection Hologram

- 1 Pre-compute basis fringe B : the interference pattern of a plane wave and a point some fixed small distance from the hologram plane.
 - 2 Render 3-D scene from d discrete viewpoints, recording luminance only.
 - 3 Pack all 2-D luminance views into a single 3-D OpenGL texture.
 - 4 Divide the hologram plane into blocks (called "Holographic Emitters").
 - 5 **for** each *holographic emitter e* **do**
 - 6 Compute the line of texture coordinates corresponding to the luminance visible over all sampled angles for one point on the hologram plane.
 - 7 Place onto the output hologram a 1-D polygon, texture-mapped with the product of the above texture slice and the fringe B
-

3.3 Initial work

The first implementations of the RIP algorithm to be deployed [Quentmeyer 2004] on dedicated 3D graphics cards (GPUs) suffered from performance limitations. Early implementations computed changes in scene geometry on the computer's CPU, and transferred these data to the GPU each time a new frame was to be rendered. For each frame, at least 100 parallax views were required to create the light-field lookup table. Parallax views of the scene were rendered and lit by the GPU, but after each view, pixel data were transferred back to the computer's main memory. After all views were generated, they were processed by the CPU, before being returned to the GPU for fringe modulation. This original work was performed on NVIDIA Quadro FX 3000G GPUs, which employed the AGP bus for communication with the host processor. Transfers over this bus were identified as a bottleneck, resulting in practical frame rates on the order of one frame per second.

3.3.1 Modifications using Framebuffer Objects

Further work on improving implementations of the RIP algorithm allowed parallax view data to remain on the GPU and be used for fringe modulation without involvement of the host processor or host memory. This implementation employed the OpenGL Framebuffer Object extension to construct a 3D "texture map" in GPU memory for storing the light-field lookup table, and render parallax views directly to sequential slices of this memory. The 3D texture-map was applied to the fringe primitives by specifying only texture coordinates corresponding to the portion of the light-field sampled by each primitive. Modulation of fringe primitives was then performed by texture-mapping unit within the GPU, allowing for no intermediate manipulation of the light-field between creation and use. With this GPU-optimized approach and a subsequent upgrade to NVIDIA GTX260 GPUs, we were able to demonstrate for the first time RIP hologram computation that could be computed at the frame update rate of the display itself.

Shortcomings of RIP Approach

Although the current implementation of the RIP approach performs well for holographic emitters placed close to the hologram plane, computation can be slow when emitters are placed at greater distances. Distant points require larger basis fringes and in order to maintain image resolution, these larger basis fringes must maintain their spacing. Larger fringes with the same spacing will overlap to a greater degree, requiring each hologram pixel to contain contributions from many basis emitters. The evolution of graphics compute architectures away from the use of accumulation buffers has decreased the relative efficiency of computing images with accumulated fragments. The original RIP algorithm performs sequential rendering of individual emitters all additively blended into the same frame buffer. Recent GPU architectures are parallel, and rely on extensive caching for buffer reads and writes. Management of hazards possible from simultaneous writes from multiple process-

ing units slows the computation of accumulation and blending operations. The RIP algorithm could be further updated for modern GPUs by writing out individual modulated fringes separately into a temporary buffer, and then accumulating overlapping fringes in a second pass.

3.3.2 Diffraction-Specific Coherent Panaramograms

Although practical for reconstructing parallax views of scenes in realtime, RIP holograms as realized did not attempt to re-create the focus affordances of those scenes. When imaged with a camera or eye, scene points appear in sharp focus only when they lie in the emitter plane. Portions of the scene in front of or behind the emitter plane exhibit blur unless the viewer focuses at the emitter plane. Stereograms are unlike two-view stereoscopic images in that objects displayed can exhibit smooth motion parallax, but like stereo displays, images have different accommodation and vergence demands, which can degrade the viewing experience. To address these limitations, joint work with Quinn Smithwick and Dan Smalley, [Smithwick et al. 2010] proposed a new approach to realtime hologram computation. In this approach, scene radiance was again sampled from a series of view directions, but distance to the radiating points was recorded as well. These distance measurements were then used to simulate small segments of interference-modeled holograms that would reconstruct to points that could focus independently. With this approach, if the scene sampling of rendered views approaches the sampling of the hologram, it was hypothesized that this approach could converge to a full interference hologram computation.

Original Implementation

The First implementation of Diffraction-Specific Coherent Panoramagram rendering software applied the following algorithm, implemented in OpenGL with the NVIDIA CG

Toolkit:

Algorithm 2: Diffraction-Specific Coherent Panoramogram

- 1 Render 3D scene from d discrete viewing directions using a horizontally-orthographic but vertically perspective camera, recording luminance and depth information for each direction.
 - 2 Pack all luminance views into a single OpenGL texture.
 - 3 Pack all depth buffers into a second texture.
 - 4 Evaluate a fragment shader at each pixel of the output hologram.
-

Where the fragment shader performs the following:

Algorithm 3: Pixel shader for Diffraction-Specific Coherent Panoramogram

- 1 Divide the hologram plane into blocks (called "Wavefront Elements") of 512 samples each, and note the block b in which the current hologram sample lies within
 - 2 **for each viewing direction d do**
 - 3 Look up the luminance and depth information for the scene point that appears at the center of block b when viewed from direction d .
 - 4 Find the 3D location of the center of the current block translated by the depth distance found in the previous step, in the direction d .
 - 5 Compute the result of an interference simulation between a reference beam, and a point falling at the location found in the previous step.
 - 6 Output sum of the hologram contributions over over the d viewing directions.
-

Limitations and improvements

In its first incarnation, the Diffraction Specific Coherent Panoramagram (DHCP) renderer succeeded at generating, in realtime, hologram representations of large geometric models. After analysis of the holograms however, two shortcomings in the original implementation were discovered. First, scene points were accurately reproduced when, in the z-direction, they were close to the hologram plane. Parts of scenes however that appeared significantly in front of, or behind the hologram plane displayed visual artifacts. After viewing and analysis of the computed holograms, we discovered that the wavefront segments computed for distant scene points were all effectively planar. This resulted in the appearance of a clear, bright scene image being visible from discrete view directions (corresponding to the directions sampled in view rendering), but images appeared dark from intermediate viewing directions. Viewing such scenes while translating the head results in flickering as if the scene was being viewed through a picket fence. This limitation was due to the inability to recover exact x position of geometry from renderings of horizontally continuous surfaces. Since this information is unavailable, the initial implementation of the DHCP approach estimated scene point x-position to be the projection of the center of a view fragment (pixel).

Second, the number of view directions and scene sampling resolution chosen for realtime performance on modern GPU hardware proved too low to realize the benefit of the theoretical convergence to fresnel holograms when both view and scene are sampled finely. Although future optimized implementations will be able to increase this sampling, our current implementation can still display artifacts at the edges of objects due to undersampling.

Improvements were made to the DHCP renderer to recover more precise position information from scenes by allowing geometry information from initial renderings to pass through the rendering pipeline and be recovered during hologram computation. Though for point objects with well-defined position within their rendered pixel this allows more correct hologram rendering, this step incurs a small performance penalty. Additionally, most scenes are represented using polygons (even our point cloud renderings are rendered as small finite-

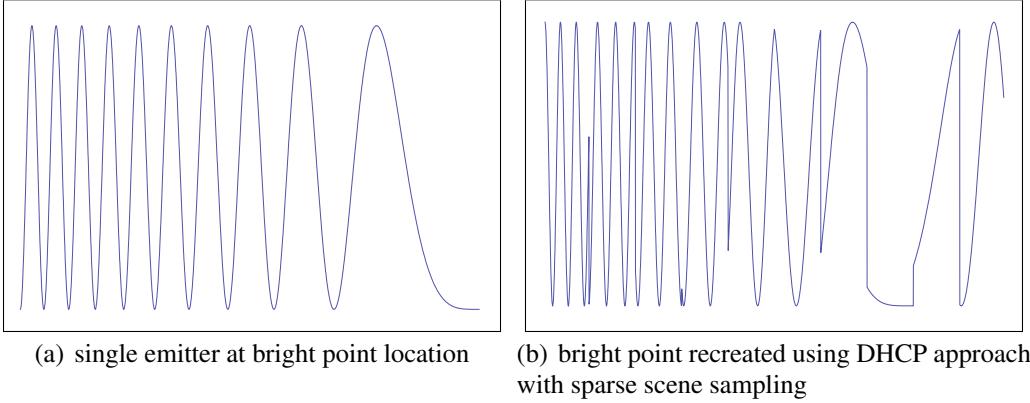


Figure 3-7: Artifacts in DHCP-rendered holograms when scene sampling is sparse.

sized sprites instead of infinitesimal points). For polygons that occupy finite area in view renderings, there is no simple way for multiple views to agree on the location of a scene point by inspecting the polygon rasterized on a given pixel. Each view sees some portion of the continuous polygon region, but agreement on a single representative point currently eludes this approach.

3.3.3 Interference-modelled Holograms

Other approaches described here create diffraction patterns that reproduce the image of a scene by discretely sampling lateral patches of scene appearance. For comparison, holograms were also produced that use scenes consisting of point emitters whose location is used directly without spatial sampling. (These point emitters can be placed at arbitrary locations in the scene, and emitter location is not quantized in hologram computation.)

A renderer was written to produce exact holograms of 3D point clouds, attempting to reconstruct as accurately as possible collections of luminous points. This renderer is not currently able to simulate scene points whose appearance changes from different viewpoints, nor is it currently able to simulate points that occlude other points. In this way, images created by this renderer have similar characteristics to images shown on a volumetric dis-

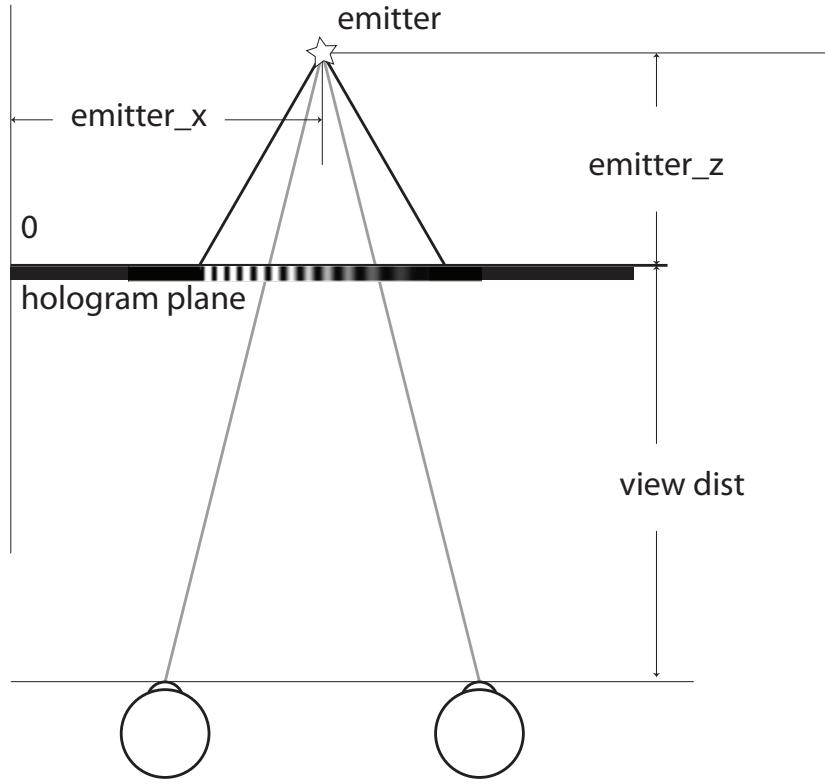


Figure 3-8: Geometry for computation of hologram of single light-emitting point

play. Images from this renderer have continuous parallax (images are consistent from any reconstructed viewing direction), and scene points come into focus at distances consistent with their parallax.

3.3.4 Two-view Holographic Auto-stereograms

In addition, I also wrote a renderer that generated two-view autostereoscopic images. The images computed by this renderer contain points that all appear in focus at a single plane, but scene points appear at different locations from leftward and rightward viewing angles. These holograms are computed using a modification of the interference modeling approach

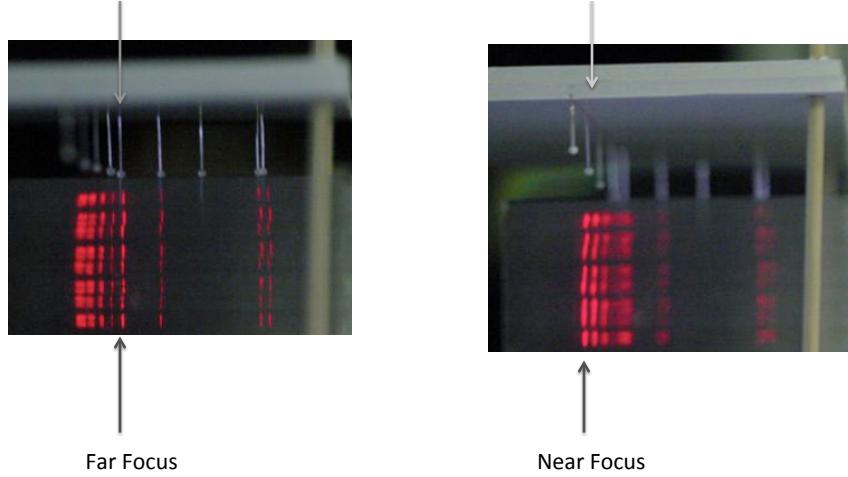


Figure 3-9: Example Holograms of vertical bars showing blur from defocus. Hologram is of a set of vertical lines arranged to match a set of straight pins (top of figures). Images shown are of same hologram with camera focused at far (left image) and near (right image) pins. The gaps in the vertical lines are due to defective RF channels. Horizontal misalignment of line segments is due to imperfect tuning of horizontal scanner control system. Line curvature is due to bow-distortion and field curvature from the design of the Mark II Display.

described above, but the effective x- and z-location of scene points is adjusted within the computation of each basis function such that each scene point results in two spatially separated half-chirps instead of a single continuous chirp (see Figure 3-10).

3.4 Software architecture

Hologram rendering software for point-cloud scenes were written to support multiple point-cloud formats. Point clouds can be obtained from the Kinect depth camera by way of several software libraries, and can also be loaded from other sources via files. Two open-source Kinect interface libraries, libfreenect and OpenNI each have relative strengths and weaknesses, so implementations for each were written. libfreenect was originally written without access to official specifications for the hardware. It is fast, and simple, but relies on hand-calibration for accurate depth and color registration. OpenNI includes contributions

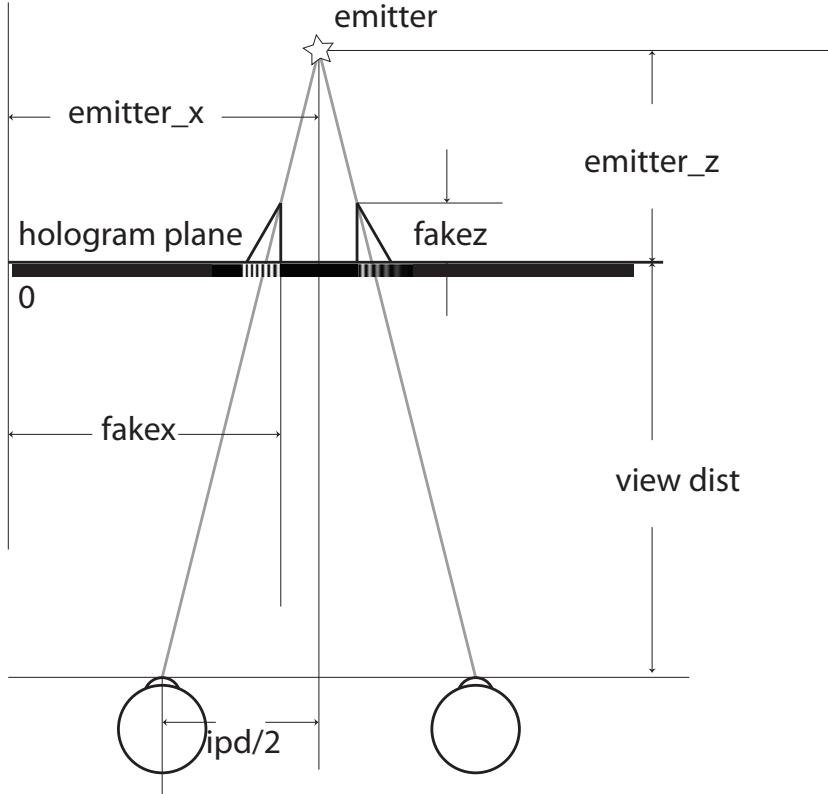


Figure 3-10: Two-view auto-stereogram for simulation of a stereoscopic display. Each scene point is displayed as two partial holographic emitters positioned just behind the hologram plane of the display. Each of the two emitters projects light towards a single eye.

from some of the developers of the Kinect hardware. OpenNI provides more accurate registration out of the box, but is complex and fragile. In addition, an Open Source software project, Point Cloud Library (PCL) provides access to registered XYZ-RGB point-clouds from the Kinect hardware (via OpenNI), allows for loading of arbitrary point clouds, and provides implementations of many computational geometry algorithms for processing these data. Point clouds used in viewing experiments experiments were pre-generated and written to files. These files were loaded during experiments using the Point Cloud Library.

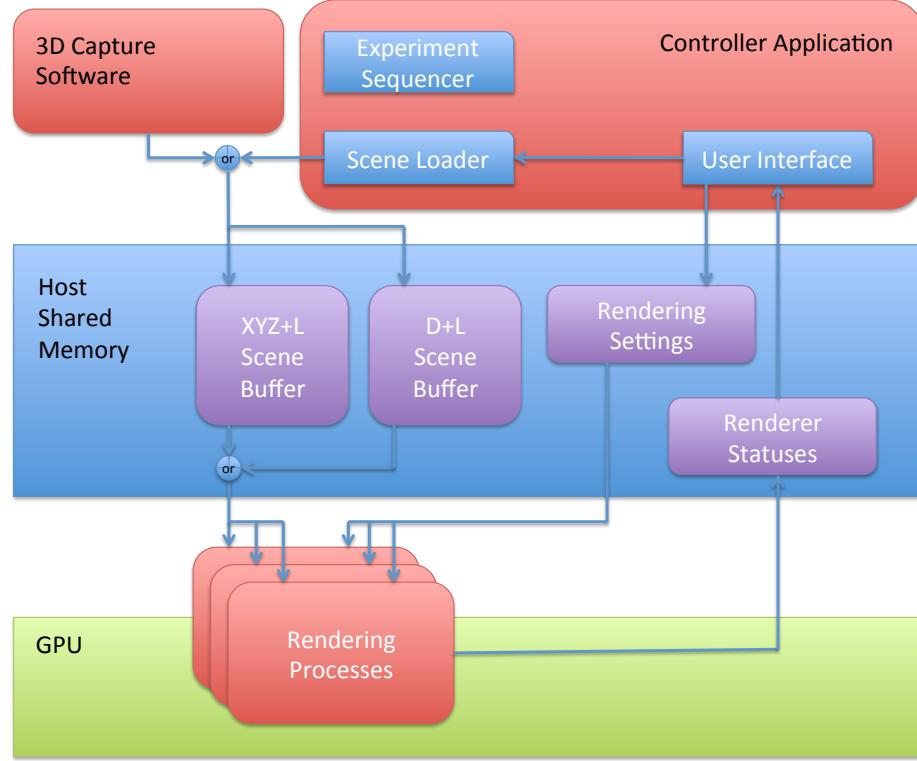


Figure 3-11: Software Block Diagram

3.4.1 Software Programs

The hologram computation software infrastructure has several modular components that have been designed to isolate and simplify the build process of each renderer and data acquisition implementation. Since these implementations use a number of large, sometimes mutually incompatible software libraries, keeping components separate makes individual modules easier to modify, upgrade and maintain.

Controller Application

An application with a graphical interface allows users to load pre-made 3D models for display, allows setup of live 3D scene streaming, and allows an experimenter to initiate and record responses from psychophysical experiments. The controller application has modules for reading point cloud files, loading them into shared memory for access by rendering programs, and for controlling the execution of the rendering programs. The graphical user interface provided by the controller application contains a series of sliders and input boxes for rotating, translating, and scaling a scene to be rendered. Selections made in this interface are written to a shared-memory buffer for reading by rendering processes. The controller application can also display statuses of the rendering processes.

Rendering programs

Rendering software takes a scene representation, along with settings describing desired viewing geometry, and computes the holographic fringe patterns for display. For the work described here, rendering using different approaches was performed using different rendering processes invoked by the controller application. For each rendering method, three instances of the rendering process are invoked, computing portions of the scene in parallel. Each rendering process creates an OpenGL context associated with a single graphics card (two displays) and outputs fringe patterns for scanlines driven by that graphics card.

RIP renderer: The RIP rendering processes load lit, triangle-based scenes from the filesystem and render those scenes according to viewing geometry read from the controller application via shared memory. Although it would not be difficult to implement, RIP renderers do not currently support streaming point-cloud input.

DHCP renderer: The DHCP rendering processes load point-cloud data from shared memory, and also load rendering settings (including scene translation and rotation) from separate shared memory. The DHCP renderer is written using OpenGL and Cg shaders.

Full interference and auto-stereogram renderers: The full interference renderer processes load point-clouds and settings from shared memory. This renderer computes hologram fringe patterns using a CUDA kernel, and writes results into an OpenGL texture. The OpenGL texture is then mapped to a fullscreen rectangle, rendering it to the framebuffer. Settings in the full-interference renderer allow the use of an alternate CUDA kernel that generates auto-stereogram fringes instead.

Live data sources

Two programs can be used for providing live scene data to the rendering processes. The first of these programs uses a Kinect camera directly attached to the hologram-rendering computer. This program initializes the Kinect camera, and reads raw depth and color information, placing it into shared memory for access by rendering processes. A second program was written to open a UDP network socket and listen for data sent by a remote computer. Data read from this network socket is copied to the depth and color buffers for rendering as with the first program. In this configuration, a third scene-streaming program was written to run on the remote computer. This scene streaming program reads frames a Kinect and streams them over the network to the second program.

3.4.2 Internal representations

Scene data are transferred to point-cloud renderers using either of two alternative formats. Data are written into shared memory by scene capture/generation modules to be further processed by renderers. Both formats include information about scene geometry, and monochromatic scene luminance.

The first scene format includes only disparity and luminance information, with 640x480 samples of each. Disparity and luminance are from separate sensors and must be transformed to create shaded 3D points. When data is transferred between modules in this

format, the image and depth data can be combined in the GPU as part of scene rendering. While this is ideal for live streaming from a Kinect sensor, it is not necessary for computer-generated or stored point clouds in XYZ format. Units of disparity are stored as floating-point representations of raw 11-bit integer disparity units. Luminance is gamma corrected and stored as 8-bits per pixel.

The second format used for transferring scene data stores lists of 3D points, pre-processed for fast rendering. Since individual scene points only appear on a single scanline, and rendering is divided by scanline across multiple GPUs, each point only needs to be transferred to a single GPU. This approach can be faster when the size of a scene sent and processed by a GPU is a performance bottleneck. For small scenes, it may be just as fast to transfer the entire scene to all graphics processors, and have the processors quickly cull points that are not needed for the scanlines being rendered. Scene pre-processing requires points to be transformed into screen coordinates. Following this transformation, it is trivial to find the scanline, if any, a point will appear within. For displays systems reproducing full-parallax holograms, each scene point may include support from multiple scanlines, so simple binning of points into scanlines is most useful for HPO holograms. Once the scanline for a scene point has been found, that point is inserted into a data structure that will serve as input to the hologram computation kernel for that scanline. In the simplest version of this approach, this data structure is a simple fixed-size array. Future versions of this project could use a geometric data structure to allow sorting of points for occlusion processing, and other optimizations.

3.4.3 Future work

Data structures, such as the interval tree [Cormen et al. 2001] could be useful for preparing scene information for hologram computation: Each scene point in an HPO hologram is "supported" by a finite range of locations in the hologram. This range of hologram

locations that may contribute to the appearance of a scene point is constrained by the maximum viewing angle of the hologram, and by the position of the scene point relative to the hologram plane. The interval tree data structure would allow scene points to be stored with the range of locations on the hologram where information about that scene point may be encoded. With this data structure, the scene points contributing to any given point on the hologram can be found in $O(\log n)$ time. Use of this approach either on the GPU, or in pre-processing on the CPU could allow for more points per scanline to be included in computation of realtime holograms.

Chapter 4

Vision Experiments

After developing a demonstration of a holographic television system, discussions of possible applications of such a system motivated a series of vision science experiments. I conducted these experiments to perform preliminary evaluation of the Mark II display for displaying 3D scenes for television. Further, I developed these experiments to demonstrate feasibility of evaluating holographic displays and systems that include these displays. This chapter describes methods for a series of short pilot experiments, and three multi-subject studies that suggest viewers can gain performance benefits when viewing images reconstructed from horizontal-parallax holograms. Some of these experiments compared the Mark II display to a Zspace Zdesk stereoscopic display, while others compared different image generation approaches that selectively remove some visual cues to depth from images displayed on Mark II. These methods are also described in the context of other experiments conducted with non-holographic 3D displays, and demonstrate feasibility of further vision science experiments employing holographic displays.

4.0.4 Introduction

Experiments described here seek to quantify the impact of one often-cited benefit of volumetric [Grossman and Balakrishnan 2006] and holographic content - the ability to present monocular depth information. 3D displays projecting scene points that focus at varying depths (so that a viewer can selectively bring some points into focus at the expense of others) provide this monocular information. Monocular depth information cannot necessarily be used by the observer to judge absolute, or even relative depth of objects, but does provide a cue to depth separation [Mon-Williams and Tresilian 2000]. These cues are present when viewing true 3D scenes, and are not available with traditional two-view stereo displays. This section will present an overview of display requirements for displaying monocular depth information and describes experiments for testing human and display performance in several viewing tasks. Though these experiments test viewing of very simple stimuli, I anticipate the results will be useful in informing other types of information and data visualization design, and for identifying applications where holographic displays improve task performance.

4.1 Pilot Experiments

I conducted several small-scale experiments to sample the large space of possible images and tasks that might be compatible with holographic displays. These experiments were used to identify interesting stimuli for deeper study. I was interested in 3D scenes that might be perceived differently when reproduced using different display systems. Other research [Johnston 1991] has suggested that scenes viewed on stereo displays can appear to be flatter when few pictorial cues are present. Though few experiments have been performed specifically with holographic displays, Akeley [2004] explored visual performance when viewing a novel blended multi-planar display. Akeley's findings suggested that timing visual performance could identify differences in holographic displays as well.

4.1.1 Motivation and Related Work

The proposed methods allow for testing of human performance at discriminating layered information using different display methods. The experiments compare display methods using emulated and traditional 3D displays to rendering methods that exploits the full capabilities of electro holography. By comparing performance using these two methods, the task-performance benefit of holographic displays can be critically evaluated.

Understanding of the impact of monocular cues on discerning details in 3D displays is useful in other aspects of information design. For example, the use of layered information is integral to geographic information system visualization where many pieces of information are available for a single location on a map. The US census dataset alone contains more layers of information than can be practically visualized simultaneously. By optically layering geographic information layers, further gains in data density over state-of-the-art visualizations [Skupin and Hagelman 2005] could be achieved. By understanding the practical limits in ability to discern layers, such displays could be further optimized.

Augmented reality applications also stand to benefit from understanding monocular cue presentation in 3D displays. In these applications, an electronic display adds information "on top of" the physical reality around the user. By placing this augmented reality in depth, these displays could make electronic annotations on physical objects easier to read. Since accommodative adjustments of the human visual system take time, labels or other virtual markers could be more quickly interpreted if they appeared in the same focal plane as the objects to which they relate. Additionally, augmentations that appear out of plane might be less likely to obscure other labels, or parts of the real scene seen through a display.

More generally, all 3D display applications that require close work (approximately at arms length or less), or seek to convey objects that appear close to the viewer, are likely to benefit most from display technology that can accurately portray monocular depth information, since these tasks are most likely to induce accommodation-vergence conflicts [Hoffman

et al. 2008] when traditional stereo display techniques are used. Related experiments have been documented for evaluating deep display technologies. Kim et. al [Kim et al. 2008] describe an evaluation of three 3D displays used in their lab. By photographing the light output of the display through a lens and pupil approximating that of a human observer's eye, they compare three displays' ability to provide monocular depth cues. As expected, they find no monocular cues in stereo images, but are able to show defocus-like blur in photographs from their Multi-Focus Display (MFD). Single directional views from the MFD do not exhibit blur when the simulated eye is refocused, but four adjacent views fall on the pupil of the camera, and the relative alignment of these views changes, creating discretely duplicated images that closely approximate the appearance of blur.

4.2 Stimulus design

Several pilot experiments explored tasks that were intended to be very difficult or impossible for subjects to complete without the use of stereopsis. Stimuli tested included dense luminous letters, as well as dot patterns.

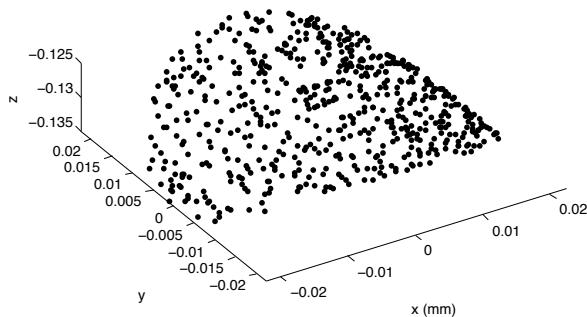


Figure 4-1: Example stimulus evaluated in one pilot experiment. In this pilot, subjects were to assess the sign of the curvature of the “chip.” Although this task is difficult without stereopsis, and easy with stereoscopic viewing, interactions with display aberrations precluded use of this stimulus in detailed studies with the Mark II display (See Appendices).

Luminous letters were useful for testing ability to discern depth order, but this stimulus type

relied on RIP holograms for real-time stimulus generation. Holograms of letters near the hologram plane could be computed in real-time (each letter consisted of closely-packed luminous points such that they appeared nearly contiguous). For deeper holograms however, the RIP approach employs blending of many overlapping primitives, which can become slow for large (deep) primitives. This limitation restricted dense-stimulus experiments to depths close to the surface of the display. For exploring deeper images, sparse stimuli were used.

Random dot stereograms consisting of collections of single monochromatic points [Julesz 1971] have been used in other work to attempt to isolate binocular 3D cues. When constructed carefully, stereograms can be devised to reveal very little information about their underlying 3D shape when viewed monocularly, or without correct binocular fusion. For stereograms to effectively conceal their underlying 3D shape, the view seen by each eye individually should appear as a random collection of points. That is to say that the view seen by each eye should contain as little evidence as possible to suggest that the points seen are not randomly distributed. Since sparse dot patterns can be quickly computed, and can be made to contain few monocular cues to depth, they present an exciting opportunity for study of perception of dynamic holograms.

Many designs for dot-pattern stimuli (including some used in other perceptual experiments [Gillam et al. 2007]) were explored . Some patterns were explored that could be useful for testing future displays, but interacted poorly with the horizontal aberrations present in the Mark II Display.

4.3 Displays

Vision experiments were conducted primarily using the Mark II Holographic Display (see Chapter 2). In addition, some experiments were also conducted using a second high-end

reference display. The reference display was a zSpace zDesk system, which provides a 24 inch 1080p monitor, along with a motion-tracked stylus, and passive polarized eyewear for head tracking and stereoscopic viewing. The zSpace system uses time-multiplexing to display full-resolution images to alternate eyes, refreshing images for each eye at 60Hz. This stereoscopic display provides high resolution, full-color images that provide binocular and head-motion parallax. When viewing this display however, the viewers eyes must be focused on the display surface for sharp images, even when images appear to be in front of, or behind the display.

4.4 Data Collection

I designed experiments using stimuli that where possible, could be performed without bias on subjects who were not naive to the purpose of the experiments. Particularly well suited to this purpose are experiments that explore the threshold of perception. In such experiments, a subject is asked a question about a stimulus that can be answered correctly if the stimulus is perceived as intended, but not when the stimulus is below a perceptual threshold. Thresholds can exist along many dimensions, including size, distance, luminance, contrast, duration, and separation.

Two-alternative forced choice (2AFC) experiments [Blackwell 1953] are widely used in perceptual experiments. In this experimental paradigm, a subject is presented with a series of stimuli, and for each, the subject selects one of two responses. Stimuli are chosen to be similar except for a single property of interest. Stimuli are presented across variations of this property, and responses are analyzed to measure the effect of this property on stimulus perception.

4.5 Data Analysis

For each experiment, responses to the set of stimuli were used to fit a three-parameter model of the subject’s perceptual performance. For experiments estimating a threshold, the model was of the probability that the subject could respond correctly for a given stimulus level. For these threshold experiments, the probability of correct response for below-threshold stimuli should be 50% (at chance), and should increase over some range of stimulus levels to approach 100%, (minus some stimulus-independent rate of error). Here, I will use the term Threshold to describe the stimulus level at which subjects respond correctly 75% of the time. I will use the term Slope to describe the rate of transition from chance guessing to reliably correct responses. Though generally low, base error rates were computed in performance models as well.

Although other functions can be used to model the psychometric response function [Wichmann and Hill 2001], for simplicity I chose the cumulative Gaussian. This model simulates an underlying threshold that is then offset on each measurement by the sum of many random offsets. For the task of perception of stereo images, these offsets could include lapses or increases in attentiveness, “lucky” or “unlucky” eye movements resulting from a subject’s prediction of where the stimulus might appear, or other complex occurrences that change the time to perform the task. With this model, the “Slope” parameter describes the standard deviation of the distribution of random offsets, and the “Threshold” value is the mean.

$$P(x) = 0.5 + 0.5 * \text{normcdf}(x, \mu, \sigma) \quad (4.1)$$

where `normcdf` is the cumulative normal distribution:

$$\text{normcdf}(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt \quad (4.2)$$

For threshold experiments, a nonlinear optimization method based on the Levenberg-Marquardt [Marquardt 1963] damped least-squares approach was employed to find parameters that fit the subjects' responses to individual stimuli. Confidence intervals on the parameters found by this approach were estimated using the sensitivity to the objective function to changes in the parameters near the computed minimum.

4.6 Time-To-Perceive Experiment

A larger experiment was conducted to explore subjects' ability to quickly perceive depth in briefly displayed images. This experiment sought to quantify differences in visual performance for objects shown at different depths, and with different display methods.

4.6.1 Methods for line-plane time-to-perceive experiment

General Methods

Subjects were positioned with their eyes 50 cm from from the screen (either diffuser for holographic display, or LCD surface for zSpace display). Subjects were given a chin rest and cushion, and instructed to maintain a stable head position, although head position was not rigidly constrained. Subjects were centered relative to the display by showing a scene containing a series of vertical bars receding in depth, and were asked to center themselves relative to the bars. This roughly placed the point between the subject's pupils at the center of the x-axis of the display. Subjects were then given a handheld keypad, and instructed to press keys indicating their best guess at the configuration of the stimulus presented. Stimuli were presented in sequences of 500 trials. In each trial, a cross was shown at the surface of the screen for one second. At the end of this second, the cross was replaced by a stimulus.

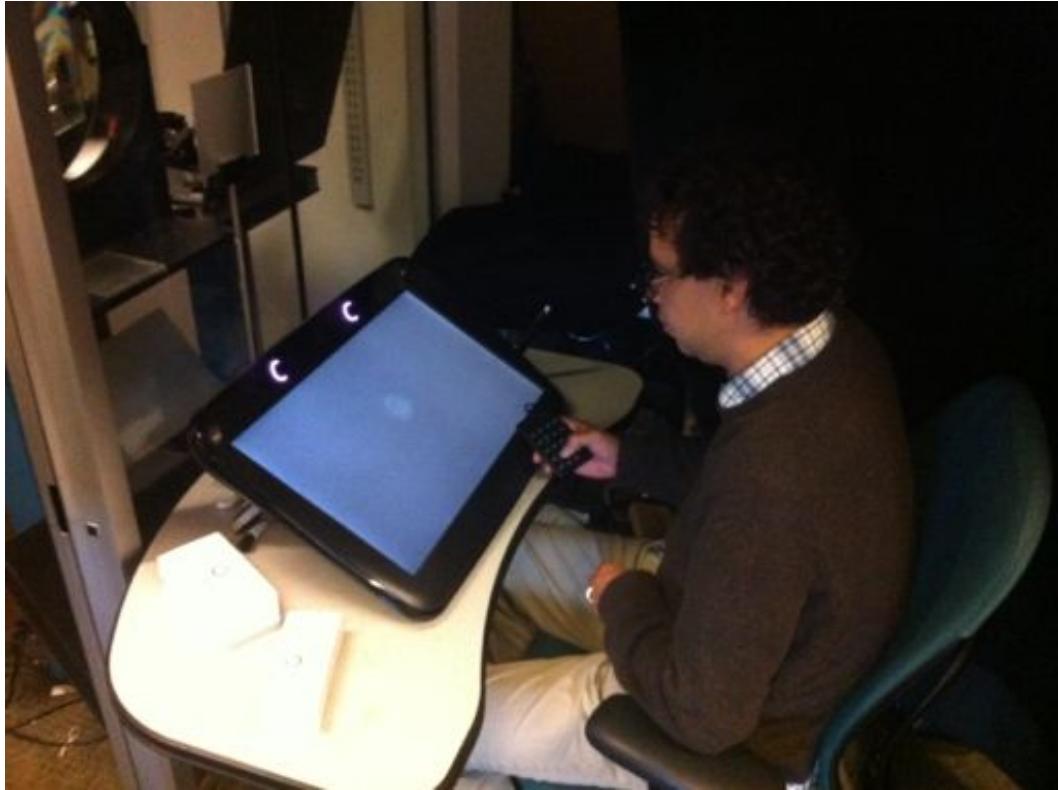


Figure 4-2: Subject performing viewing experiment shown using zSpace display. Mark II display also visible above zSpace display at subject's eye level. Similar subject positioning was used for stimuli shown on Mark II display.

The stimulus was visible for a pre-determined period of time, after which a blank screen was shown. Subjects were instructed that they could respond before or after the stimulus was removed. If subjects responded while the stimulus was still visible, the blank screen was shown immediately after the subject's key-press was detected. If the stimulus was shown for the intended time without detecting a key-press, a blank screen was shown until the subject responded. After the subject pressed a key, their response was recorded, and the next trial was initiated. Log files were generated recording subjects' responses, along with the timestamps (in milliseconds) of the start of the trial, the time that the stimulus was removed, and the time that the subject's response was detected.

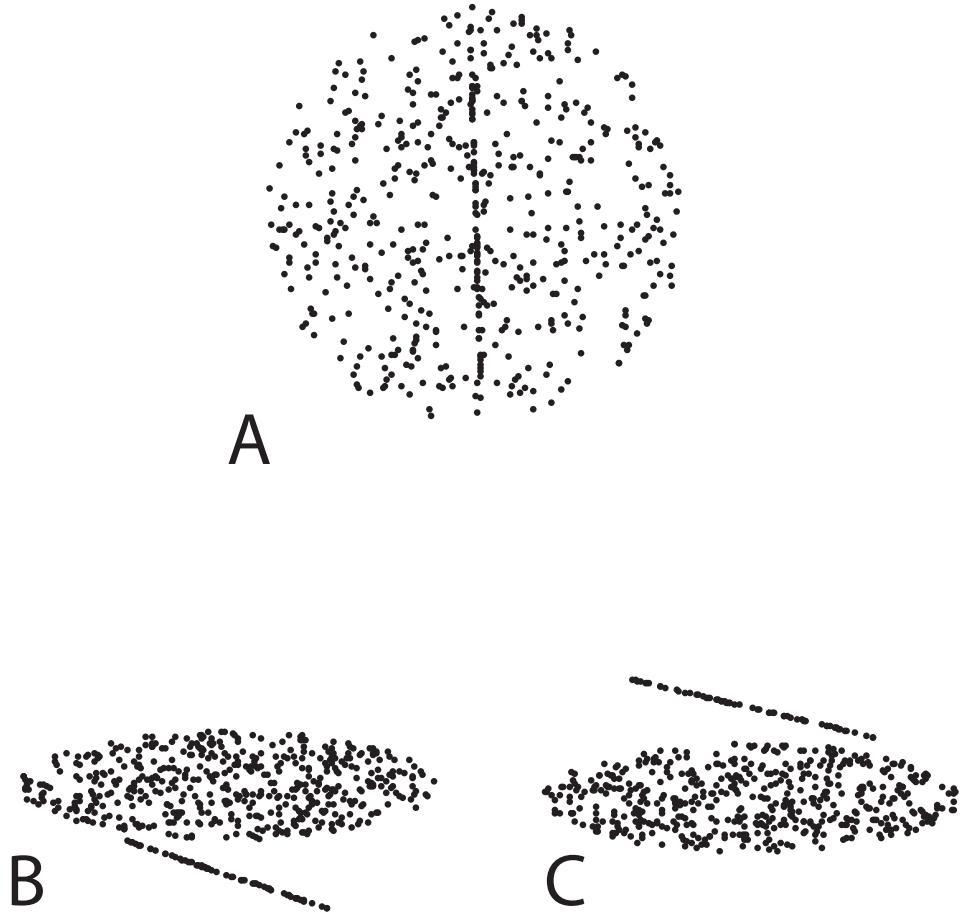


Figure 4-3: Stimuli for time-to-perceive experiments. A: front view of stimulus. B and C: Oblique views with line either in-front or behind plane.

Stimuli

Stimuli for this experiment were all combinations of a disk consisting of 500 points, and a vertical bar, both composed of 60 dots. For these experiments, the vertical bar was placed either two centimeters in front of, or two centimeters behind the disk. The disk-bar combination was displayed simultaneously for a series of durations, ranging from zero to one second. Position of the stimuli along the z-axis (into/out of the display) was also varied. For experiments comparing rendering methods, stimuli were shown at five different depths.

For experiments comparing holographic images to LCD images, stimuli were shown at two depths. All combinations of front/back position of bar, duration and depth were enumerated and shuffled, such that subjects viewed all combinations but in randomized order.

4.6.2 Results comparing performance viewing holographic and autostereoscopic images on the Mark II display:

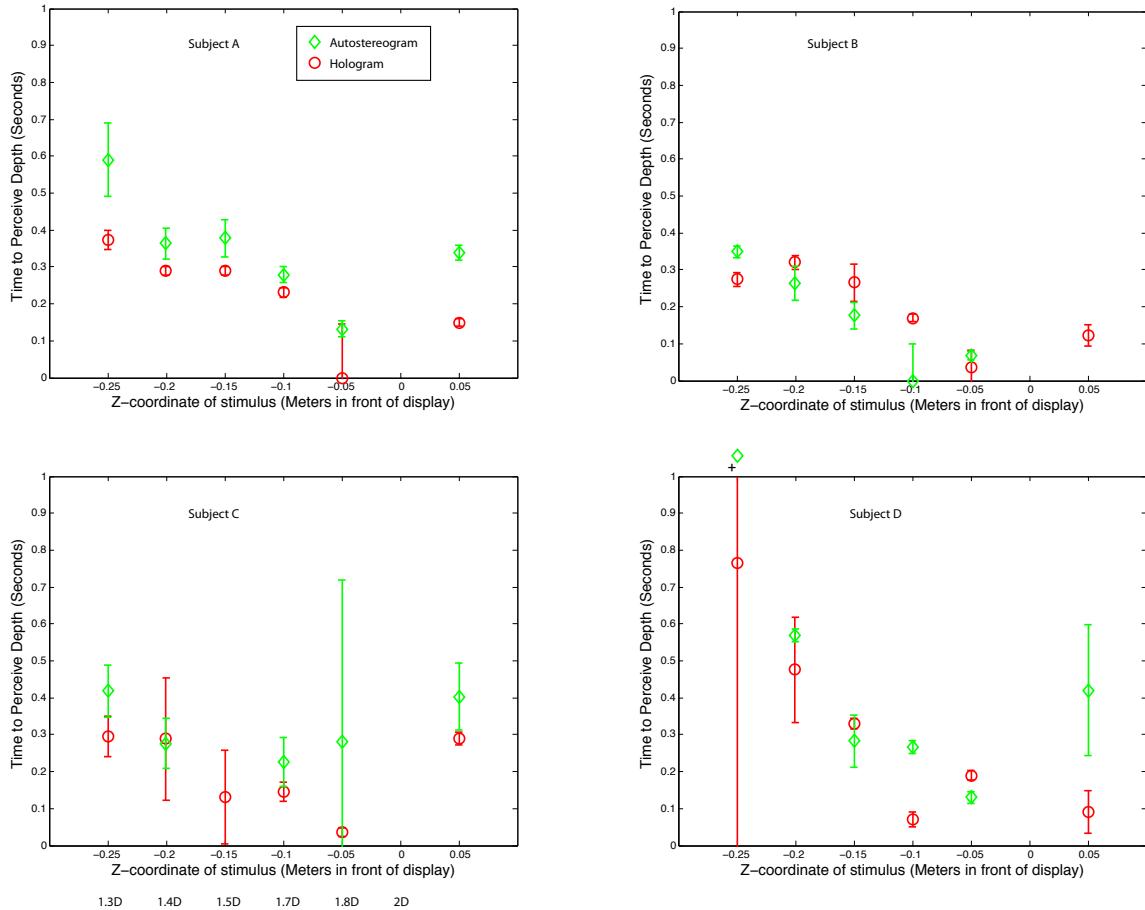


Figure 4-4: Results for four subjects in “time to perceive depth” experiment using Mark II Holographic Display. Experiment stimuli shown using point-based interference hologram renderer (red), and point-based auto-stereogram renderer (green). Subjects are able to perceive depth more quickly deep stimuli (leftmost condition in plots) when stimuli rendered as HPO holograms.

I conducted a multiple-subject experiment comparing stimuli presented holograms to stimuli shown as auto-stereograms. Auto-stereograms simulate the appearance of a traditional stereoscopic display, but are shown using the same display hardware as the hologram stimuli. This experiment design has the advantage of testing a potential advantage of holographic displays while controlling for effects due to defects in the Mark II display. Stimuli were the bar and disk images described above, with disk location ranging from 0.5 cm in front of the display to 2.5 cm behind the display. The bar presented was either 2 cm in front of, 2 cm behind the disk, and subjects responded with the perceived relative position of the bar. Of the six subjects tested, two subjects were not able to reliably complete the task (correct answers for both conditions were no better than chance for any stimulus location). Of the four remaining subjects, I observed a trend for all rendering methods of longer time to perceive for stimuli farther from (both in front and behind) the display surface. Especially of interest was a trend towards faster perception of holograms at the largest stimulus distance (2.5 cm behind the diffusing screen). More experiments to validate this effect will be helpful, but the small confidence intervals for fitted model parameters for most subjects at this distance suggest that this effect will be repeatable. Particularly promising is the results of subject B, where performance with auto-stereograms was faster for stimuli at the display surface, but holograms were still faster 2.5 cm behind the display.

4.6.3 Results comparing performance viewing on holographic and LCD displays:

An experiment was conducted comparing time to perceive holograms on the Mark II display with stereoscopic images shown on the zSpace display. For this experiment, stimulus presentations were brief, so subjects had little or no time to take advantage of the dynamic head-tracking capabilities of the zSpace display. For this experiments, subjects were positioned close to the center each display, and were asked to limit head movement. Data were collected for two subjects viewing holograms and stereoscopic presentations of the bar and

disk stimuli described above. Disk was shown at two depths: 30 cm behind and 2 cm in front of the display surface. The bar was either 2 cm in front of, or behind the disk. Two disk depths were used so that subjects could not preemptively adjust gaze to the location of the stimulus as it could appear at either one of two very different depths. Though intended as a distractor condition, results for this 0.5 cm condition, showed a trend towards improved performance using the zSpace display. For the 30 cm condition, subjects reliably performed more quickly when viewing images on the holographic display. This result is interesting as the Mark II display exhibits a number of artifacts (see Conclusions chapter) that appear to degrade stereoacuity for subjects using that display, but for a depth difference that is substantially above stereoacuity threshold, subjects were able to overcome that shortcoming and perform more quickly than with a traditional display.

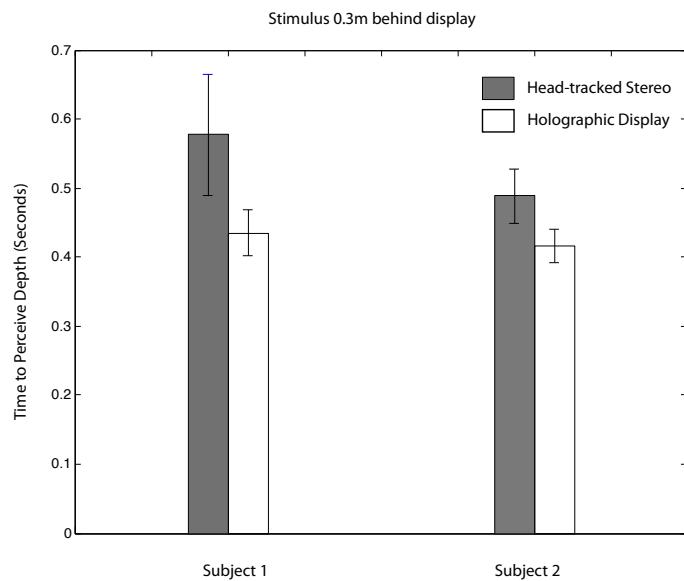


Figure 4-5: Results for two subjects in validation experiment comparing stimuli shown on holographic display to stereographic stimuli shown on head-tracked stereo zSpace display. Subjects are able to perceive depth of distant objects more quickly when viewing holograms.

4.7 Interactive pointing experiment



Figure 4-6: Subject performing an interactive pointing experiment using the Mark II display. Stylus is tracked by the zSpace system located on the table in front of the subject.

Teleoperation requires both perception of the conditions present in a remote scene, as well as action in response to those perceptions. Although the focus of this dissertation is the display and perception portion of telepresence, future evaluations of telepresence technologies could explore interactions between visual performance and motor tasks. Studies of prehension [Bingham et al. 2007] and teleoperation [Radi and Nitsch 2010] have suggested that live visual feedback, even without haptic feedback, can result in good performance. To demonstrate one such exploration, I conducted a simple experiment to measure subjects' performance at a 3D hand-eye coordination task. In this task, subjects were asked to interactively manipulate a cursor seen on a display so that it aligned with another object also seen in the display. For this task, no difference was found between holographic

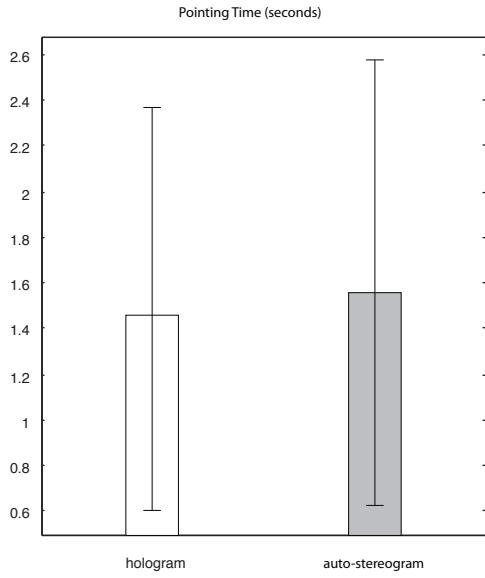


Figure 4-7: Results for one subject comparing free-hand pointing for targets displayed using holograms and using auto-stereograms. Difference in pointing time is not significant.

and stereoscopic representations of the scenes, but subjects performed significantly slower when viewing the scene as a flat (non-stereoscopic) projection. These results are consistent with improved performance found when viewing a volumetric display when compared to another head-tracked stereoscopic display [Grossman and Balakrishnan 2006].

4.7.1 Task

Subjects were positioned with their eyes 50 cm in front of the center of the diffusing screen of the Mark II Holographic Display. Subjects were given a cushion for stabilizing their head position, and were asked to avoid head movement. Subjects were then presented with a sequence of trials in which they were asked to use a hand-held stylus to adjust the position of an on-screen cross until it was at the center of an on-screen circle. 1 cm Target circles were presented at different distances along the z-axis of the display, approximately at eye level, centered on the display. Position of the tip of the stylus was tracked using an infrared camera system, and position of the cross was updated at 30 fps. Although orientation of

the stylus was also tracked, the cross was translated only, maintaining upright and face-on orientation. The cross image was generated to appear approximately 30 cm in front of the tip of the stylus so that for a range of target circle positions, the stylus could be manipulated within the space between the subject's face, and the screen surface. For each trial, a circle appeared on the screen at a pseudorandom depth, and subjects were asked to first position the cross within one target-circle-radius (in three dimensions) of the target center, and then to press a button on the stylus once they were confident that the markers were aligned.

4.7.2 Discussion

For this pointing task, I was not able to detect a difference in performance (either pointing accuracy, or pointing time) between holographic and stereoscopic presentations. Although this result could suggest that holographic displays provide no benefit for interactive tasks, it is also possible that the type of task also plays a role. For this experiment, pointing took more than one second on average. Since other experiments described here show a roughly 20 millisecond benefit for fusing holographic images of similar depth, the long time-scale and large intra-trial variation of the pointing task may be masking this benefit. Additionally, for interactive tasks, even roughly estimating target position could be enough to start a hand movement in the direction of the target, followed by more precise adjustment after the target has been fused. The time taken to physically move the hand to the area of the target may well be the limiting factor for this task. Future pointing experiments could be designed to explore this further. Exploration of pointing at targets in the presence of visual clutter could be informative, as pilot experiments have suggested that stimuli can be more difficult to fuse in the presence of distractor objects. Also, conditions could be tested where subjects must point without visual feedback, preventing the use of any 2-D cues for alignment of target and cursor.

4.8 Layer Discrimination Experiment

Images shown on displays may be sparse as explored in the experiments above, but may also be dense, continuous objects. This section describes a comparison of rendering methods using RIP holograms for holographic display to explore possible tradeoffs between rendering methods for spatial and non-spatial tasks. In this experiment, I evaluated human performance across different dense holographic rendering methods, by requiring the identification of overlapping letters separated in depth.

By performing a combined overlapping letter identification task and a depth discrimination task, this experiment emulates a situation likely to appear, for example, when visualizing layered information with text labels. For such visualizations to be useful, viewers should be able to read text on multiple layers, and will benefit further by being able to discriminate the depth order of layers. Understanding the critical depth separations at which identification performance improves will be useful in designing visualizations that seek to exploit overlapping information. Likewise, observing performance over a range of depths will inform designers of the practical limits for displaying text on an astigmatic display. By comparing performance at viewing stereograms and two-layer holograms, we hope to evaluate the potential benefit of monocular depth cues.

In the interest of exploiting focusing ability to read text, designs have been presented for employing multi-layer LCD screens for text editing [Masoodian et al. 2004]. Although no experiments examining performance at reading overlapping letters have been documented, Experiments have been performed examining subject's ability to read text when overlaid on textured backgrounds. Scharff and Ahumada [Scharff and Ahumada 2002] examined human performance at searching for specific words hidden in passages of text placed on different backgrounds, emulating the appearance of web pages that use background images. They found a decrease in performance when text contrast was low, and when backgrounds contained high spatial frequencies. In an unpublished study using a two-layer LCD display,

Wong et al [Wong et al. 2003] tested performance at a similar task while placing text and background at different layers. The authors suggest that performance at reading text over depth-separated backgrounds is different from when both are presented at the same depth. These studies, indicate that monocular cues to depth can increase viewer's ability to identify depth-overlapped letters on a screen and that the combination of monocular and binocular cues allows viewers to discriminate their relative depth ordering of these letters.

4.8.1 Methods

Display

For this experiment, Images were shown on the Mark II Holographic Display. The display was configured as described in Chapter 2.

Rendering

Computing the scanlines to display an arbitrary horizontal-parallax-only computer-graphics 3D scene can be computationally intensive. Full optical simulations of 3D scenes are currently impractical for realtime applications, so approximations are used to render holograms for the display. Approximate rendering methods currently in use [Plesniak et al. 2006] on the Mark II display subdivide the output of the display into discrete angular bins and for each bin, sample the luminance of the 3D scene using traditional computer graphics rendering techniques. The scene is then displayed as a cloud of discrete light-emitting points, whose brightness, is modulated in a discrete set of directions according to the 2D renderings. Each light emitter, or "hogel" is created by computing a hologram of a single point, and the holograms for points covering the desired scene are added together to form a single full-screen hologram. Most published photographs of Mark II display output

have placed these light emitters in a single plane, making the images more similar to high-density stereograms in that all light-emitting points appeared in focus in the same plane, which sits at the display's surface.

Scene appearance was previously demonstrated [Plesniak et al. 2006] for 383 emitters over 140 different viewing directions (spanning 30 degrees), which places emitters 0.19 mm apart, with each view spanning about 0.21 degrees. This geometry was chosen to place views about one pupil-width wide at a viewing distance of 0.6 meters. Even with this view spacing, the full fan of rays from a single view entering the viewer's pupil creates a strong monocular depth cue at the location of the emitter. This fan of rays is in focus only when the viewer is accommodating at the plane of the emitter, and will appear blurred otherwise. It may be possible for such a configuration to also provide other bundles of rays that come into focus at other distances. When rays from separate emitters whose view zones cross before entering the pupil, focusing the eye at the depth-plane of intersection causes these rays to overlap on the retina. The following experiment is used to comparatively evaluate the legibility of letters rendered as a single plane of emitters whose appearance changes with angle, as compared with a condition where light emitters are placed at two separate planes.

Experiment

In order to test subjects' performance at reading layered letters using the two rendering methods, subjects are asked to identify and compare the relative depth positions of two letters presented overlaid on each other, but separated in depth. To prevent subjects from using occlusion as a cue to relative depth, the two letters were presented as transparent, self-illuminated surfaces. Each was visible "through" the other irrespective of their depth order.

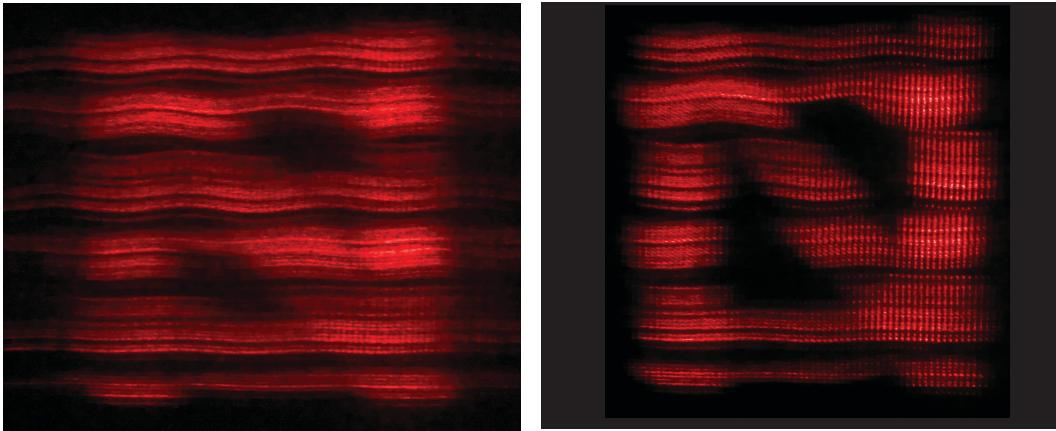


Figure 4-8: Centermost views of example stimuli. Left image (N overlaid on S) generated using layered planar holograms. Right image (N overlaid on D) by stereogram rendering.

Stimuli Letters for the task are rendered using two methods, using capitals from the SLOAN typeface [Pelli et al. 1988], approximately 70 mm tall, displayed as luminous characters on a dark background. The first method uses the Reconfigurable Image Projection algorithm [Plesniak et al. 2006] to render a single set of 64 parallax views of the two overlaid transparent letters. The views are then used to modulate a single plane of emitters that lies 4 mm in front of the display’s diffuser plane. This method will be referred to as "Stereogram" rendering. The second rendering method to be evaluated creates two layers of isotropic emitters, each in the shape of one letter. For this "Two-Layer" method, the image is effectively that of two planar holograms added together. For both rendering methods, emitters are placed on a grid, 128 emitters wide, with 144 lines of emitters.

Task Subjects are seated in front of the Mark II Holographic display, in a chair adjusted to place their eyes 600 mm from the diffuser screen. Subjects are presented with a series of letter pairs and are asked to discern the letters’ depth order. Letter pairs consist of an N 75 mm behind the diffuser plane, and a second letter placed either in front or behind the first. Subjects respond with key presses, first to discriminate whether the non-N letter is in front, or behind the N, and second to identify the second letter presented. One hundred letter pairs, ranging in depth separation from -75 to 75 mm, and rendered with the two methods

(Stereogram and Two-Layer), are presented in random order. After viewing a dark screen, a pair of letters appears, and remains visible for 750 ms. and then the screen goes blank. Subjects are instructed to respond to the depth order of the non-N letter by using the 8 and 2 keys on the numeric keypad (8 indicating behind and 2 indicating in front). Following this, subjects are asked to press the key on the keyboard corresponding to the non-N letter that appeared. Experimental software controls timing and display, and records responses, and response latencies.

Results The experiment described above was performed on three subjects. For both rendering methods, performance at judging relative depth of the two letters was poorest with small depth separations. Errors in letter recognition were more uniform across the depth range. Performance in depth perception was significantly better with the Two-Layer rendering method (81% correct) compared to the Stereogram method (61% correct) (Wilcoxon signed-rank test: $z=-2.832$, $p=0.005$). Performance in letter recognition was not significantly better with the Two-Layer rendering method (88% correct) compared to rendering method #1 (83% correct) (Wilcoxon signed-rank test: $z=-1.219$, $p=0.223$).

4.8.2 Discussion

Subjective evaluations showed that subjects had little or no ability to independently bring into focus layered presentations rendered using the Stereogram method although this method presents dense-parallax views conveying strong binocular depth cues. Combined with significantly higher error rates for depth perception, this method appears to be less suited for displaying layered information than the Two-Layer method. When viewing images rendered by the Two-Layer rendering method, some subjects suggested that they could focus independently on discrete depth layers although further evaluation using an optometer could verify this. Although the Two-Layer method supported better depth discrimination performance, the stereogram method may still be preferable for rendering small details, as it

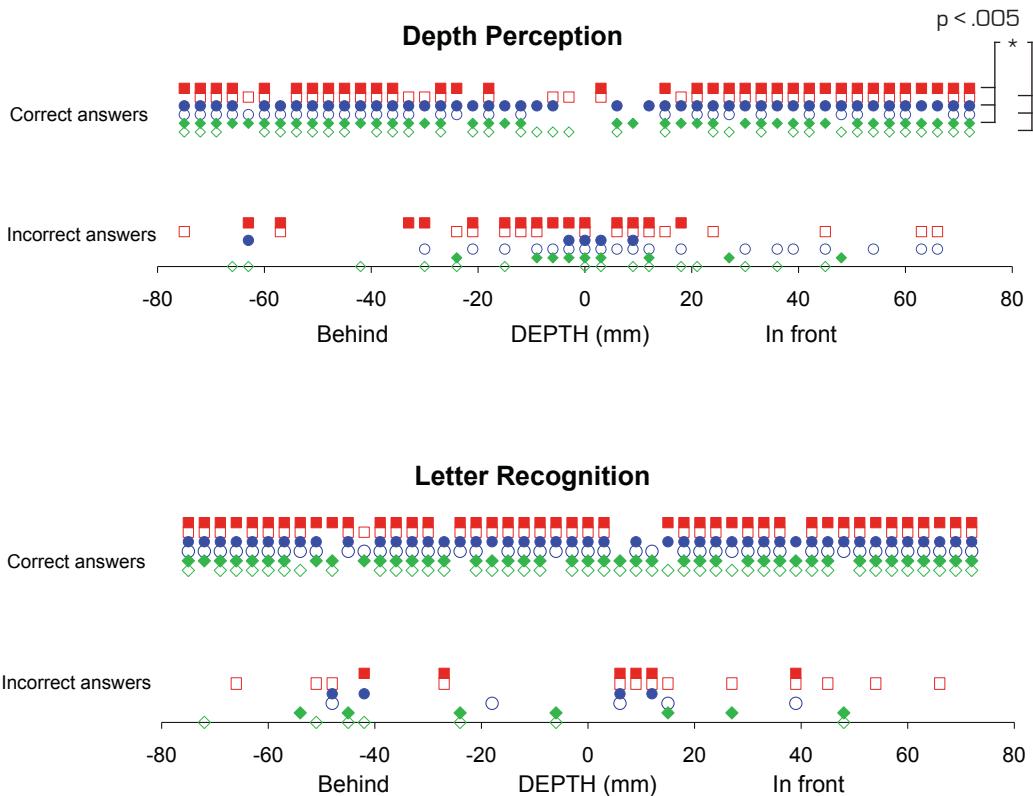


Figure 4-9: Results of depth discrimination and letter recognition for three subjects. Filled symbols are for Two-Layer rendering method, open symbols for Stereogram method.

allows objects at all depths to be simultaneously in-focus. Since the Mark II holographic display is only capable of displaying horizontal parallax, larger depth separations result in more difference between horizontal and vertical focus, and may impede letter identification. Additional experiments using larger depth separations could investigate this. Future work could also include experiments with other rendering methods including the phase-added stereogram [Yamaguchi et al. 1993] or wavefront element rendering [Smithwick et al. 2010].

Chapter 5

Conclusion

5.1 Contributions

Experiments and Methodology

The work documented here demonstrates the feasibility of several methods for assessing the performance of display systems in communicating spatial information to viewers. Results, while gathered using a small number of subjects, are suggestive of a performance benefits from even the most crude holographic images. Specifically, while no consistent difference in performance was observed for viewing images near the surface of a display, holographic images did allow faster perception of depth information when representing objects deep within a display.

Real-Time Holographic Rendering Architecture

This work also contributes designs and implementations for rendering holograms for real-time display. Several approaches are described which can be selected for different tasks.

RIP holograms can be generated from 3D models or sets of 2D view images. I have contributed GPU optimizations to this rendering approach allowing for real-time hologram computation and display. For scenes that appear far within or in front of a display, I have contributed a CUDA accelerated Fresnel hologram rendering approach that displays sparse scenes, but with accurate horizontal focus information. Further, I have contributed an architecture and reference implementation for capture, transmission and rendering of real-world 3D scenes reproduced remotely as holograms. To my knowledge, this is the first demonstration of real-time hologram display of its size where images were computed from a 3D camera.

Evaluation of Mark II design for quantitative experiments

In using the Mark II holographic display for the research described here, a few shortcomings have become apparent. Some of these are inherent to the original design of the display system, and some are limitations of our current implementation of that design.

The horizontal scanning system of the Mark II display introduces several artifacts that limit the quality of images displayed. The horizontal scanning system is required for both stabilizing the moving holograms present in acousto-optic light modulators, and for spatially multiplexing the small light modulators to create a large image and view zone. (Alternative approaches to scanning include tiling [Hahn et al. 2008], which while expensive, may prove more suited to perception research.) The scanning system is composed of six smaller scanning mirrors that are aligned and synchronized with each other. At the same time, these mirrors must be spaced slightly so that they do not contact each other while moving. The static and dynamic alignment of the mirrors is fragile, and any space between, or angular misalignment of mirrors results in gaps between angular sections of the view volume. During display operation, a servo controller actuates the mirrors to move at a high angular rate. In the Mark II display, the servo controllers are aging analog systems, and were difficult to tune reliably to produce synchronized actuation of all mirrors. Difficulties encountered

included thermal instability of controllers and glitching due to degraded components (capacitors). Scanning defects were not noticed by most participants of the vision experiments conducted on the Mark II display as viewing geometry was selected to mitigate these problems. When viewing the display from a distance, both eyes may sit inside the view cone of a single scanning mirror, so misalignment or gaps can be invisible except when moving the head laterally. When viewing from close to the display however, each eye can see contributions of rays that come from a larger range of angles. Closer than about 40cm, the eye always sees simultaneous contributions from multiple scanning mirrors. For these reasons, I was not able to achieve an alignment that produced continuous images for viewers close to the display.

The vertical scanning system is driven using a stair-step waveform that causes the vertical scanning mirror to continue to vibrate long after the servo has stopped moving. This vibration introduces some minor horizontal optical aberrations. These aberrations are somewhat visible at the horizontal Fourier plane, but are not immediately apparent in the image. This suggests that the images shown on this display have some time-varying angular error, which would have little impact on images displayed at the hologram plane, but would impair the sharpness of images placed in front of, or behind the hologram plane.

5.2 Closing Thoughts

Holographic displays have unsurpassed flexibility for conveying visual information. Currently, these displays are complex, fragile research instruments. Despite this, I have been able to conduct a few controlled experiments that begin to demonstrate the potential for these displays. As these displays mature, we will have the opportunity to experiment with a broader range of applications, including applications that benefit from full-color and from larger displays. The methods documented here however have been designed to work with both fully functional, but also flawed research displays. It is my hope that as further re-

search is conducted with holographic and other new displays, researchers will build upon these methods and learn even more about human vision, and techniques for information display.

Appendix A

Illustration of Distortion from Mark II Horizontal Optical Design

Although the Mark II holographic display can successfully reconstruct holograms that are viewable over a large range of head positions, the optical design of the horizontal scanning system introduces distortion. For image points deep within the display, this distortion is minimal. The distortion becomes more pronounced with distance from the output lens. Distortion is small along the axis of the display, but becomes severe for points at the extremes of the horizontal scanning sweep.

Table A.1: Magnitude of displacement in Z direction for edges of hologram planes

Plane Center z-location (mm)	Unintended displacement of plane edge along z axis (mm)
-189.7	8.5
-94.8	10.7
0	28.5
94.0	60.7
186.9	86.5
369.3	171.8

An illustration was generated by ray-tracing through a model of the horizontal optical sys-

tem. Rays were traced for seven de-scan mirror positions spanning the horizontal raster scan, and for object points at a series of planes in depth. Planes were simulated for holograms reconstructing 1m, 2m and 4m in front of the AOM plane, at the AOM plane, and 1m and 2m behind the AOM plane. With a demagnification of about 10.3, the ideal reconstructions of these planes would fall 97cm apart (with the closest plane separated by 194cm). Actual reconstructions of centers of planes are more closely spaced for images placed far from the viewer, and distances are increased for points near the viewer. Future hologram computation systems could modify holograms to compensate for this distortion.

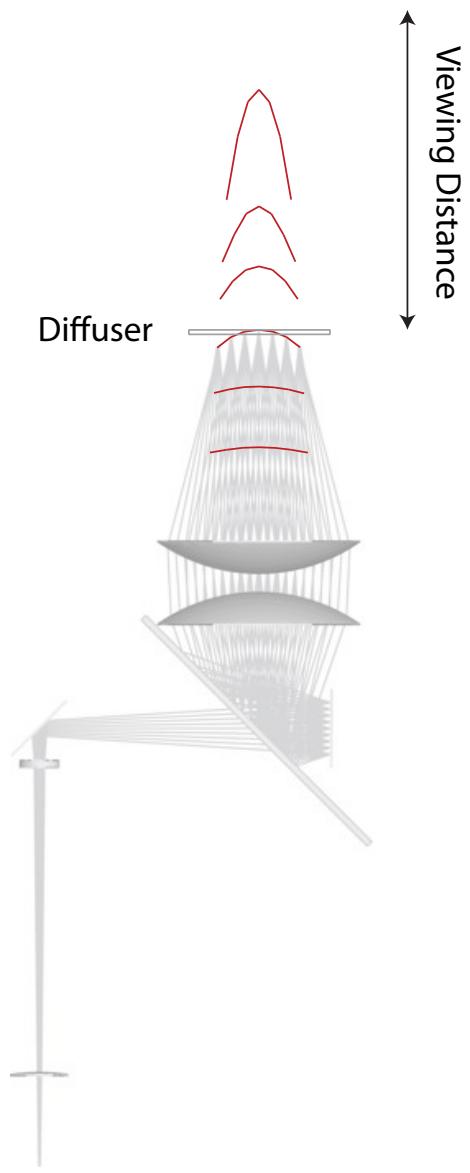


Figure A-1: Contours of reconstructions of planes after magnification by the optical system of the Mark II Display

Bibliography

- ADELSON, E. H., AND BERGEN, J. R. 1991. The plenoptic function and the elements of early vision. *M. Landy and J. A. Movshon, (eds) Computational Models of Visual Processing.*
- ADELSTEIN, B., LEE, T., AND ELLIS, S. 2003. Head tracking latency in virtual environments: psychophysics and a model. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 47, SAGE Publications, 2083–2087.
- AKELEY, K. 2004. *Achieving near-correct focus cues using multiple image planes*. PhD thesis, Stanford University.
- ARSENAULT, R., AND WARE, C. 2004. The importance of stereo and eye-coupled perspective for eye-hand coordination in fish tank vr. *Presence: Teleoper. Virtual Environ.* 13 (October), 549–559.
- BALLANTYNE, G. 2002. Robotic surgery, telerobotic surgery, telepresence, and telementoring. *Surgical Endoscopy And Other Interventional Techniques* 16, 10, 1389–1402.
- BARABAS, J., SMITHWICK, Q. Y. J., AND BOVE, JR, V. M. 2010. P-3: Evaluation of rendering algorithms for presenting layered information on holographic displays. *SID Symposium Digest of Technical Papers* 41, 1, 1233–1236.

- BARABAS, J., JOLLY, S., SMALLEY, D., AND BOVE JR, V. 2011. Diffraction specific coherent panoramagrams of real scenes. *Proc. SPIE Practical Holography XXV* 7957, 795702.
- BENTON, S. A., AND BOVE, JR, V. M. 2008. *Holographic Imaging*. Wiley, Hoboken, NJ.
- BINGHAM, G., COATS, R., AND MON-WILLIAMS, M. 2007. Natural prehension in trials without haptic feedback but only when calibration is allowed. *Neuropsychologia* 45, 2, 288 – 294.
- BLACKWELL, H. R. 1953. *Psychophysical thresholds: Experimental studies of methods of measurement*. No. 36. Engineering Research Institute, Univ. of Michigan.
- BLANCHE, P. A., BABLUMIAN, A., VOORAKARANAM, R., CHRISTENSON, C., LIN, W., GU, T., FLORES, D., WANG, P., HSIEH, W. Y., KATHAPERUMAL, M., RACHWAL, B., SIDDIQUI, O., THOMAS, J., NORWOOD, R. A., YAMAMOTO, M., AND PEYGHAMBARIAN, N. 2010. Holographic three-dimensional telepresence using large-area photorefractive polymer. *Nature* 468, 7320 (11), 80–83.
- BLEICHER, A. 2010. The gulf spill's lessons for robotics. *Spectrum, IEEE* 47, 8, 9–11.
- BOURDON, J. 2000. Live television is still alive: on television as an unfulfilled promise. *Media, Culture & Society* 22, 5, 531–556.
- CHARMAN, W. N., AND WHITEFOOT, H. 1978. Astigmatism, accommodation, and visual instrumentation. *Appl Opt* 17, 24 (Dec), 3903–10.
- COLLEWIJN, H., AND ERKELENS, C. 1990. Binocular eye movements and the perception of depth. *Reviews of oculomotor research* 4, 213–61.
- CORMEN, T. H., STEIN, C., RIVEST, R. L., AND LEISERSON, C. E. 2001. *Introduction to Algorithms*, 2nd ed. McGraw-Hill Higher Education.

- ELLIS, S., AND MENGES, B. 1997. Judgments of the distance to nearby virtual objects: interaction of viewing conditions and accommodative demand. *Presence (Cambridge, Mass.)* 6, 4, 452.
- FARADAY, M. 1846. Thoughts on ray vibrations. *Philosophical Magazine XXVIII*, S.3 (May), N188.
- FREEDMAN, B., SHPUNT, A., MACHLINE, M., AND ARIELI, Y., 2008. Depth mapping using projected patterns, Apr. 2. US Patent App. 12/522,171.
- GABOR, D. 1948. A New Microscopic Principle. *Nature* 161 (May), 777–778.
- GERSHUN, A. 1936. The light field. *Journal of Mathematics and Physics XVIII*, 51–151.
- GILLAM, B., BLACKBURN, S., AND BROOKS, K. 2007. Hinge versus twist: the effects of 'reference surfaces' and discontinuities on stereoscopic slant perception. *Perception* 36, 4, 596–616.
- GOODMAN, J. W. 1976. Some fundamental properties of speckle. *J. Opt. Soc. Am.* 66, 11 (Nov), 1145–1150.
- GROSSMAN, T., AND BALAKRISHNAN, R. 2006. An evaluation of depth perception on volumetric displays. In *Proceedings of the working conference on Advanced visual interfaces*, ACM, New York, NY, USA, AVI '06, 193–200.
- HAHN, J., KIM, H., LIM, Y., PARK, G., AND LEE, B. 2008. Wide viewing angle dynamic holographic stereogram with a curved array of spatial light modulators. *Opt. Express* 16, 16, 12372–12386.
- HALLE, M. W. 1997. *Multiple viewpoint rendering for three-dimensional displays*. PhD thesis, MIT. Supervisor-Benton, Stephen A.
- HASHIMOTO, N., HOSHINO, K., AND MOROKAWA, S. 1992. Improved real-time holography system with lcds. SPIE, S. A. Benton, Ed., vol. 1667, 2–7.

- HÄUSSLER, R., SCHWERDTNER, A., AND LEISTER, N. 2008. Large holographic displays as an alternative to stereoscopic displays. In *Proc. SPIE*, vol. 6803.
- HINCKLEY, K., PAUSCH, R., GOBLE, J. C., AND KASSELL, N. F. 1994. A survey of design issues in spatial input. In *Proceedings of the 7th annual ACM symposium on User interface software and technology*, ACM, New York, NY, USA, UIST '94, 213–222.
- HOFFMAN, D. M., GIRSHICK, A. R., AKELEY, K., AND BANKS, M. S. 2008. Vergence-accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of Vision* 8, 3 (03).
- JEPSEN, M. L. 1989. *Holographic Video: Design and Implementation of a Display System*. Master's thesis, Massachusetts Institute of Technology.
- JOHNSTON, E. B. 1991. Systematic distortions of shape from stereopsis. *Vision Res* 31, 7-8, 1351–60.
- JULESZ, B. 1971. *Foundations of cyclopean perception*. University of Chicago Press, Chicago.
- KIM, W., AND BEJCZY, A. 1993. Demonstration of a high-fidelity predictive/preview display technique for telerobotic servicing in space. *Robotics and Automation, IEEE Transactions on* 9, 5, 698–702.
- KIM, T., POON, T.-C., AND INDEBETOUW, G. 2002. Depth detection and image recovery in remote sensing by optical scanning holography. *Optical Engineering* 41, 6, 1331–1338.
- KIM, S.-K., KIM, D.-W., KWON, Y. M., AND SON, J.-Y. 2008. Evaluation of the monocular depth cue in 3d displays. *Opt. Express* 16, 26, 21415–21422.
- LEE, J. H., STARK, L. R., COHEN, S., AND KRUGER, P. B. 1999. Accommodation to static chromatic simulations of blurred retinal images. *Ophthalmic and Physiological Optics* 19, 3, 223–235.

- LEITH, E. N., AND UPATNIEKS, J. 1962. Reconstructed wavefronts and communication theory. *J. Opt. Soc. Am.* 52, 10 (Oct), 1123–1128.
- LEITH, E., UPATNIEKS, J., HILDEBRAND, B., AND HAINES, K. 1965. Requirements for a wavefront reconstruction television facsimile system. *Journal of the SMPTE* 74, 10, 893–896.
- LEVOY, M. 2006. Light fields and computational imaging. *Computer* 39, 8 (Aug.), 46–55.
- LUCENTE, M. 1994. *Diffraction-specific fringe computation for electro-holography*. PhD thesis, Massachusetts Institute of Technology, Dept. of Electrical Engineering and Computer Science, Cambridge, Mass.
- MANIA, K., ADELSTEIN, B. D., ELLIS, S. R., AND HILL, M. I. 2004. Perceptual sensitivity to head tracking latency in virtual environments with varying degrees of scene complexity. In *Proceedings of the 1st Symposium on Applied perception in graphics and visualization*, ACM, New York, NY, USA, APGV ’04, 39–47.
- MARQUARDT, D. W. 1963. An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial & Applied Mathematics* 11, 2, 431–441.
- MASOODIAN, M., MCKOY, S., ROGERS, B., AND WARE, D. 2004. Deepdocument: use of a multi-layered display to provide context awareness in text editing. In *AVI ’04: Proceedings of the working conference on Advanced visual interfaces*, ACM, New York, NY, USA, 235–239.
- MON-WILLIAMS, M., AND TRESILIAN, J. R. 2000. Ordinal depth information from accommodation? *Ergonomics* 43, 3, 391 – 404.
- MOORE, K. S., GOMER, J. A., PAGANO, C. C., AND MOORE, D. D. 2009. Perception of robot passability with direct line of sight and teleoperation. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 51, 4, 557–570.

- MURPHY, R. 2004. Trial by fire [rescue robots]. *Robotics Automation Magazine, IEEE* 11, 3 (sept.), 50 – 61.
- O'SHEA, R., AND GOVAN, D. 1997. Blur and contrast as pictorial depth cues. *Perception* 26, 599–612.
- PALMER, S. 1999. *Vision science: Photons to phenomenology*, vol. 1. MIT press, Cambridge, MA.
- PELLI, D. G., ROBSON, J. G., AND WILKINS, A. J. 1988. The design of a new letter chart for measuring contrast sensitivity. *Clinical Vision Sciences*, 2, 187–199.
- PLESNIAK, W., HALLE, M., V. MICHAEL BOVE, J., BARABAS, J., AND PAPPY, R. 2006. Reconfigurable image projection holograms. *Optical Engineering* 45, 11, 115801.
- QUENTMEYER, T. 2004. *Delivering Real-Time Holographic Video Content With Off-The-Shelf PC Hardware*. Master's thesis, Massachusetts Institute of Technology.
- RADI, M., AND NITSCH, V. 2010. Telepresence in industrial applications: Implementation issues for assembly tasks. *Presence: Teleoper. Virtual Environ.* 19, 415–429.
- RASKAR, R., WELCH, G., CUTTS, M., LAKE, A., STESIN, L., AND FUCHS, H. 1998. The office of the future: A unified approach to image-based modeling and spatially immersive displays. In *Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques*, ACM, New York, NY, USA, SIGGRAPH '98, 179–188.
- RAWSON, E. G., NAFARRATE, A. B., NORTON, R. E., AND GOODMAN, J. W. 1976. Speckle-free rear-projection screen using two close screens in slow relative motion. *J. Opt. Soc. Am.* 66, 11 (Nov), 1290–1294.
- SCHARFF, L. F. V., AND AHUMADA, ALBERT J., J. 2002. Predicting the readability of transparent text. *J. Vis.* 2, 9 (12), 653–666.

- SHAKED, N. T., KATZ, B., AND ROSEN, J. 2009. Review of three-dimensional holographic imaging by multiple-viewpoint-projection based methods. *Appl. Opt.* 48, 34 (12), H120–H136.
- SHAKED, N. T. August 2009. A review of incoherent digital fresnel holography. *Journal of Holography and Speckle* 5, 124–140(17).
- SKUPIN, A., AND HAGELMAN, R. 2005. Visualizing demographic trajectories with self-organizing maps. *GeoInformatica* 9, 2, 159–179.
- SMITHWICK, Q. Y. J., BARABAS, J., SMALLEY, D. E., AND V. MICHAEL BOVE, J. 2010. Interactive holographic stereograms with accommodation cues. SPIE, H. I. Bjelkhagen and R. K. Kostuk, Eds., vol. 7619, 761903.
- ST-HILAIRE, P. 1994. *Scalable optical architectures for electronic holography*. PhD thesis, Massachusetts Institute of Technology, Program in Media Arts and Sciences.
- TAKAKI, Y., AND YOKOUCHI, M. 2012. Accommodation measurements of horizontally scanning holographic display. *Opt. Express* 20, 4 (Feb), 3918–3931.
- WATSON, B., WALKER, N., WOYTIUK, P., AND RIBARSKY, W. 2003. Maintaining usability during 3d placement despite delay. In *Virtual Reality, 2003. Proceedings. IEEE*, IEEE, 133–140.
- WICHMANN, F. A., AND HILL, N. J. 2001. The psychometric function: I. fitting, sampling, and goodness of fit. *Perception & psychophysics* 63, 8, 1293–1313.
- WONG, B. L. W., VILLANUEVA, R., AND NEES, A. 2003. Colour and transparency on the multi-layer display (mldTM). Tech. rep., Discussion Paper 2003/05. Department of Information Science, University of Otago, Dunedin, New Zealand.
- YAMAGUCHI, M., HOSHINO, H., HONDA, T., AND OHYAMA, N. 1993. Phase-added stereogram: calculation of hologram using computer graphics technique. SPIE, S. A. Benton, Ed., vol. 1914, 25–31.

YAMAMOTO, K., MISHINA, T., OI, R., SENO, T., AND KURITA, T. 2010. Real-time color holography system for live scene using 4k2k video system. SPIE, H. I. Bjelkhagen and R. K. Kostuk, Eds., vol. 7619, 761906.