Carnegie Mellon

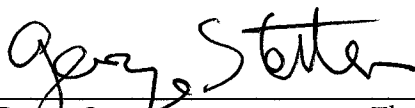## Thesis

Near-field holographic display for in situ merger of real-time tomographic data with direct vision

John Galeotti

Submitted in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
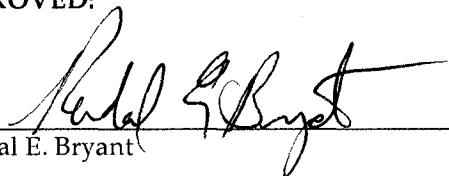in the field of Robotics

ACCEPTED:

George Stetten   Thesis Committee Chair $8/17/07$ Date

Reid G. Simmons   Program Chair $9/23/07$ Date

APPROVED:

Randal E. Bryant   Dean $9/25/07$ Date

Near-Field Holographic Display for In-Situ Merger of Real-Time Tomographic Data with Direct Vision

John Michael Galeotti

CMU-RI-TR-07-30

George Stetten, Chair
Mel Siegel
Martial Hebert
Roberta Klatzky

UMI Number: 3281229

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI®

Abstract

John Michael Galeotti (The Robotics Institute, Carnegie Mellon University, Pittsburgh, PA 15213)
Near-Field Holographic Display for In-Situ Merger of Real-Time Tomographic Data with Direct Vision

Real Time Tomographic Reflection (RTTR) is a technique developed by our laboratory that permits in situ visualization of near-field tomographic images, enabling natural hand-eye coordination to guide invasive procedures. Originally developed for the Sonic Flashlight, RTTR uses a half-silvered mirror to merge the direct view of a patient with a real-time tomographic image from within the patient by reflecting a viewpoint-independent virtual image precisely into its actual anatomic location. For example, the Sonic Flashlight displays a real-time virtual image of an ultrasound data slice at its real location *inside* the patient. This work presents a derivative concept that replaces the half-silvered mirror with a Holographic Optical Element (HOE). This new technique, called Real Time Tomographic Holography (RTTH), eliminates a disadvantage of RTTR, namely, that the virtual image is restricted to the size and shape of the source display. *RTTH enables the tomographic projection of a large virtual image using a small LCD image source*, as required for many medical and non-medical applications. The perceived location of the projected virtual image is essentially *independent of viewpoint*. There are no optical elements between the HOE and the virtual image, allowing *unobstructed viewing of—and interaction with—both the virtual image and the patient or physical object(s) within which the virtual image may be located*. This dissertation presents the design and construction of the first RTTH optical system. It also presents alignment and validation methods for the RTTH system to evaluate its inevitable distortions, in particular the blur and positional instability that must be present, to some degree, in any non-trivial RTTH system.

2

Acknowledgments

I would like to thank my adviser, Dr. George Stetten, for his confidence in me and for the assistance he provided through my graduate research. His guidance is appreciated. I am also grateful to my committee, Dr. Mel Siegel, Dr. Martial Hebert, and Dr. Roberta Klatzky, for their patience and support. Dr. Damion Shelton and Dr. David Wang, prior fellow graduate students at CMU, along with Dr. Bing Wu (post doc) have also been of help, as have Dr. Wilson Chang, Dr. Robert Tamburo, Aaron Cois, Gaurav Shukla, and David Weiser (all University of Pittsburgh). I would also like to thank Dr. Andreas Nowatzyk, who first proposed the idea of real time tomographic holography and provided guidance during my initial research.

My parents, Dr. Gary and Sharon Galeotti, have provided endless love and encouragement through the years. They have cultivated within me a curiosity and a love for knowledge for which I am grateful. More recently, my life has been blessed by the love and companionship of my wonderful wife, Amanda. Despite enduring many late nights, she has continually supported and encouraged me, helping not only with my morale, but also with my grammar and experimentation. During the last year of this work, she has also dealt gracefully with pregnancy, moving, giving birth, and taking care of a newborn. I cannot imagine life, fatherhood, or completion of this dissertation without her. David, the aforementioned newborn and our first child, has also been a wonderful source of joy in the midst of my completing this work.

Finally, I thank my Lord and Savior Jesus Christ for making the ultimate sacrifice, so that I might live. He gives meaning to my life and work, and He sustains me through difficult times. Without Him, this thesis would never have been written.

3

Contents

4

List of Figures

7

List of Tables

9

Nomenclature

λ the operating wavelength of an HOE, page 23

AR   augmented reality, page 7

BZP   binary zone plate, page 28

DOE   diffractive optical element, page 24

DV   direct vision, page 11

FD-2   Final Design 2 of the RTTH system we present, page 59

HMD   head-mounted display, page 11

HOE   holographic optical element, page 21

JPL   NASA's Jet Propulsion Laboratory, page 30

PGD   planar geometric distortion, page 47

RTTH   Real Time Tomographic Holography, page 21

RTTR   Real Time Tomographic Reflection, page 15

SG   stereo graphic, page 11

VHOE   volume phase holographic optical element, page 24

WP   Wasatch Photonics, Inc., page 30

Chapter 1

Background

1.1 Motivation: near field, in situ visualization of virtual images

In the current practice of medicine, images are routinely acquired by ultrasound, computerized tomography (CT), magnetic resonance imaging (MRI) and other modalities. These images are viewed on a film or screen, rather than by looking directly into the patient. This separation between image display and the patient workspace requires the surgeon using the images for real-time guidance to mentally integrate two disparate frames of reference. The difficultly of performing such mental integration is problematic when performing invasive procedures, where direct physical interaction with the region being imaged is required.

Similar problems arise in other areas where it is advantageous to augment the human visual system. Security services routinely screen luggage, and if a bomb is found, it may be advantageous to disarm it without having to open a potentially rigged suitcase. Emergency rescue services (especially firefighters) may use infrared laser scanning technologies to see through smoke. Finally, military personnel benefit greatly from an ability to use sonar to see in murky water, infrared to see at night or through smoke, or in a recent development micro-impulse radar to see through walls.

All of these application areas require human interaction with an environment that must be scanned and visualized in real time. The requirement for physical interaction has two important consequences:

- The virtual image must be **stable** in space. That is, *the perceived location of each point in the virtual image must be essentially independent of viewpoint.*

- The virtual image must be located close to the human, optically in the **near field**.

For reasons discussed later, near field visualization is much more difficult than far field visualization "at infinity." (Optically speaking, an image can be treated as if at an infinite distance when it is far enough away that significant translational movement by the observer does not change the angles at which the image points' light rays strike the observer.)

11

1.2 Augmented reality

A number of researchers have worked to develop more natural ways to merge images with the perceptual real world, thereby, in the case of medicine, removing the clinician's need to shift their gaze between the patient and the image [2, 3, 4, 5, 6, 7, 8, 9]. These techniques fall into the broad category of *augmented reality* (AR). AR is part of what Drascic and Milgram define as the Reality-Virtuality Continuum [10]. At one end of the continuum are purely virtual environments, while at the other end are completely real environments. Between them lie two classes of mixed reality environments. The first of these is augmented virtuality (AV), where a predominantly virtual environment is enhanced by the addition of images of objects from the real world. The second mixed reality technique, AR, enhances what is predominantly a real scene with virtual objects.

Reality
direct view (DV)
stereo video (SV)

Mixed reality

Virtual Reality
stereo graphic (SG)

Augmented
Reality
DV with SG
SV with SG

Augmented
Virtuality
SG with DV
SG with SV
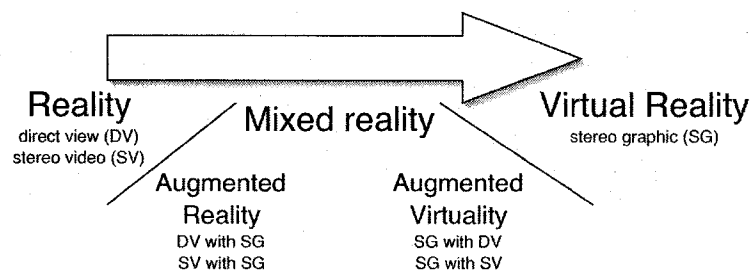
Figure 1.1: Drascic and Milgram's reality–virtuality continuum. Note that there are other types of Augmented Reality, including our own visualization methods, that do not fit into their two example subcategories listed here.

In the task domains mentioned above, all four portions of the continuum are useful (pure virtual environments are useful for training, for example). However, AR techniques are clearly the most appropriate for cases requiring primarily direct physical interaction with the real world but simultaneous real time visualization of scans of some or all of the workplace. In this context, AR seeks to project as virtual objects the images from these scans, and it seeks to project these scans *in situ* (in the location from which they were scanned). For example, in the case of guiding needle biopsy, both the patient and needle are real. The "virtual" addition to the scene may consist entirely of the ultrasound image being used to guide the biopsy procedure.

There are two fundamentally different methods for displaying a virtual object in the context of AR. Unfortunately, the AR community and the optics community ascribe different meanings to some key terms, and so to disambiguate this dissertation will use somewhat unique terminology. The traditional method for displaying an object in virtual reality or AR systems, used almost exclusively in current commercial and research systems, is to display what this dissertation will term a *pseudo virtual image*. The other method, used by our lab and a few others, is to create, in the parlance of optics, a true *virtual image* (as defined in the following section).

Throughout this dissertation, "virtual image," when not prefixed by "pseudo," will mean a true virtual image. Otherwise, by "virtual" we will mean only simulated or synthetic, and accordingly "pseudo virtual image" will mean a perceived 3D scene containing one or more virtual (i.e., simulated or synthetic) objects, but not a true virtual image. (A pseudo virtual image is normally created from two or more physical images mentally combined using stereo vision into the pseudo virtual

image's single perceived 3D scene.)

1.3 Optics basics

At this point, it is appropriate to briefly review the optical properties of *real images* and (true) *virtual images*. We begin by describing *source images*, which the optics literature commonly refers to as *objects*. Source images are directly viewable. Source images can either directly emit light or reflect and scatter incident light. Examples of the former include a TV screen, an organic light-emitting diode (OLED) display, or an LCD (with backlight). The latter would include all externally illuminated physical objects, such as a person, a chair, a rock, or a tree in a field.

A real image is an image that occurs as the result of the in-focus convergence of light rays in space (the direction of light in these "rays" can either be towards or away from the point of convergence). Technically, a source image is a type of real image because each point on it either emits or diffuses light rays that necessarily converge at that point. This dissertation, however, reserves the term "real image" for real images that are *not* source images. Accordingly, for a real image to be visible, a sensor or a diffuse screen must be placed at the location of the real image. For example, a projector produces a real image on a projection screen (Figure 1.2), and the human eye lens produces a real image on the retina.[1] Real images are not suitable for in situ visualization due to the unacceptable requirements of placing a diffuse screen inside the the patient (or suitcase, etc.) and providing visually unobstructed paths from the real image projector and the viewer to that screen.
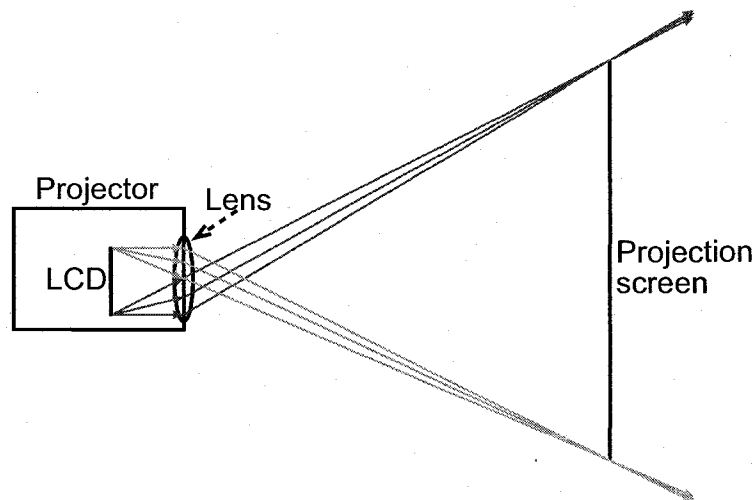
Figure 1.2: The LCD is the source image for the lens of the projector, which forms a larger real image of the original LCD source image. For the real image to be viewed, however, it is necessary to place a projection screen at the location of the real image, where all the light rays come into focus.

---

[1]If a real image is projected onto a diffuse screen the result will be a source image for an observer's eyes to see. Without the screen present, the real image will still exist, but it will not be visible and accordingly will not produce a source image.

In contrast, a virtual image occurs as the result of the *apparent* in-focus convergence of light rays, when in fact the virtual image's light rays do not focus at (or typically even travel through) the virtual image. Accordingly, if an image sensor (such as a camera) is placed at the location of a virtual image, it will not see the virtual image. An object's reflection in a mirror is an example of a virtual image. The magnified image viewed through a magnifying glass is another (Figure 1.3). Virtual images are ideally suited for in situ visualization because of their absence of equipment at the virtual image location, and because they do not require a visually unobstructed path to the virtual image (the light rays never actually travel to, through, or from the virtual image).
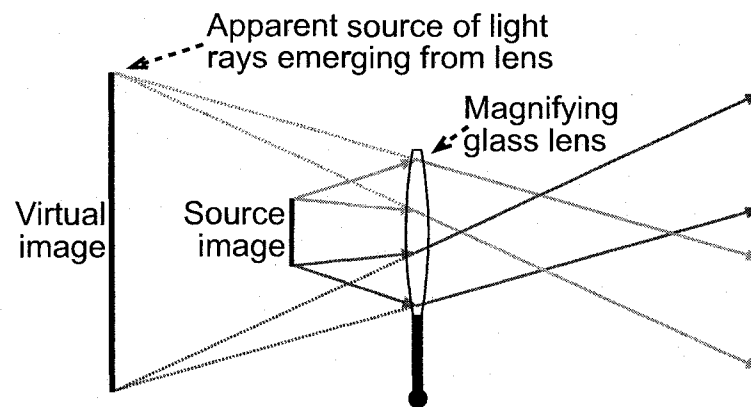
Figure 1.3: The magnifying glass lens forms a magnified virtual image of a smaller source image. The virtual image only appears to exist when viewed through the magnifying glass. The magnifying glass forms the virtual image by bending the light rays from the source image such that they appear to have come from a different location, which is the point where the emerging light rays "back-focus," i.e., where straight-line extensions backwards beyond the origins of the rays converge. Note that no actual light rays from the source image ever pass through the virtual image.

It is now useful to describe a theoretically perfect virtual image (in practice, virtual-imaging optical systems are designed to produce virtual images that are as "perfect" as possible). A perfect virtual image has the key property that it is optically indistinguishable from a real object.[2] No matter how a virtual image is viewed, all light rays appearing to originate from a perfect virtual image are geometrically indistinguishable from the equivalent light rays that would originate from the contents of the virtual image if those contents were real and being viewed directly. Hence, any observer of a perfect virtual image, no matter their vantage point, will see the virtual image in its correct location and will need to focus at the apparent depth of the virtual image to see it clearly. In practice, a mirror is able to produce an almost perfect virtual image; multiple people all looking in a single flat mirror all see each other clearly and at the appropriate distance(s). Virtual images produced by other optical elements, such as lenses or curved mirrors, are not as perfect for reasons discussed later.

---

[2] This does *not* necessarily mean that it is optically indistinguishable from the source object upon which it is based. For example, the virtual image produced by a fun-house mirror that shortens your reflection is clearly different than yourself, but if perfect it would be indistinguishable from a "shorter" version of yourself.

1.4   AR systems displaying a pseudo virtual image

The most common method of achieving pseudo virtual image overlay is to either fully or partially replace the operator's direct vision (DV) by means of a head-mounted display (HMD). Several varieties of HMD exist, with the two most common being optical see-through (Figure 1.4) and video see-through (Figure 1.5) [11]. In the former, a partially reflective (half-silvered) mirror combines direct vision of an object with a stereo graphic (SG) overlay. In the latter, the user's direct vision is completely occluded by video displays that show a combination of images from head-mounted video cameras and the virtual environment generated by a computer. A third HMD design that is just emerging is retinal scanning [12]. Retinal scanning projects the overlaid image directly onto the user's retinas using lasers mounted on the rim of a special pair of glasses.

All HMD implementations described above function by computing how each displayed virtual object should appear to each eye, and then overlaying an appropriate rendering of the virtual object for each eye to see. Because each eye is presented with a different viewpoint-specific rendering, the human stereoscopic visual system is able to deduce the apparent 3D size, position, and orientation of the virtual object(s) based on the difference between the two renderings of the virtual object. To perform such viewpoint-specific rendering, it is necessary to track the position and orientation of the operator's head, any tools or other physical objects necessary to determine the position of the the virtual objects, and the object or person being scanned in order to present a correct augmented image to the user.
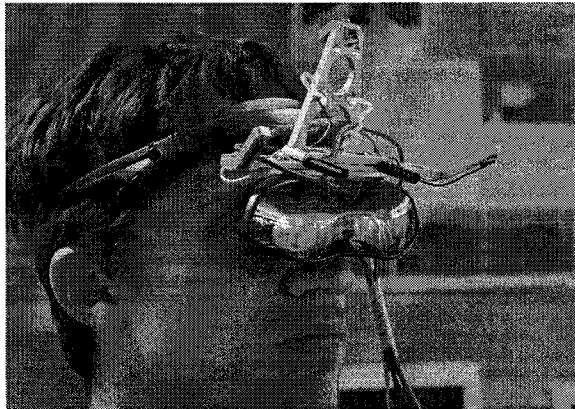
Figure 1.4: Fuch's and State's HMD [1].

It should be noted at this point that the present writing classifies as a pseudo virtual image system any visualization system for which the perceived virtual objects are not directly represented as true virtual images, since to some degree most or all of the problems associated with pseudo virtual images systems apply in such cases. Because it is impossible for a human to focus on a screen placed immediately in front of their eyes, most pseudo virtual image systems do, in fact, project true virtual images floating at an (often short) fixed distance from the user by placing mirror(s) or magnifying lens(es) in front of the HMD's LCD screen(s) [4, 6, 12]. However, those virtual images are *not* coincident with the virtual objects. For example, one of Fuch's HMD systems uses semi-transparent mirrors in the goggles to overlay the renderings of the virtual objects into the wearer's field of view. The mirrors produce virtual images of the source LCD screens, so that the LCD

Figure 1.5: Sauer's HMD and tracking system [2].

screens appear to be floating transparently in front of the wearer's eyes, but there is no virtual image formed at the actual perceived location of the virtual objects.

Drascic and Milgram have identified a number of difficulties with the pseudo virtual image approach [10]. For example, even without stereoscopic vision, each eye is capable of gauging distance based on how its lens automatically focuses (accommodates) in order to see an object sharply. In the real world, accommodation and stereoscopic depth perception work together and are usually in agreement. With an HMD, however, the eyes must accommodate to the fixed distance of the projected LCD images, rather than the stereoscopically perceived distance to the virtual objects being viewed. Optical see-through HMD systems can exacerbate this problem by requiring the user's eyes to accommodate differently for virtual objects than for the real objects located physically at the same range. Other problems Drascic and Milgram identified with pseudo virtual image-based AR systems include interposition failures, over-expanded depth of field, and accommodation-vergence conflict. Interposition failures result from incorrect occlusions or perceived occlusions. An over-expanded depth of field creates problems correctly visualizing objects visible by one eye only. Vergence measures how much the eyes rotate to directly gaze on a common point, and accommodation-vergence conflict results when the eyes must accommodate to a different distance than their vergence point, which for example happens when the HMD's eye-pieces are not precisely separated to match the user's eye spacing. (It should be noted that State developed an HMD system that somewhat reduced the accommodation-vergence conflict, but it was only able to do so at the expense of introducing disparity-vergence conflict [5].)

All of these problems result because the desired illusion of depth in a pseudo virtual image is produced almost exclusively by use of stereo vision, and the other depth cues, which are often either uncontrollable or less than optimally controlled, present conflicting depth information to each other and to stereo vision. In addition, there are significant implementation problems associated with pseudo virtual image systems:

- Extensive tracking hardware is typically required to compute how each displayed virtual ob-

ject should appear to each eye of each observer, which subsequently introduces a number of registration and lag issues as well.

- Those HMD systems that replace direct vision with video suffer a significant loss of resolution.

- The requirement for a head-mounted apparatus can be unwieldy and uncomfortable.

1.5 AR systems displaying a true virtual image

Due to the viewpoint independence possible with a virtual image, in situ visualization systems producing true virtual images of the virtual objects do not need to track one or more observers or to compute in real time how a virtual object will appear from multiple different vantage points. The problems associated with pseudo virtual image systems (listed above) do not apply to (true) virtual image systems. It is for this reason that the present research concentrated on virtual image systems. Virtual image systems do have their own limitations, which will be analyzed in section 2.4.

Currently, aside from this work, there are only three other implementations known to us of true virtual image near-field AR systems [13, 7, 9]. All of these systems project a virtual image by use of a semi-transparent (half-silvered) mirror. It follows from the work of James Clerk Maxwell that *a perfectly flat mirror is the only known optical device theoretically able to produce a perfect virtual image* [14, 15]. By saying a virtual image is "perfect", we mean that it has several properties, including that it is in perfect focus and that it always appears to be in its correct location, regardless of viewing angle. A lens can, properly placed, produce a virtual image, and it has the advantage that it can scale the virtual image relative to the image source (i.e., it can introduce non-unity magnification into the optical system). But as Maxwell and others have proven, *it is impossible to make a perfect magnified virtual image of a real image source* [16].

Because all of these systems use a mirror to project the virtual image of a 2D display, they are restricted to visualizing tomographic data[3] lying on a 2D cross-section. However, in some cases 3D volumes can be visualized by rapidly sweeping the scanned cross-section back and forth, as will be discussed in section 2.2.

A primary reason for the dominance of pseudo virtual image visualization systems is their ability to visualize arbitrarily sized and shaped virtual objects without requiring unwieldy equipment of matching size. This research has utilized holography to extend virtual image-based systems such that they are more capable of visualizing arbitrarily sized and shaped planar[4] virtual objects without requiring unwieldy equipment of matching size. Holography allows:

- A clear, optically unobstructed view of the person or object being scanned

- Simultaneous, viewpoint-independent direct viewing of an in-situ virtual image, with the virtual image physically larger than the source image used to project it

- Lack of any physical objects between the holographic "viewport" and the person or object being scanned

---

[3]Tomographic data is 2D or 3D data that represents points in space, rather than projections (like X-ray images).
[4]Future research may be able to extended our work to visualization of volumes.

Chapter 2

Real Time Tomographic Reflection

As described in Chapter 1, current AR visualization systems producing true virtual images do so by use of half-silvered mirrors. Since those techniques are designed for visualization of tomographic data, we have termed this group of techniques *Real Time Tomographic Reflection* (RTTR).

RTTR eliminates the need for tracking either the observer or the patient. The lack of tracking is possible because of the nature of virtual images.[1] Given a 2D source display, RTTR can only be used to visualize a single tomographic slice at a time.[2] Ultrasound produces a tomographic slice within the patient representing a set of 3D locations that lie in a plane. Three dimensional imaging modalities, such as CT and MRI, can also be sampled to extract an arbitrarily oriented tomographic slice through the 3D data. The image of any of these tomographic slices can be displayed at its correct size on a flat display, and may be reflected to appear to occupy the same physical space as the actual slice within the patient. If a half-silvered mirror is used, the patient may be viewed through the mirror with the reflected image of the slice superimposed, independent of viewer location. The reflected image is truly occupying its correct location within the patient, and therefore does not require any particular perspective to be rendered correctly.

To accomplish RTTR, certain geometric relationships must exist between the tomographic slice being scanned, the monitor displaying the ultrasound image, and the mirror. As shown in Figure 2.1, the mirror must bisect the angle between the slice and the monitor. On the monitor, the image must be correctly translated and rotated so that each point in the image is paired with a corresponding point in the slice, to define a line segment perpendicular to, and bisected by, the mirror. By fundamental laws of optics, the ultrasound image will thus appear at its physical location, independent of viewer position.

---

[1] As a reminder of the notation established in section 1.2, "virtual image" here denotes a true, optical virtual image, not a "pseudo virtual image" such as one stereographically rendered by a HMD.

[2] Autostereoscopic displays based on stacked, transparent LCDs could potentially mitigate this restriction, but their cumbersomeness is less suitable for hand-held devices. A lenticular display (an autostereoscopic parallax panoramagram that uses lenses rather than physical barriers [17]) is a 3D display that could also potentially mitigate this restriction [18], but lenticular displays currently have low refresh rates, low resolution, and render only a small number of predetermined viewpoints.

18
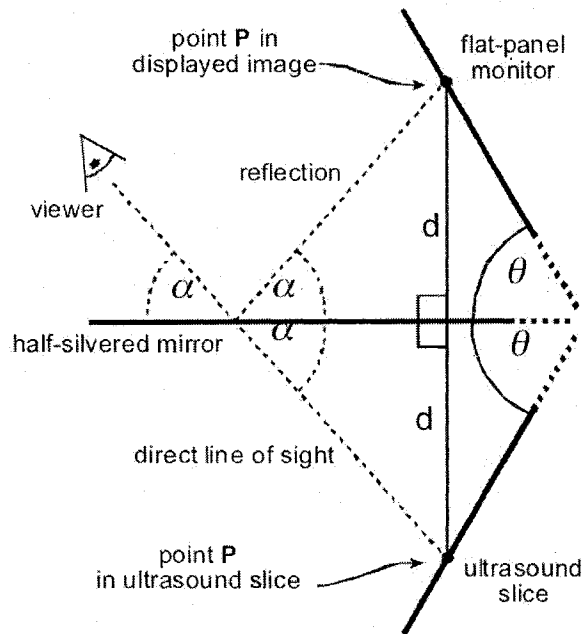
Figure 2.1: Configuration of the Sonic Flashlight: A half-silvered mirror bisects the angle between the ultrasound slice (within the target) and the flat-panel monitor. Point P in the ultrasound slice and its corresponding location on the monitor are equidistant from the mirror along a line perpendicular to the mirror (distance = d). Because the angle of incidence equals the angle of reflectance (angle = α) the viewer (shown as an eye) sees each point in the reflection precisely at its corresponding physical 3D location, independent of viewer location.

2.1   Hofstein's display apparatus

Hofstein proposed the first RTTR system for in situ visualization in 1980 [13]. He displayed an ultrasound slice in real time, strategically positioning an ultrasound transducer, a half-silvered mirror, and a display, such that the virtual image produced by the mirror was "spatially superimposed within the exact region of the patient's body being scanned." Hofstein's invention was intended to make it obvious to a doctor what anatomy was being scanned, for the purpose of simplifying interpretation of the ultrasound data. It consisted of a large box that rested directly on the patient, thereby physically blocking access to the region being scanned, generally precluding his device's use for any interventional procedures. Accordingly, his device was for visualization purposes only, and therefore it had much less stringent requirements for registration accuracy between the reflected virtual image and the actual ultrasound scan.

2.2   The Sonic Flashlight

The Sonic Flashlight represents an independent rediscovery of the idea of RTTR and was the first to apply it to the guidance of interventional procedures [19]. The apparatus of the Sonic Flashlight is

Figure 2.2: Schematic representation of Hofstein's original system for RTTR

depicted in Figure 2.3. In Figure 2.4, a human hand is seen with the transducer pressed against the soft tissue between the thumb and the index finger. While not a common target for clinical ultrasound, the hand was chosen because it clearly demonstrates successful alignment. The ultrasound image is consistent with the external landmarks of the hand. The photograph cannot convey the strong sense–derived from stereoscopic vision, convergence, and accommodation–that the reflected image is located within the hand. This perception is intensified with head motion because the image remains properly aligned from different viewpoints. To one experiencing the technique in person, anatomical targets within the hand visible in the ultrasound would clearly be accessible to direct percutaneous injection, biopsy or excision.

Another key feature of the Sonic Flashlight is that it is the only known RTTR system that is handheld. Accordingly, it is rapidly and easily possible to acquire and visualize a wide range of *arbitrarily oriented* slices. Stated differently, the user can view an arbitrary slice simply by "placing" the virtual image produced by the Sonic Flashlight in the desired location. Furthermore, such an ability to visualize arbitrary slices in real time allows for the possibility of visually interrogating an entire volume of data, simply by sweeping the Sonic Flashlight through the volume. The human visual system is, to some extent, capable of integrating the in situ *slices* over time, building up an internal model of the entire volume. Hence even though the Sonic Flashlight can only display one plane of data at a time, it can be used to visualize entire volumes. Our laboratory is currently conducting research comparing this ability with that provided by conventional ultrasound displays [20, 21, 22].

Superimposing ultrasound images onto human vision using RTTR improves an operator's ability to find targets while avoiding damage to neighboring structures and generally facilitating interpretation of ultrasound images by relating them spatially to external anatomy. As such, it holds promise for increasing accuracy, ease, and safety during percutaneous biopsy of suspected tumors, brain surgery [23], insertion of catheters [24], eye surgery [25], and many other interventional procedures. Our laboratory has been developing the Sonic Flashlight for eight years. We have tested the Sonic Flashlight on vascular phantoms [24] and have recently conducted clinical trials on more than 70 patients to place vascular catheters [26]. We have also developed a virtual Sonic Flashlight for human factors experiments and analysis [27, 28]. The virtual Sonic Flashlight has no actual scan-

Figure 2.3: Schematic representation of the Sonic Flashlight apparatus. A flat-panel monitor and an ultrasound transducer are placed on opposite sides of a half-silvered mirror such that the mirror bisects the angle between them.

Figure 2.4: Photograph, from the viewpoint of the operator, showing a scan of a hand using the apparatus depicted in Figure 2.3. The reflected ultrasound image is merged with the direct visual image.

ning apparatus, but rather is tracked and the "ultrasound data" synthesized by a computer according to the position and orientation of the probe. This device allows us to increase our understanding of psychophysical phenomena and help improve the design of the standard Sonic Flashlight.

2.3 Masamune's CT-based system

Masamune, et al., are the only other group we are aware of to apply RTTR to image guided intervention. They have demonstrated RTTR using a CT scanner, rather than ultrasound [29]. The application to CT was also earlier proposed, but not implemented, by Stetten [19]. By properly mounting a flat-panel display and a half-silvered mirror above the gantry of a CT scanner, a slice displayed on a flat panel monitor can be reflected by the half-silvered mirror to its correct location within the patient, as shown in Figure 2.5. Assuming the patient remains motionless between the

time of the CT scan and the viewing, no tracking is required. However, without repeated scans, the CT image will not be correctly updated during any invasive procedure that changes anatomical structures. The practicality of providing sufficiently continual updates during a procedure is questionable, given the presence of ionizing radiation. Ultrasound does not pose this problem.[3]
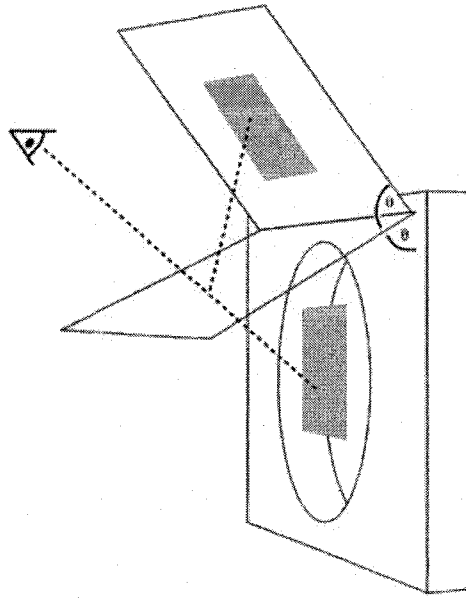
Figure 2.5: Schematic representation of RTTR for CT apparatus

2.4 RTTR analysis

The above examples demonstrate the concept of RTTR for *in situ* visualization of virtual images. In practice, there is a major limitation of RTTR, which is its inability to display a large virtual object without requiring an *equally* large display source, such as the large flat-panel display used with a CT scanner by Masamune. Many applications, including ultrasound, require a portable system in which the physical display is not so large as to get in the way. Ultrasound guidance for deeper procedures such as liver biopsy or amniocentesis would be facilitated by a hand-held device that produces a large virtual image reaching deep within the patient, provided that such a device is neither heavy nor unwieldy. In non-medical applications, firefighters or scuba divers would benefit from an ability to see images of virtual objects scanned using various technologies such as laser range finders or sonar, that can interrogate through smoke or turbid water a few meters in front of them, but they must necessarily wear all the optical equipment used to produce such large virtual images. RTTR is insufficient for these cases, because it cannot produce virtual images larger than the original display source. This insufficiency limits the application of RTTR in medical areas such as deep abdominal ultrasound or other potential non-medical areas such as luggage inspection for bomb detection and disarmament using micro impulse radar.

---

[3]While sufficiently powerful ultrasound could be dangerous, the power levels used in medical ultrasound transducers pose no known risk.

Another limitation of conventional RTTR is that the shape of the virtual image must be identical to that of the display source. Most ultrasound scanners produce a wedge-shaped region in which the narrow end of the wedge (closest to the probe tip) has much higher resolution than the broad section of the wedge (deepest into the patient). Conventional displays, on the other hand, are rectangular with uniform resolution throughout. In some cases, it may be desirable to map the pixels of the display to the actual scanned pixels[4] of the ultrasound, which would require that something similar to a projective transform be applied to the display source.[5] Moreover, some catheter-based ultrasound probes produce tomographic slices that are curved rather than flat. An RTTR implementation using such a probe would require that the source display be curved to match the virtual image.

When developing an RTTR system, there are several parameters that must be optimized and/or fall within certain bounds. Fortunately, because modern glass mirrors are very precise devices, most of the error metrics typically associated with optics design can safely be ignored. However, the thickness of the glass in the half-silvered mirror should be minimal to keep *refractive error* within acceptable bounds; this is the only variable that affects the view-point independence of the system. That is, the glass interferes with the viewer's perception of the location of the virtual image, limiting the location stability of the perceived virtual image when viewed from a wide range of acceptable viewing locations. It should be noted that in the design of the clinical version of the Sonic Flashlight, our laboratory has had to use plastic mirrors, for reasons of safety and sterility, and these are far less precise than their glass equivalents.

Other important parameters for RTTR development include:

1. The fraction of light (intensity) reflected by the mirror

2. The fraction of light (intensity) transmitted through the mirror

3. The resolution of the virtual image

4. The non-occluded range of acceptable viewing angles

One final aspect of RTTR that has not been discussed yet is that of calibration. All RTTR implementations require calibration so that each pixel in the virtual image occupies its correct location in space. RTTR calibration procedures have been developed for the cases of both the real and virtual Sonic Flashlights [30, 31], with the virtual Sonic Flashlight requiring the more complex mathematics of the two.

---

[4]Technically, an ultrasound slice represents a co-planer collection of voxels rather than pixels.

[5]While such transforms could be accomplished in software prior to generating the original source image, software is not capable of preserving the original scanning resolution by remapping the relative locations of individual pixels.

Chapter 3

Real Time Tomographic Holography

This research uses holography to partially overcome the fundamental limitations of RTTR to allow true virtual-image-based visualization of arbitrarily sized and shaped virtual objects without requiring unwieldy displays of matching size. We use a *holographic optical element* (HOE) in place of a mirror for in-situ real time tomographic image projection, so we term the technique *Real Time Tomographic Holography* (RTTH). To be useful, RTTH requires an HOE capable of producing a near field virtual image that is larger and possibly shaped differently than the source LCD image. As such, RTTH holds the potential of an expanded set of applications relative to those of RTTR.

As shown in Figure 3.1, a simple configuration for a RTTH display consists of a source of monochromatic light (a laser), which passes through an LCD shutter[1] similar to those found in most laptop computers (though smaller and without the usual backlight). The shutter contains many individual pixels that control the passage of the laser light, based on an image derived in real time from a scanner rigidly attached to the apparatus. This could be an ultrasound scanner as in the Sonic Flashlight. The HOE works something like a mirror, creating a virtual image. The difference is that, unlike a planar mirror, the HOE can produce a virtual image whose size, shape, and relative position differ from the source image. The other difference is that the virtual image produced by an HOE will never be perfect, because only a mirror can produce a perfect virtual image (as discussed in section 1.5). Accordingly, *the primary goal of RTTH is to produce a system in which each pixel in the LCD translates to a pixel in the virtual image that is viewpoint independent to within some small tolerance when observed from within a specified field of view.*

Due to its flexibility in the size and shape of the virtual image, RTTH has potential non-medical applications, some of which require very large virtual images. Unlike RTTR, RTTH is well-suited for the applications listed in section 2.4. It could enable scuba divers to see in dark or turbid water, by mounting an RTTH display and a sonar device to a face mask. Mounting an infrared laser rangefinder to a fireman's helmet, it should be possible to project a virtual image several meters in front of an individual, by affixing an HOE to the inside of the visor. This could be very useful for finding a victim in smoke, for example, as shown in Figure 3.2.

Accordingly, this research has focused on overcoming the aforementioned limitations of RTTR by use of one or more custom-designed holographic optical elements, potentially coupled with additional conventional optics, such as lenses, mirrors, and/or fiber optic faceplates.

---

[1]Because an LCD is grayscale rather than black and white, it is better described as a spatial light modulator (SLM) than as a shutter, but "shutter" is a more accessible term.

24
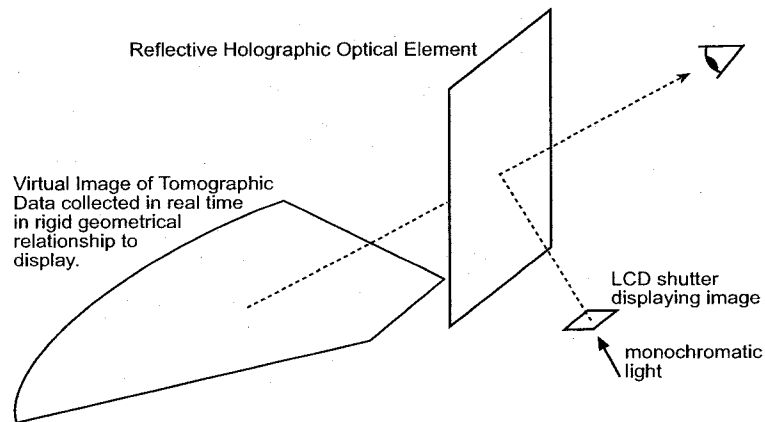
Figure 3.1: A simple configuration for an RTTH display. The LCD shutter creates a real time image from laser light. Each pixel is mapped by the HOE into a virtual image whose size and shape is determined by the particular HOE. The virtual image is independent of viewer location.
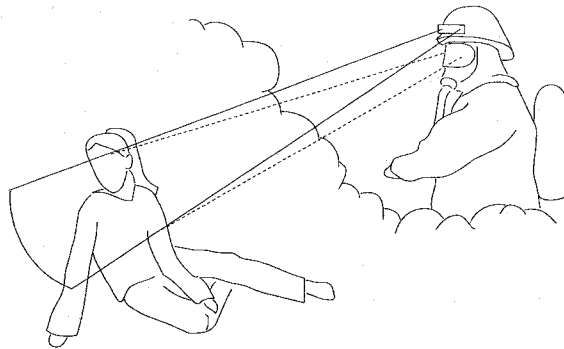
Figure 3.2: An RTTH display projects a nearly stable virtual image at the actual location of the data in real time, as shown here displaying infrared laser rangefinder data through smoke using an HOE in a fireman's visor.

3.1 The need for holography

As previously mentioned in section 1.5, a perfectly flat mirror is the only known optical device able to produce a perfect virtual image floating in 3D space. Unfortunately, a flat mirror is incapable of magnification, and it is therefore unable to produce a virtual image of an object larger than the source display. Likewise, a flat mirror is unable to produce a reflection of a different shape than the source image.

A properly placed lens, on the other hand, can produce a virtual image of a different size than the source display. Unfortunately, it has been theoretically proven that it is impossible to make a perfect magnified virtual image of a real image source [16]. In addition, because of the need to view the virtual image by looking through the lens, the lens would distort the view of reality on the other

side of the lens. One of the important features of most AR systems is that they do not significantly distort the appearance of reality other than for the addition of the virtual objects. In the case of RTTR applied to guiding invasive procedures, it is critical to be able to clearly see both the exterior of the subject under examination and the virtual object(s) inside the subject, since all physical tools will interact with both the exterior and the interior of the subject. A half-silvered mirror nicely allows viewing both the subject's exterior (by the mirror's transparency) and the virtual image inside the subject (created by the mirror's reflection). In effect, a half-silvered mirror acts like a mirror for some fraction of the rays striking it, but like a clear piece of glass for the remaining fraction of rays. A lens, on the other hand, bends *all* of the light rays passing through it, including those from the subject. Accordingly, a lens presents the viewer with a distorted view of the subject's exterior. A lens is also very limited in its ability to transform the shape of a source display into a differently shaped virtual image, although it does have the ability to perform some concentric distortions (such as pincushion) and, when used in an off-axis configuration, to skew an image.

An HOE, on the other hand, simultaneously operates at multiple diffractive orders. *Accordingly, an HOE can be designed to exhibit the behavior of a co-planar collection of lenses for one pre-chosen wavelength of light (first order diffraction), while allowing wavelengths other than the operating wavelength to pass through more-or-less unaffected (zeroth order diffraction).* Unlike a "half" silvered mirror, in which there is a fixed ratio (often not actually 50%) of reflection-to-transmission over a broad spectrum of frequencies, a transmissive HOE can be essentially transparent except to its operating wavelength, λ^2. The HOE transmits almost all of the light from the direct visual image of the object, and adds to that the monochromatic virtual image, potentially resulting in more light directly from the object as well as a brighter (but inherently "grayscale") virtual image. Provided the source image is illuminated at the operating wavelength, a magnified virtual image of the source can be viewed while looking through the HOE, and yet the rest of reality viewed through the HOE will appear almost entirely undistorted. The allowed decrease in size of the LCD relative to the virtual image can reduce the weight of a hand-held device, and avoid occlusion caused by having the actual display in the line of sight. *HOEs also have the advantage of allowing a geometric transformation to be applied between the source image and the virtual image, such as that performed by a prism or diffraction grating.* Each of these factors may make RTTH also preferable to RTTR for certain applications where only a small virtual image is necessary.

3.2 Theory of hologram operation

Holography is typically used to create a static virtual image, such as the security logos embossed on major credit cards. Holography works by manipulating the amplitude and/or phase of incident light. An HOE is usually designed to operate on monochromatic light. We denote by λ the operating wavelength of such an HOE. The manipulation of a monochromatic wave front by an HOE can be fully characterized by the HOE's transfer function (TF), which describes how each location on the HOE attenuates and shifts the phase of incident light. The present writing uses a Cartesian coordinate system, (x_h, y_h), to index all locations across the surface of an HOE. An arbitrary HOE transfer function containing an attenuation component $A(x_h, y_h)$ and a phase shift component $\phi(x_h, y_h)$ can

---

[2]Accordingly, the ambient lighting should not contain a high percentage of light with a wavelength near λ, but in practice this is not usually a concern.

be written in the form:

$$T(x_h, y_h) = A(x_h, y_h) e^{i\phi(x_h, y_h)} \tag{3.1}$$

An HOE that operates only by manipulating the phase of incident light is classified as a diffractive optical element (DOE). The most general types of DOE are the kinoform (see subsection 3.3.2) and volume phase holographic optical element (VHOE) (see subsection 3.3.3), both of which are able to implement the transfer function of any other type of diffractive optic, such as a lens or a diffraction grating. The transfer function of any DOE is generally regarded as having zero attenuation, which is optimal for maximizing virtual image brightness:[3]

$$T(x_h, y_h) = e^{i\phi(x_h, y_h)} \tag{3.2}$$

A few terms will now be defined before proceeding. A *wave front* is a mathematical surface in space corresponding to a contour of constant phase, which for a point light source is the locus of all points at a fixed number $n \in \Re$ of wavelengths from that source. A light ray is a conceptual construct perpendicular to a wave front. Changing the shape of a wave front is therefore equivalent to changing the direction of the rays associated with the wave front. When a ray passes through or reflects off a surface, it is always the case that the input ray, the surface normal, and the output ray all lie within a single plane called the *plane of incidence*. When working with diffractive optics, it is more convenient to measure the propagation of a ray from one point to another using *optical path length* along the ray instead of counting the number wavelengths along the ray. The optical path length of a ray n wavelengths long can be thought of as the distance the light ray would traverse through air, which is typically longer than the distance the ray would actually traverse if some or all of its n wavelengths passed through an optical lens. This is due to the higher index of refraction of the lens, which slows down the light causing more wavelengths to fall within a fixed distance. Finally, a *thin lens* is an optical element that can be approximately modeled as having zero thickness, so that an incident ray may be regarded as leaving the lens at the same point at which it entered. The operating portion (not including the substrate) of virtually any HOE is a thin lens, and so if a ray is incident at coordinates (x_h, y_h), then that ray will also leave from the coordinates (x_h, y_h).

Both kinoform and VHOE types of DOE work by changing the phase of incident light at each point on the DOE's surface. A DOE can be either reflective or transmissive, but we will only analyze the transmissive type since it is simpler to explain (the analysis of the reflective type is very similar). If an incoming ray with wavelength λ is incident to a DOE at the point (x_h, y_h), then the DOE increments the optical path length of the incident ray by the amount:

$$P(x_h, y_h) + c, \quad 0 \le P(x_h, y_h) < \lambda \tag{3.3}$$

where $P(x_h, y_h)$ represents the manufactured holographic profile of the DOE, and c is a much larger but constant term due to the refractive index and underlying thickness of the DOE substrate. Because adding the constant c to the optical path length of all incident rays has no effect on the shape of the wave front passing through the DOE, the c term can be ignored, and the behavior of the DOE is fully specified by $P(x_h, y_h)$, which is called the *phase function* and is analogous to the phase shift $\phi(x_h, y_h)$ in Equation 3.2, except that $\phi(x_h, y_h)$ is in units of radians while $P(x_h, y_h)$ is in units of

---

[3]Like any optic, a DOE introduces some overall attenuation, but it does not use an attenuation function to bend the light rays passing through it, and so the typically small and constant attenuation is ignored.

optical path length. Also, just as phase shift can be taken modulo 2π without changing its effect, likewise a phase function can be taken modulo λ without changing its effect on a wave front, while greatly simplifying its manufacture as a DOE. In practice, $P(x_h, y_h)$ may be discrete valued (especially in the case of the kinoform, described in subsection 3.3.2), quantized to anywhere between 16 and 256 different levels, but in such cases it can still be approximated by a continuous function, especially when, as in our case, 256 levels would be available if discretization were necessary.
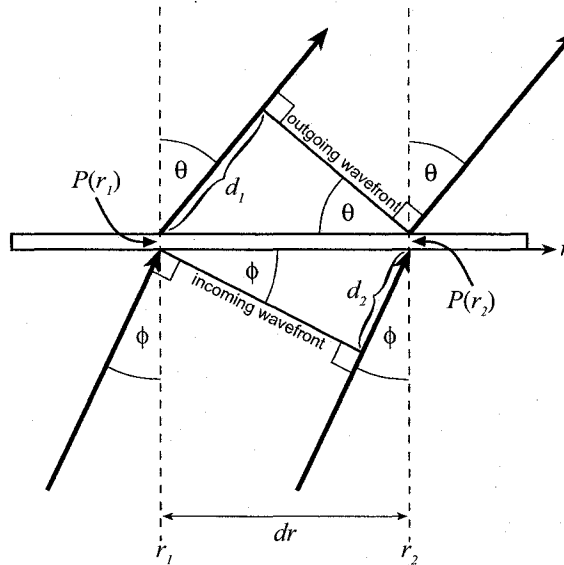
Figure 3.3: The effect of a DOE on an incoming wave front and associated rays
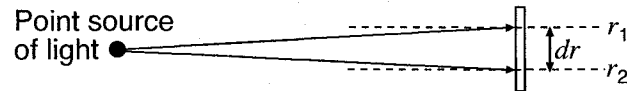
Figure 3.4: These two rays are parallel in the limit as dr approaches 0.

A ray entering a DOE at position (x_h, y_h) will exit the DOE at (x_h, y_h) in the plane of incidence with a direction completely determined by the incoming ray's direction and the gradient of the phase function with respect to (x_h, y_h). To see why, we examine Figure 3.3, which shows the behavior of two rays of light incident to a thin lens transmissive DOE at r_1 and r_2. Both rays are from the same point source of light, as shown in Figure 3.4. We are only interested in deriving the behavior of the ray incident at r_1, and we will do this by taking the limit as the neighboring ray from the same point source (but instead incident at r_2) moves closer to the point r_1 until the two rays converge such that they are both incident at r_1 ($\lim_{dr \to 0}$). Because we are taking the limit, we can assume that both rays lie in the same plane of incidence. This is because in the limit as dr (the distance between r_1 and r_2) approaches 0, the two incoming rays become coincident, resulting in a single direction for the surface normal. Because the surface normal of the DOE at r_1 and the direction of the incoming ray at r_1 are sufficient to determine the plane of incidence for the r_1 ray, we can, without loss of generality, derive the system behavior based only on the angles and coordinates within that common plane of

incidence. In Figure 3.3, the r axis is the intersection of the plane of incidence and the plane of the DOE. As the two incoming rays approach the same angle of incidence ϕ (such that the rays are parallel), the section of the wave front between the two rays approaches a straight line orthogonal to the two rays, as is depicted for the incoming and outgoing rays and wave fronts. Since each depicted wave front represents a common crest along the two different rays, the optical path lengths along both rays from one crest in the incoming wave front to another crest in the outgoing wave front must be the same. Therefore, the following equality must hold for the various optical path lengths labeled in Figure 3.3:

$$P(r_1) + d_1 = d_2 + P(r_2) \tag{3.4}$$

$$\therefore$$

$$d_1 = d_2 + P(r_2) - P(r_1) \tag{3.5}$$

Now, by construction we know that in Figure 3.3

$$\sin(\phi) = \lim_{dr \to 0} \frac{d_2}{dr} \tag{3.6}$$

$$\sin(\theta) = \lim_{dr \to 0} \frac{d_1}{dr} \tag{3.7}$$

The limits in the above equations are necessary because the geometry of Figure 3.3 only holds in the limit as dr approaches zero. Substituting Equation 3.5 into Equation 3.7, we have

$$\sin(\theta) = \lim_{dr \to 0} \frac{d_2 + P(r_2) - P(r_1)}{dr} \tag{3.8}$$

Splitting apart the right-hand side of Equation 3.8 into two terms, we have that

$$\sin(\theta) = \lim_{dr \to 0} \frac{d_2}{dr} + \lim_{dr \to 0} \frac{P(r_2) - P(r_1)}{dr} \tag{3.9}$$

Substituting Equation 3.6 into Equation 3.9, we have

$$\sin(\theta) = \sin(\phi) + \lim_{dr \to 0} \frac{P(r_2) - P(r_1)}{dr} \tag{3.10}$$

Now we will investigate the behavior of the second term in the right-hand side of Equation 3.10, making reference to Figure 3.5, which graphs some particular phase function $P(r)$, along with its tangent at the point $(r_1, P(r_1))$. It should be apparent in Figure 3.5 that

$$\lim_{dr \to 0} \hat{P}(r_2) = P(r_2) \tag{3.11}$$

and that

$$\left. \frac{dP(r)}{dr} \right|_{r=r_1} = \frac{\hat{P}(r_2) - P(r_1)}{dr} \tag{3.12}$$
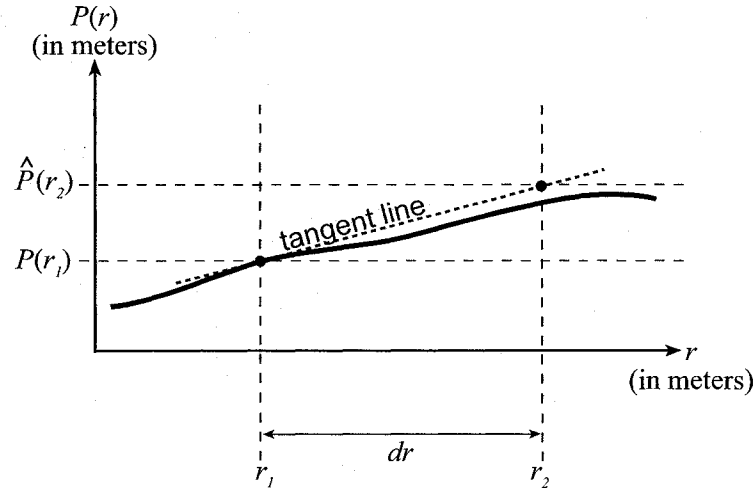
Figure 3.5: Plot of the phase function

Substituting Equation 3.11 into Equation 3.12 we get

$$\left. \frac{dP(r)}{dr} \right|_{r=r_1} = \lim_{dr \to 0} \frac{P(r_2) - P(r_1)}{dr} \tag{3.13}$$

Finally, we are now able to substitute Equation 3.13 into Equation 3.10 to achieve the following result:

$$\sin(\theta) = \sin(\phi) + \left. \frac{dP(r)}{dr} \right|_{r=r_1} \tag{3.14}$$

Given the angle ϕ of an incoming ray within the plane of incidence, the point at which it enters the DOE, and the partial derivative of the DOE's phase function along the plane of incidence at that point, Equation 3.14 allows us to compute the angle θ of that ray relative to the DOE's surface-normal within the plane of incidence as it leaves from the same location.

3.3 HOE production techniques

There are multiple ways to produce an HOE. At this point it is appropriate to briefly discuss the form and basic production process for a few key types of HOE: the binary zone plate, the kinoform, and the VHOE.

3.3.1 Binary zone plate

Figure 3.6 illustrates the basic operation of the reflective binary zone plate (BZP) type of HOE, optically created in this case by photographic means. During its creation, an interference pattern between spherical wave fronts from two point sources of in-phase monochromatic light (A and B) is captured photographically on the surface of the BZP as regions of permanent local reflectivity at the constructive peaks. Subsequent monochromatic light emitted only from point A will be reflected by

the photographically captured interference pattern in the HOE, appearing to emanate from point B independent of viewer location. Points A and B are referred to as the *construction points*. A BZP is a binary recording; each location on the BZP is either fully reflective or not reflective at all.
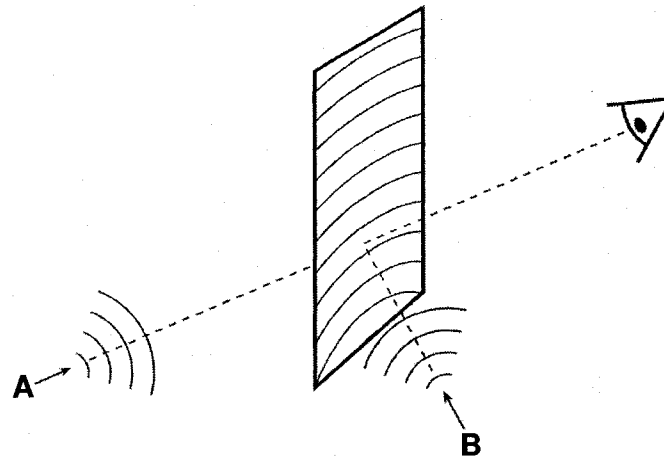
Figure 3.6: The interference pattern produced by two sources of laser light, A and B, is stored in the reflective BZP. Thereafter, a pixel of laser light emitted from B is reflected by the interference pattern in the BZP so as to appear to emanate from A.

A BZP can be produced photographically. As shown in Figure 3.7, a laser can be directed through a variable (polarizing) beam splitter to provide a pair of in-phase beams whose relative amplitudes can be adjusted. The beams can then be reflected by two mirrors (M1 and M2) to illuminate two pinhole spatial filters that are positioned along the desired optical axis of the BZP. The first pinhole should be positioned at the center of where the LCD shutter will be located, while the second pinhole should be positioned at the mirror-image (for a reflective BZP) of the center of the desired virtual image. The two optical path lengths should be matched. The resulting spherical wave fronts create an interference pattern on the photographic plate. Figure 3.8 shows such an apparatus in operation by others in our lab in an early attempt to produce a holographic plate.
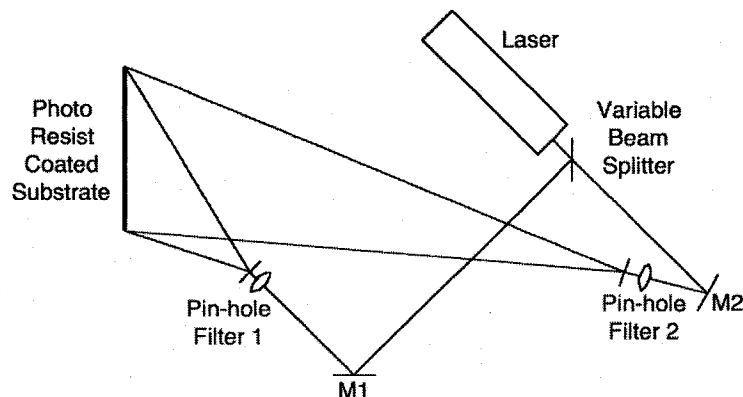
Figure 3.7: Optical layout for generating a reflective binary zone plate
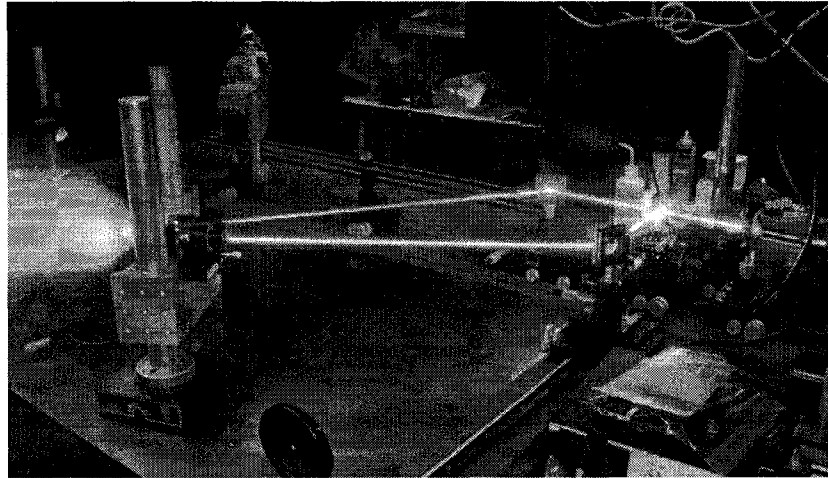
Figure 3.8: Apparatus to produce a reflective binary zone plate. Laser is split (right) and directed through 2 pinholes (left) to create an interference pattern, for projection onto a photographic plate (laser beams enhanced for visibility by long exposure and physically moving a target along the beams).

3.3.2 Kinoform

A transmissive kinoform is physically similar to a sheet of glass, except that one of the surfaces is not planar when viewed at a microscopic level. Instead, the non-planar surface varies in thickness, possibly being very jagged at the sub-wavelength scale. At each point a kinoform adds to the optical path length of any incident ray an amount proportional to the kinoform's thickness at that point. Accordingly, $P(x_h, y_h)$ indirectly represents the profile of the kinoform's non-planar surface. (A kinoform's substrate thickness is the thickness at points where $P(x_h, y_h) = 0$.) NASA's Jet Propulsion Laboratory (JPL) has one of the few facilities capable of manufacturing a large, high resolution kinoform. They do this by using analog electron-beam lithography [32], allowing very precise quantization of the kinoform's surface, both in terms of thickness (256 different levels) and feature size (approximately $1/4$ wavelength). For some time our laboratory negotiated with JPL to produce an HOE for this research, but the efforts were abandoned because of irreconcilable issues involving intellectual property.

3.3.3 VHOE

A volume phase holographic optical element (VHOE) works by changing the refractive index of every point on a very thin layer of a special material, which is sandwiched between two glass sheets. The effect is to change the optical path length at each point, similar to the effect of a transmissive kinoform but achieved by local, microscopic manipulation of refractive index rather than of glass thickness. Accordingly, $P(x_h, y_h)$ indirectly represents the profile of the VHOE's refractive index. A VHOE is often preferable to a kinoform due to its completely uniform thickness, smooth exteriors, and (typically) higher effective recording resolution.

Wasatch Photonics, Inc. (WP) is the manufacturer we have used for VHOE production. WP records the holographic profile into a layer of dichromated gelatin (DCG) sealed between two sheets of glass. The recording is made using the simultaneous exposure of two laser sources, as in section 3.3.1, but WP records an entire phase function into a VHOE, as opposed to the binary recording of a BZP. WP makes use of proprietary techniques and a DCG recording medium to produce an evenly exposed VHOE with high diffraction efficiency (allowing for a bright virtual image) and low loss (high transparency), all of which are important for an RTTH system.

WP works closely with the Ralcon Development Lab, run by Richard Rallison, Ph.D. (co-founder of WP), which graciously provided us with three demonstration VHOEs designed for applications somewhat similar to ours. The sample VHOEs clearly demonstrated the basic optical quality achievable with a VHOE produced by WP. All three VHOEs were capable of producing a magnified off-axis virtual image of a 532 nm image source. Unfortunately, not only did these VHOEs have limited magnifying power, but they were also designed to project virtual images that appear to be an infinite distance away (i.e. from any vantage point all the rays appearing to come from the virtual image are parallel). Nevertheless, we were able to use the demonstration VHOEs to produce a proof-of-concept demonstration of RTTH, as shown in Figures 3.9-3.10. When a demonstration VHOE was properly oriented between the viewer and the LCD, it would project a small but nevertheless magnified virtual image of the LCD. Because each demonstration HOE was designed to project a virtual image at infinity, when it was positioned to produce a near-field virtual image, that virtual image was not stable. Rather, the virtual image would appear to move in space as one changed viewing position.

3.4 Existing real time holographic systems

One of the most well known real time holographic systems is the head-up display (HUD) system used primarily in combat aircraft to locate vital information where the pilot would naturally be looking, outside of the cockpit itself. Such systems are critical because of the potentially fatal consequences of the pilot's removing his eyes from enemy aircraft, even for a moment. Similar but much simpler systems are currently being developed for luxury automobiles, both to display current vehicular speed and, in many cases, to display infrared night-vision video.

Aircraft HUD systems also have the job of projecting virtual images of markers onto actual enemy and friendly aircraft, to help the pilot track and distinguish between them. As with our application, in such a case it is very important that the markers actually appear to be located at the same position in space as the aircraft the pilot is viewing. However, unlike our application, HUD systems are designed to be used by viewers looking at objects in the distance, and so the optics of HUD systems are designed to project virtual images at infinity. Projecting everything to an infinite distance away is acceptable because there is little difference, when viewed through a small aperture (such as the cockpit glass) between an object at infinity (for which the light rays will be parallel) and an object several meters away (for which the light rays will be almost parallel). In such cases, the viewer will use other depth cues (such as scale) to determine the depth of the virtual image.

These HUD systems have the additional design simplification of needing only a relatively narrow field of view. The pilot/driver is seated in a chair at a relatively fixed position relative to the HOE. Accordingly, there is not a large expected range of head movement and so the space of expected eye locations subtends a small angle relative to the HOE.
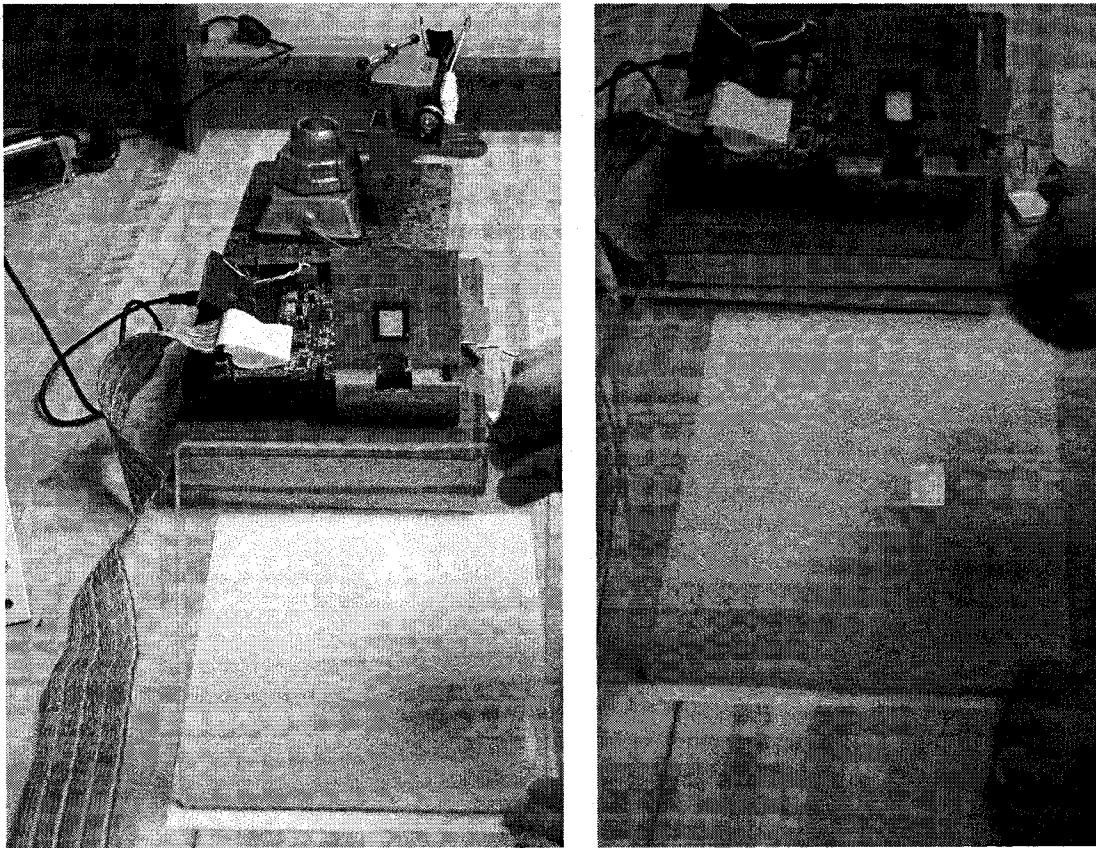
Figure 3.9: *Left:* The complete proof-of-concept demonstration apparatus. A laser shines through a beam-spreading lens, and both are held by the vise at the top of the picture. An LCD is positioned vertically such that it is back-lit by the laser. The VHOE is held by hand in front of the LCD, such that a virtual image of the LCD is visible through the HOE (if viewed in black and white, the virtual image may not be clearly visible). *Right:* A slightly zoomed in view of the proof-of-concept. The resulting virtual image is shown up close in Figure 3.10.

The other major existing real time virtual image projection application of HOEs is to HMDs. Because HMDs generally use a separate optical system for each eye, they can tolerate a very narrow field of view, since the user can adjust the HMD to have the optics properly positioned in front of each eye. HOE-based HMDs are currently used in applications such as military combat for which it is acceptable and even desirable to project the virtual image to infinity.

In summary, there are existing HOE-based real time virtual image projection systems, but they all are designed to project a virtual image to infinity, and they all benefit from a relatively small field of view requirement. While these existing systems demonstrate the potential of HOE-based systems for projecting high quality virtual images, they do not address some of our specific concerns. The RTTH applications discussed at the beginning of this chapter require a focused virtual image in the near field, generally within a few meters, and many of these applications place significant value on achieving a much larger field of view.
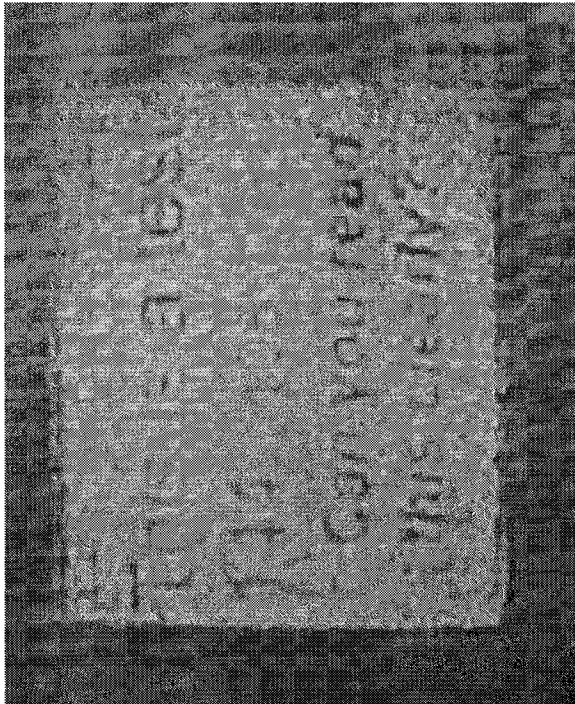
Figure 3.10: A close-up view of the virtual image produced by the demonstration VHOE in Figure 3.9. The pixelization is a result of cropping the previous image, not of the HOE itself. (We had difficulty getting a camera to properly expose when zoomed in on the small HOE; humans however seem to have no problem.)

Chapter 4

Methodology

The original attempts at RTTH resulted from work with Andres Nowatzyk, Ph.D., who first proposed replacing the mirror in RTTR with an HOE [33]. The initial attempts at RTTH were made by Nowatyk, Shelton, and Stetten. They sought to directly record a binary zone plate on an optical table by use of two in-phase spherical wave fronts (see Figure 3.8). They had great difficulty getting proper exposure of the zone plate, however, and so the process was abandoned in favor of computer-based design and outsourced manufacturing of the HOE.

Modern optical design makes heavy use of computer simulation and optimization [34], but nevertheless it remains very much an art. As with most design processes, optical design begins with the establishment of a basic design specification, including such things as:

- The basic physical shape and form of the optical system

- The key desired properties of the optical system

- Other preferences for the optical system.

The design specification is used to create both a computer-adjustable model of the optical elements in the system and a *merit function* that quantifies the "goodness" of any variation of that model. The merit function is based on quantities derived from simulation (typically ray tracing, see section 4.2) of the model, and the merit function is used to guide an optimization procedure that adjusts the model in an effort to improve its merit function score.

It is usually insufficient to perform a single optimization run, and so the final adjusted model of one optimization is typically used as a starting point for a new optimization run. The new optimization run typically uses a manually updated merit function and/or a manually updated model with new or different adjustable parameters. A final design is usually arrived at only after many such iterations.

Much of the art of of optical design lies in three key areas:

- Model choice, which specifies the optimization search space

- Merit function creation, which guides the optimization process

- Manual tuning of the model and/or the merit function between optimization runs

36

Chapter 5 will discuss the specifics of how the above design process was followed. The remainder of this chapter will cover the basics of optical modeling and the processes of computer-based simulation and optimization of optical systems.

4.1 Optics modeling

One must be able to model an optical element in order to simulate, optimize, design, and manufacture it. The optical model is:

- Simulated to evaluate the merit function.

- Adjusted accordingly by the optimization algorithm.

- The final result of an optical design.

- Used as the basis for manufacturing.

Accordingly, an ideal model is:

- Easy to simulate.

- Succinct, with minimal redundancy between its adjustable parameters.

- Sufficiently flexible[1] to represent an acceptable design solution.

- Straightforward to manufacture.

As a relevant example, three commonly used models for a simple, two-surface optical lens are:

- The *thin lens* model

- the *paraxial* model

- the full *geometrical* model

The thin lens model has only one parameter: focal length [35]. The paraxial approximation model has only six parameters (known as the cardinal points) [35]: two focal points, two principal points, and two nodal points). The full geometrical model represents a lens as two geometric surfaces on a specific type of glass (the glass is often modeled by its refractive index for each analyzed wavelength), and accordingly the full geometrical model can be accurately simulated using full ray tracing according to Snell's law [35].

The thin lens model is trivially easy to simulate (even with pencil and paper), but it does not account for the optical aberrations present with any real lens. The full geometrical model, on the other hand, can fully represent any physically-realizable lens, and it is a precise description for manufacturing. The difficulty with a full geometrical model is that it lacks a small set of intuitive parameters that encapsulate its functionality. Instead, the geometrical description of a surface can be arbitrarily complex, and it is up to the designer to determine the type(s) of geometry to which each surface will be restricted. Examples include:

---

[1]Meaning it has enough adjustable parameters (degrees of freedom).

- Planer, with no parameters.

- Spherical, with one radius parameter.

- Cylindrical, with one radius parameter.

- More complex models, such as a surface described by an eighth-degree polynomial.

Just as there is great freedom in modeling a geometrical lens surface, there is also great freedom in modeling a DOE. Useful DOE models include:

- An algebraic model of the phase function.

- The 3D locations of two construction points.

- The *optical fabrication* model, which includes the 3D locations of two construction points, plus intervening construction optics.

An algebraic model of the phase function is the most generic model, but it can be difficult for a designer to determine what form an algebraic model should take. The algebraic model works well for designing a kinoform, since it is straightforward to sample and quantize any algebraic phase function into a surface-relief pattern to be etched into a kinoform.

The second model above, utilizing two construction points, can only model the zone plate class of holograms (see subsection 3.3.1). This model contains only six parameters to optimize, consisting of three real numbers for each construction point. Some of these parameters may be fixed at zero from a-priori knowledge, further reducing the number of parameters to optimize, such as in an axial system, which would require that both construction points be on-axis, thereby reducing the number of parameters to two. This model is straightforward to manufacture, either optically from two physical construction points, or as a kinoform by precomputing the phase function that an optical recording would produce.

The last model above is capable of representing any optically fabricated hologram. It models an HOE by the optical setup used to photographically record the HOE, and accordingly this model is best suited for optical fabrication. For a simple zone plate recording, this model would be the same as the second model, but this model allows and accounts for extra optics to be placed between one or both of the construction points and the HOE during photographic recording of the HOE. These extra optics are referred to as *construction optics*, since they are only used to construct the HOE, and are not used during HOE playback. Each construction optic must be fully specified as to its position, orientation, and internal optical model. The optical fabrication model can be complex to design and optimize, because there is no a-priori constraint on the number or type of construction optics. It is entirely possible to model (and manufacture) a hologram constructed using multiple other intervening construction holograms, some of which are algebraically modeled kinoforms and others of which are themselves modeled by optical fabrication, allowing recursive complexity. A VHOE is well suited to implementing this model.

4.2 Simulation

Computer simulation of an optical system allows computer optimization of the optical system, because simulation is the means of performance evaluation used by the merit function. While there are

multiple methods of optical simulation, the methods used for this work all fall under the category of basic *ray tracing*.

Figure 4.1 graphically demonstrates ray tracing from one image point on an LCD through an HOE, such that after exiting the HOE, the apparent virtual source of each of the rays can be determined. Each point on the LCD is a source of rays striking the entire surface of the HOE. When all of the rays from a single point on the LCD exit the HOE, they should ideally all appear to have come from a *single* virtual point in the holographic image, so that the virtual point will be in focus. Once simulated, it is up to the merit function to measure the extent to which each virtual point is not in focus.
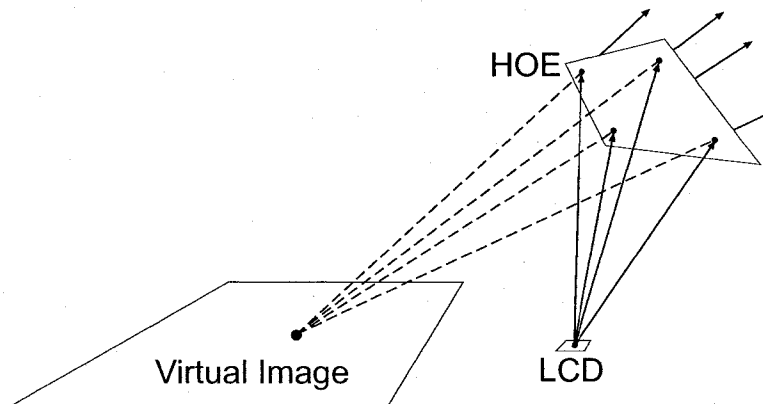
Figure 4.1: A virtual image point is found by ray tracing from one point on the LCD through many points on the HOE.

The process of ray tracing follows the path of one or more light rays (see section 3.2) through an optical system, each light ray beginning at some specified starting point. Typically, multiple starting points are used, and each starting point emits multiple rays in different directions, each of which is independently traced. Ray tracing requires a working model of the path a light ray will take as it interacts with each simulated optical surface. Most such models of light ray paths are simple algebraic equations based on:

- the light ray's associated wavelength

- the ray's angle relative to the surface normal

- properties of the optical surface, such as its curvature and the refractive index on either side of it

Section 3.2 developed such a model for a flat DOE surface based on the DOE's phase function. The model developed there is appropriate for *sequential ray tracing*. Sequential ray tracing does not simulate the multiple exit trajectories a single ray may take when it enters a single surface, and accordingly it does not simulate the fraction of incident intensity present in any of the exiting rays. In practice, there are almost always multiple exit trajectories, but usually there is only one of primary concern to the developer. For example, when a ray strikes a simple glass lens surface, most of the light ray's energy will pass through the surface as a single ray, but some of the incoming

ray's intensity will reflect off of the surface as a separate outgoing ray. Often, the reflecting ray can be ignored because it will contain only a small fraction of the incident ray's intensity and because it will not interfere with the normal operation of the optical system. Likewise, a DOE normally has multiple orders of operation, the first of which was developed in section 3.2^2. The higher orders of operation are similar, but bend rays more due to in-phase geometry that is "off" by integer multiples of λ. The first order of a DOE is typically the one with the greatest percentage of light, and accordingly *diffraction efficiency* is a measure of the percentage of incident light that exits in the first order. Sequential ray tracing was used exclusively for this work, and all DOEs were simulated in their first order of operation. It was up to the manufacturer of our DOE designs to produce them with a high and approximately uniform diffraction efficiency.

Finally, it is also worth mentioning that rays can be traced both forwards and backwards through an optical system. Making reference to Figure 4.1, it is perfectly valid to trace rays starting at a point in the virtual image rather than at a point in the LCD. If tracing from a point in the virtual image through the HOE, the simulation would then need to back-trace through HOE to determine from where on the LCD each ray would actually have come. In a perfectly focused system, just as all the rays from a point on the LCD converge at a point in the virtual image, likewise all the rays from a point in the virtual image back-trace to converge at a single source point in the LCD. When optimizing an actual system, the merit function can measure lack of focus in the system by analyzing either how well the virtual points are focused (for forward tracing) or the source points are focused (for reverse-tracing).

4.3 Optimization

Design optimization is used to iterate continually towards a "better" optical system. There are many optimization methods that may be utilized in optical design, but they all require the designer to supply a merit function (sometimes referred to as an *error metric*) that deduces a single scalar numeric score for the performance of an optical system, based on quantities derived from simulation. Optimization procedures compare the merit scores of differing optical designs to derive a potentially better design; the new design is then simulated, and its merit score is then compared to the previous merit scores to determine the effect of the design changes. The merit function must not only measure whether one system is better or worse than another, but the degree to which it is so. *Accordingly, much of a designer's work lies in crafting a merit function that accurately distills all of the combined design goals, preferences, and constraints into an single representative scalar score for an optical system.*

One of the greatest difficulties in crafting a merit function is in establishing the relationships between each of the desirable properties in terms of the overall optimization. These relationships can be far from simple or linear. For example, how stable must the virtual image be to permit no "appreciable" change across the field of view? What, if any, benefit is derived from further increases in stability beyond this threshold of acceptability? If it is determined that the stability criteria is met if each pixel in the virtual image is never perceived more than 1.0 mm from its appropriate location, then how beneficial is it to slightly sacrifice other desirable properties to reduce the above 1.0 mm down to 0.5 mm?

---

[2]The zeroth order of a DOE is the one in which it has no wavelength dependence, typically doing nothing and appearing transparent as it passes light straight through.

It is usually straightforward to come up with a formula that scores an optical system according to a *single* design goal or preference. How multiple such formulas should be combined, however, is generally a non-linear and non-quadratic function. If each such formula constitutes a single term in the merit function, then most such terms will have some range within which they are acceptable, but when multiple terms are outside of their acceptable ranges it is usually unclear how much weight should be given to each out-of-range term. Weight assignment is important, because it has a direct influence on which properties an optimization algorithm will seek to "satisfy" first, and the degree to which it will seek to keep those properties "satisfied" while optimizing other properties.

It is usually intractable to globally optimize an optical system, and so a designer must work to guide an optimization procedure around and over undesirable local optima. A well-designed merit function can help greatly, but it is still usually necessary to perform multiple optimization runs. Typically, the final adjusted model of one optimization is used as a starting point for a new optimization run. The new optimization run generally makes use of a manually updated merit function and/or a manually updated model with new or different adjustable parameters. By gradually introducing complexity into the model over many optimization runs, and by adjusting the merit function to focus on different design goals as the model becomes more complex, it is often possible to eventually arrive at a suitable (and locally optimal) optical design.

Chapter 5

The Evolution of a Design

The previous chapter presented the design methodology which we have applied in many different ways over time to ultimately arrive at our successful prototype. We progressively revised our intentions for the basic physical shape and form of our system, while holding to our key required optical properties. Our methods of modeling the HOE and evaluating the merit of our optical designs co-evolved along with our simulation and optimization algorithms. This chapter follows that evolution. The journey was somewhat convoluted, buffeted by changing relationships with manufacturers and the development of significant new in-house expertise in optical engineering by the author and two of his advisers, Andreas Nowatzyk and Mel Siegel.

5.1 Basic physical layout

5.1.1 Early layout with a reflective HOE

An early proposed layout of the RTTH optics for the SF is depicted in Figure 3.1. Note that this layout uses a reflective HOE and is very similar to the layout of the conventional Sonic Flashlight, except that the HOE is larger than the mirror it replaces, so as to accommodate a larger virtual image, while the LCD generating the display image is much smaller than the virtual image produced (25 mm diagonal LCD vs. approximately 280 mm diagonal bounding-box around the virtual image). Early on it become apparent that the increased size of the HOE would physically block access to the surface of the region being imaged. Since performing invasive procedures is a primary function of the holographic Sonic Flashlight, the HOE's location was moved toward the viewer to allow sufficient room on the virtual-image side for the use of tools, clinical or otherwise. Unfortunately, moving the HOE toward the viewer makes the Sonic Flashlight much longer and therefore potentially more difficult to manipulate. Moving the HOE closer to the user does, however, cause the HOE to cover a larger part of the user's field-of-view, thereby allowing for a smaller HOE to achieve the same field of view. A smaller HOE not only reduces the cost of the device, but also makes the device less unwieldy.

5.1.2 Early layout with a transmissive HOE

Because the primary advantage of RTTH over RTTR is that RTTH can produce a larger virtual image without requiring an unwieldy device, it is important to avoid having an RTTH system un-

42

necessarily long. To compensate for the increase in length inherent with moving the HOE toward the viewer, the LCD was repositioned to the virtual-image side of the HOE, thereby reducing the length of the device. Such repositioning requires the use of a transmissive instead of a reflective HOE. Furthermore, the LCD must be positioned out of the direct line of sight, to maintain an unobstructed direct view of the object being imaged. Positioning the LCD out of the line of sight can be more difficult to achieve when using a transmissive HOE, but it is possible by using the HOE to bend the light rays as a diffraction grating. Unfortunately, tilting the HOE and moving the LCD to an acceptable position on the virtual-image side of the device prevents the Sonic Flashlight from having a single-optical axis, which would have been possible with the original, simpler design in Figure 3.1. We initially sought to mitigate this added design challenge by allowing the LCD to be tilted as well.

Given that the HOE had to be tilted and moved closer to the viewer, David Wang (an MD/PhD student involved in the clinical trials of the Sonic Flashlight) was consulted to help determine a desirable layout for performing ultrasound-guided invasive procedures with an RTTH device. David used foam-board to make a physical mock-up of the HOE relative to the virtual image. His basic findings were that to provide the best view of the virtual image, the HOE should be elevated above the plane of the virtual image, and tilted to an orientation orthogonal to the user's expected line-of-sight (see Figure 5.1).
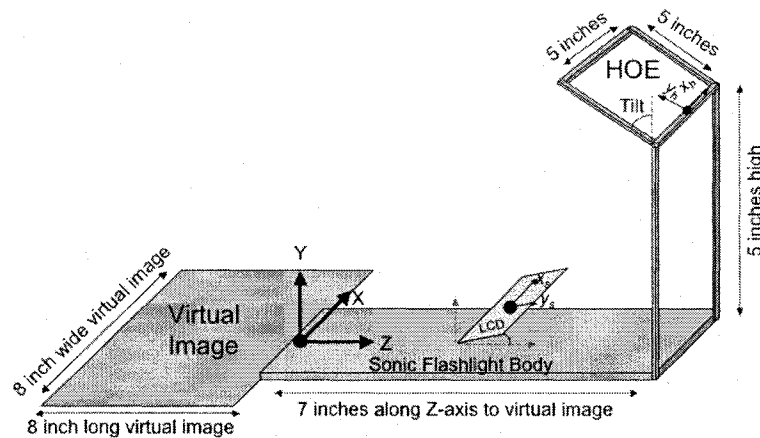
Figure 5.1: Transmissive HOE layout for ultrasound guidance (most dimensions were determined empirically)

Unfortunately, as discussed later, we were not successfully able to find an HOE design appropriate for the layout depicted in Figure 5.1, due in significant part to its lack of a single optical axis. Our next design step was to allow a range of angles for the HOE and to allow for additional optics between the LCD and the HOE. Permissible additional optics included lens(es), curved mirror(s), additional HOE(s), etc. Through this design stage we deliberately remained abstract as to other details of a complete RTTH system layout, including:

- What type and form-factor of scanner (e.g. ultrasound) would actually be used?

- How would the laser-light be spread to illuminate the entire LCD?

- How would the laser-light be diffused, either before or after the LCD, so that each LCD pixel illuminated the entire HOE?

5.1.3 Final layout with a transmissive HOE

Shortly after we started exploring designs with extra optics between the LCD and the HOE, we took a design-turn when we made contact with WP and Richard Rallison. We decided that our layout was too ambitious for a first RTTH system, and based on Rallison's experience we modified our layout to something presumably easier to design, while still preserving the key properties of a tilted, transmissive HOE with the LCD located out of the line of sight. In particular, we moved the virtual image further away from the HOE, to a distance of 1 meter, and we changed the orientation of the virtual image to be orthogonal to the user's line-of-sight, as depicted in Figure 5.2. We also decided to include a fiber optic faceplate in the design, to help produce a flat virtual image (reduce field curvature) by changing the apparent surface shape of the LCD image.
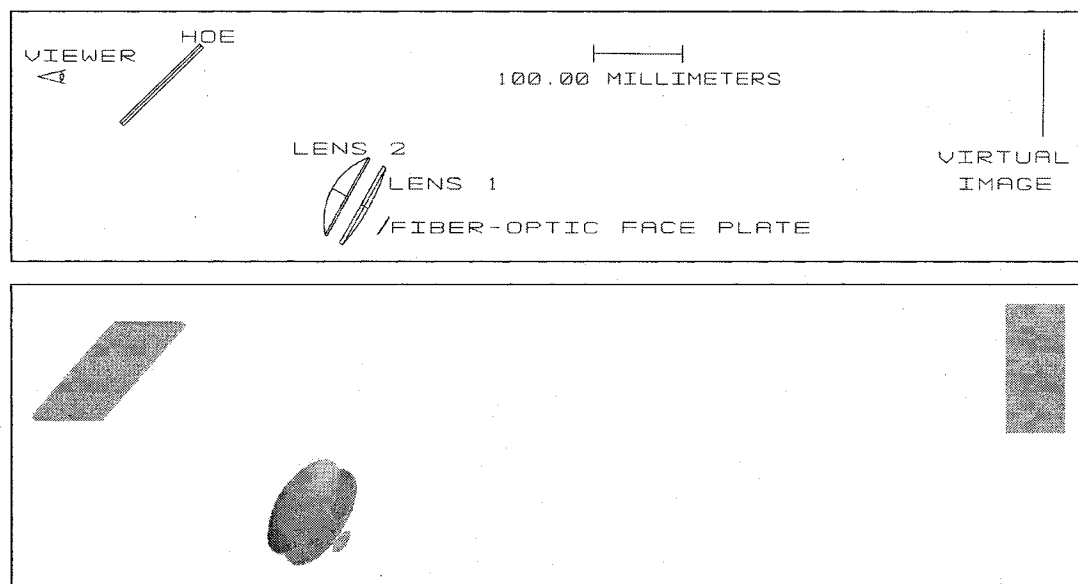
Figure 5.2: Final layout of the RTTH system: side view (top) and angled view (bottom)

A fiber optic faceplate is an optical device made of parallel, fused fiber optics, as depicted in Figure 5.3. A faceplate can be ground and polished to a curved surface, like a lens, but unlike a lens a faceplate produces a source image "floating" on its curved surface that duplicates the real image applied to its flat surface. Thus, if a fiber optic faceplate is placed flat-side against a flat LCD, it will re-create the LCD image, mapped onto the faceplate's curved surface, effectively producing a non-planar LCD, as shown on the right in Figure 5.3. Each of the fibers in a faceplate diffuses the light it emits, which is why a faceplate can turn one real image into another (real) source image.

Since this was our final basic layout for the present work, it eventually became necessary to address the design questions listed at the end of the previous section. Ultimately, we decided that the scanner should be simulated, that a beam expander would be used to spread the laser beam to a diameter larger than that of the LCD, and that the fiber optic faceplate would function as the primary means of diffusing the laser light to the entire surface of the HOE.
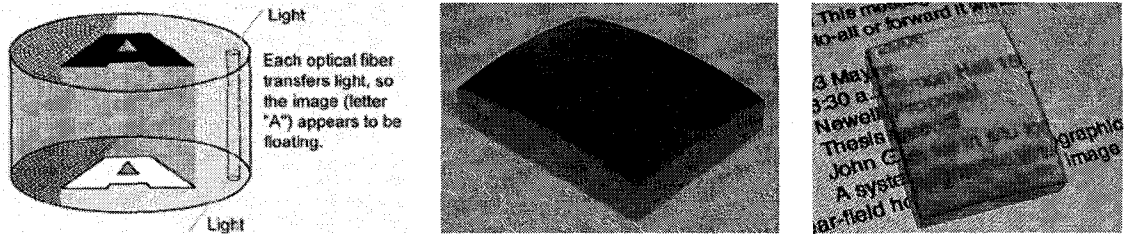
Figure 5.3: Fiber optic faceplate diagram (left), photo (center), and in operation (right)

5.2 Required optical properties and preferences

Implementation of an RTTH system is dependent upon the design of an HOE optimized according to a different set of criteria than those used for previous real time holographic systems. In particular, the following design goals had to be met:

- The HOE must produce a **near field** virtual image.

- The virtual image must be very **stable**. There must be no appreciable change in the virtual image (including its apparent location) across the entire range of viewpoints.

- The **field of view** over a significant range of viewpoints must be large enough to allow stereoscopic human viewing.

In addition, there were other design preferences:

- **Focus/sharpness** should be maximized.

- **Virtual-image curvature** (*field curvature*) should, for most RTTH purposes, be minimized.

- The virtual image **size, shape**, and **position** should match that of the tomographic scan as much as possible, with the caveat that it is much worse to not cover part of the scan than it is to extend beyond the edge of the scan.[1]

At this point it is worth pointing out that virtual image stability and focus are actually measures of the same thing, just perceived over different scales. *Virtual image stability is just focus measured across the entire field of view at once.* If all of the rays from a single point on the LCD diverge from exactly one point in the virtual image (a *virtual point*), then that point will appear both perfectly focused and perfectly stable across viewing positions. On the other hand, if a portion of those same rays passing through the center of the HOE diverge from a very small area around their ideal virtual point location, but a different portion of those same rays passing through the side of the HOE diverge from another very small area centered around a different (typically close) location, then that virtual point will still appear well-focused to an observer but it will not be stable.

Finally, it was decided that all of our designs should operate at the $\lambda = 532$ nm wavelength, because 532 nm is the standard wavelength of a green solid state laser (especially of class IIIa green

---

[1]Since for the present work it was ultimately decided that the scanner should be hypothetical and thus simulated, this last criteria was ignored, but it must be considered when designing a complete RTTH system.

lasers, the type used for "safe" laser pointers). We specified the use of a green wavelength due to the human eye's peak sensitivity to green light, providing the greatest perceived brightness within safe light intensity levels.

5.3 The HOE model

Over the course of approximately five years we explored various aspects of practical HOE production, including the various models used to specify an HOE for manufacturing, as presented in this section. Our exploration was highly educational (although at times meandering), and it eventually led to the successful prototype validated in this dissertation.

Before the present dissertation work began, others in our lab attempted to implement an RTTH system using a BZP (refer subsection 3.3.1 and Figure 3.8), which was modeled (and constructed) using two construction points. As mentioned in Chapter 4, however, they had great difficulty getting proper exposure of the zone plate, and so the process was abandoned in favor of outsourced HOE manufacturing.

The present work began by modeling the HOE's phase function algebraically, as a low-order polynomial for production as a kinoform. Subsequently, it was decided that additional optical elements would be necessary between the HOE and the LCD, and so it was hoped that with such additional optics, a simpler two-construction-point HOE model could be used for production as a VHOE. Ultimately, a working system was obtained only when the HOE was modeled (and produced) using additional construction optics, in addition to extra playback optics located between the HOE and the LCD. Refer to section 4.1 for background on these HOE models.

5.3.1 Algebraic phase function model

The author's first contribution to the RTTH project involved modeling the HOE as an algebraic phase function. At this point, we were planning to have JPL produce our HOE as a kinoform, and a kinoform is most naturally modeled directly by the phase function it implements. We chose as the phase function a parametrically defined polynomial which evolved over a number of iterations into the form:

$$
\begin{aligned}
P(x_h, y_h) = \; & p_0 \sqrt{p_1 + r_a^2} + p_2 \sqrt{p_3 + p_4 (y_a) + r_a^2} + \\
& p_5 \times y_a + p_6 \times \tfrac{1}{4} y_a^4 + p_7 \times \tfrac{1}{5} y_a^5 + \\
& p_8 \times \tfrac{1}{3} r_a^3 + p_9 \times \tfrac{1}{4} r_a^4 + p_{10} \times \tfrac{1}{5} r_a^5
\end{aligned}
\tag{5.1}
$$

defined in terms of optical-axis-centered coordinates:

$$
\begin{aligned}
x_a &= x_h \\
y_a &= y_h + c \\
r_a &= \sqrt{x_a^2 + y_a^2}
\end{aligned}
\tag{5.2}
$$

where the optimization process was only allowed to directly adjust each of the parameters p_0 through p_{10} in the phase function. (The c parameter was indirectly and automatically adjusted when the optimization process changed the geometric layout of the optical system.)

There are several advantages to using a smooth, low-order phase function such as Equation 5.1:

- The behavior of the HOE will not change drastically from one point on the HOE to the next. Therefore, a relatively sparse sampling of all the possible rays from a single point on the LCD to all the points on the HOE should be sufficient to determine the general behavior of the HOE as a whole. (We ray-traced from each sampled point of the LCD through 51x51 regularly spaced locations on the HOE to determine their best-fit focal point in the virtual image.) Ray tracing would not be practical without this smoothness property.

- Likewise, the behavior of the HOE will not change drastically from one point on the LCD to the next. Therefore, a relatively sparse sampling of points on the LCD should be sufficient to determine the general behavior of the HOE for any point on the LCD. (We used a 21x21 sampling of the LCD.)

- The order of points in the LCD will be preserved in the virtual image. In other words, we do not need to worry about the virtual focal point of one LCD pixel being out of sequence in the virtual image with respect to its surrounding pixels. This is because we are basically modeling the HOE as a combination of lenses (a simple zone plate for the first line in Equation 5.1, and one lens for each of the remaining terms in that equation), and a lens is bidirectional. Therefore, if the order of points in the LCD were not preserved, there would have to be at least two different points on the LCD that mapped to the same point in the virtual image. However, this would mean that a single point in the virtual image would map to two different points on the LCD, which is physically impossible for a collection of lenses.

- The HOE should not be highly sensitive to over-fitting. Otherwise, slight misalignment of the LCD during physical assembly could result in significant degradation in virtual-image quality.

- The lack of odd-order powers of x_a preserves the "left-to-right" symmetry of the HOE, since the even powers of x_a are by nature symmetric, and radial basis functions are as well.

5.3.2 Two construction point model

Once it was decided that additional optical elements would be necessary between the HOE and the LCD, attempts were made to use the simple two construction point model of an HOE, since it was hoped that the additional optics would reduce the minimally-sufficient complexity of the HOE. The two construction point model could also be easily produced as a VHOE by WP.

5.3.3 The optical fabrication model

Ultimately, a working system was obtained only when the HOE was modeled (and produced) using additional construction optics, while additionally using extra playback optics located between the HOE and the LCD. It proved sufficient to model two construction lenses, each with two spherical surfaces, as shown in Figure 5.4. One construction lens was located between each construction point and the HOE. Furthermore, once a workable design was reached using this model, it was then possible to use a series of optimization runs to reduce the model to a single construction lens while preserving the optical performance of the system. The final construction layout is shown in Figure 5.22.
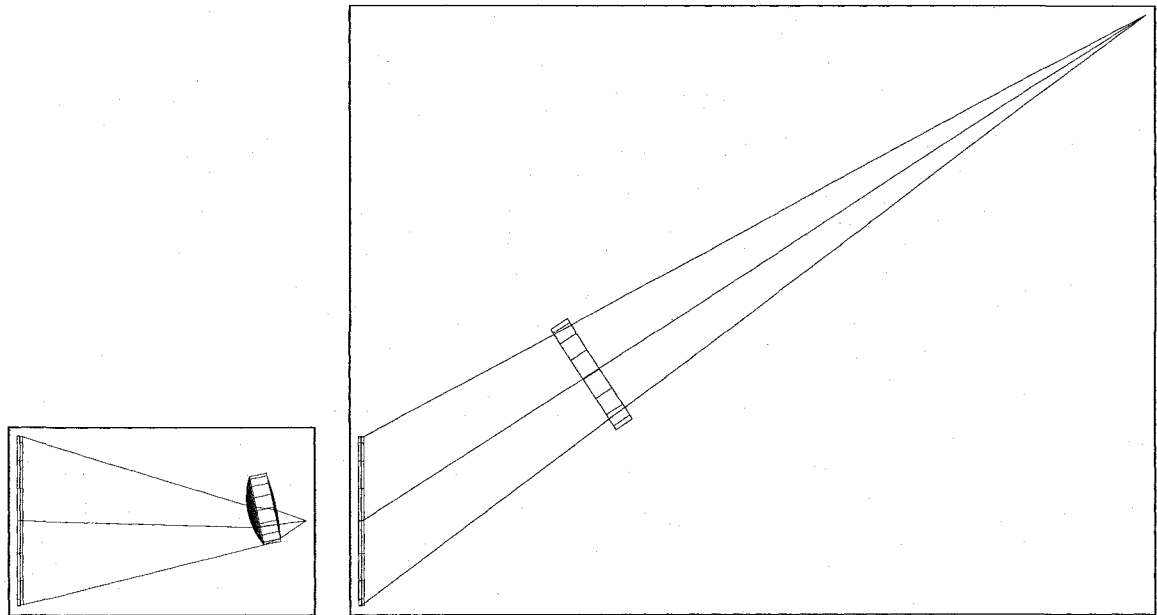
Figure 5.4: In this example of an optically fabricated HOE design, the light-rays diverge from two construction points, passing through each construction point's construction optic before striking the surface of the HOE.

5.4 The merit function (error metric)

The choice of error metric to evaluate virtual image quality was not simple. Using the traditional mean squared error (MSE) between the simulated and "ideal" paths of rays does not work well for the optimization of our HOE because it highly penalizes even small planar geometric distortions (PGDs), which are very difficult to avoid. In fact, we can easily tolerate small or even moderate PGDs because we can compensate by pre-warping the image displayed on the LCD[2]. Therefore, as is typical with modern optical design, we have defined our own error metric consisting of several terms.

Each of the design goals listed in section 5.2 introduces a corresponding error metric with a hard upper limit that must be optimized when designing the HOE. In addition, the design preferences listed in that section should be optimized as well.

As already mentioned, a particularly noticeable error would be lack of focus, otherwise known as *cross talk*. As previously discussed, a perfectly focused virtual image will be a stable virtual image, and so *minimizing cross talk maximizes both focus*[3] *and image stability (which is just a measure of focus over a pupil aperture large enough to encompass all viewpoints)*. Accordingly, we have an error term based on the root mean squared (RMS) focal error across all the simulated image

---

[2]Pre-warping the LCD image to compensate for PGDs increases the effective resolution of the LCD in areas of the virtual image that are "contracted," but it has the disadvantage of reducing the effective resolution of the LCD in areas of the virtual image that are "stretched." Accordingly, an RTTH system needs to have an LCD with a resolution high enough to still be adequate when mapped to the most "stretched" areas of the virtual image.

[3]Technically, it maximizes the average focus taken by measuring focus over any given pupil aperture from every viewpoint.

points. To measure RMS error, it is necessary to measure how close each individual ray comes to its proper focal point.

Determining the proper focal point for a ray can be tricky, and different methods were used depending on how the system was ray-traced. For the initial work using our custom simulation software, rays were traced forward from the LCD, through the HOE, and then back-projected toward their virtual image point. With this method of ray tracing, the proper location of a virtual point was defined (computed) as the one that minimized the RMS distance between that point and each of the rays that should focus at that point. Later, using commercial simulation software, rays were traced in reverse from a set of a-priori specified virtual points, through the HOE, to the LCD. Using this later method, the front of the LCD (or its fiber optic faceplate) provided a fixed surface onto which the rays should focus, and so each ray had a corresponding point at which it intersected that surface. The 2D centroid of these intersection points defined the proper location on that surface for the particular virtual point.

Errors in the position, orientation, size, and shape of the virtual image also needed to be minimized. (The required property of being near field was measured by position errors, and the desired lack of field curvature was measured by shape errors.) These criteria were enforced in two very different ways, depending on how the system was ray-traced. With forward tracing (as used initially), the convex hull of the virtual-image focus points established the basic position, size, and shape of the virtual image, and field curvature was measured using the distance from each virtual-image focus point to the desired plane of the virtual image, which in Figure 5.1 would be the y=0 plane. When we switched to backwards tracing, the virtual image was fully specified with regard to its size and shape prior to any ray tracing, as part of the system's basic geometry, and error metric terms penalized system geometries in which the virtual image position and orientation did not lie within a-priori established bounds. When traced backwards, error terms also measured the approximate degree to which the rays covered (without extending past) the active area of the LCD. Also, with backwards tracing, any field curvature of the system would manifest itself by making rays out-of-focus relative to the surface of the LCD (or its faceplate). Accordingly, minimizing the backwards-traced focal error would automatically minimize the field curvature, and so field curvature was not measured directly.

The field of view of a system is determined primarily by the HOE's size (which was fixed for all optimization runs) coupled with its position and orientation relative to the virtual image (which were allowed to vary). The field of view would become further restricted if every point in the virtual image was not visible (with appropriate viewpoint) through any point on the surface of the HOE (a condition known as *vignetting*). Accordingly, the merit function also contained terms to penalize the geometry of the system if the position and orientation of the HOE relative to the virtual image did not lie within an a-priori established finite range, and the merit function also contained a term to penalize vignetting.

Various other error metrics were also used to maintain physically-realizable and usable system geometry. For example, one error term penalized system layouts in which either the LCD or its associated optics were (partially) located within the system's line of sight from the HOE to the virtual image. Other terms maintained reasonable distances between the LCD, its associated optics, and the HOE. In practice, many such error terms relating to the geometry of the system were ultimately required.

5.5 Simulation and optimization methods

Now that we have discussed how the HOE was modeled for simulation, how the merit function was formulated, and how both of these changed over time, it remains to discuss the evolution of our simulation and optimization methods. We began by using custom implementations of both ray tracing and simulated annealing. We later switched to using the various proprietary simulation and optimization methods in the standard edition (SE) version of the commercial *Zemax* optical-engineering program. Finally, we upgraded to the engineering edition (EE) version of Zemax to make use of its advanced capability to simulate and optimize optically fabricated holograms.

5.5.1 Simulated annealing

The original attempts with RTTH resulted from work with Andreas Nowatzyk, who first proposed the use of the HOE [33]. Andreas and the author together wrote custom ray tracing software that uses simulated annealing to optimize the phase function of a kinoform. Simulated annealing [36, 37, 38, 39] is a well-understood optimization process that is fairly robust even in the presence of multiple extrema in the merit function. Such robustness is important since a change to any part of the HOE will affect how well each of the pixels in the LCD is focused in the virtual image, creating a very complex optimization space.

Our original intent was for JPL to produce our completed kinoform design. JPL also provided some basic guidance as to how we should proceed. We had originally planned on modeling and simulating the kinoform as a bitmap image representing the phase function that would be physically recorded onto a kinoform. JPL advised, however, that we should restrict ourselves to an algebraic representation of the phase function, as described in subsection 5.3.1. Although simulated annealing has considerable power, it is not realistically capable of adjusting the potentially billions of parameters that would be necessary to optimize an almost arbitrary phase function. Therefore, it is necessary to restrict the phase-function search-space to smooth and continuous (i.e. "nice") functions, such as parametrically defined polynomials. Thus we converged on the phase function given in Equations 5.1-5.2, in subsection 5.3.1. The advantages of using such a smooth, low-order phase function have already been presented.

The optimization process was allowed to adjust each of the parameters in the phase function (p_0 through p_{10} in Equation 5.1), along with some of the parameters defining the geometry of the system. If there had been a term in this phase function that increased the error metric for all combinations of its internal parameter values, then simulated annealing should have set its weight to approximately zero. All of the terms use optical-axis coordinates, (x_a, y_a), which differ from the HOE coordinates, (x_h, y_h), in the location of the origin along the y_h-axis. HOE coordinates have their origin fixed at the bottom center of the HOE as shown in Figure 5.1, whereas optical axis coordinates have their origin at the point where the optical axis from the center of the LCD to the plane of the HOE perpendicularly intersects the plane of the HOE. Note that this is not based on the optical axis to the virtual image, and there is no constraint that the LCD-based optical axis origin actually be located anywhere physically on the HOE, but rather only in the plane of the HOE. Some of the terms in Equation 5.1 also use a radial basis. That is, they are based on the radius of a polar coordinate system (r, $angle$) centered at the origin of the optical axis coordinate system. These terms are used for their rotational symmetry about the optical axis, and so the angle of the polar coordinate system is not required.

As mentioned above, there were additional optimization parameters corresponding to certain geometric properties of the layout. Referring to Figure 5.1, these are (1) the tilt of the HOE, (2) the (Y,Z) position of the LCD, and (3) the angle of the LCD. When adjusting these geometric parameters, the optimization procedure constrained the new geometry so as not to place the LCD in the user's line of sight to the object being imaged. Also, every time the geometry changed, the origin of the HOE's optical axis to the LCD was recomputed.

We soon realized that it is optically impossible to produce a single HOE that will work with the modified layout depicted in Figure 5.1. The optimization process was able to reduce all error terms to acceptable values except for the term that penalizes incorrect virtual image area. If the area-based error term was removed, then the optimization process was able to produce a phase function for the HOE that yielded a small but focused and planar virtual image, as shown in Figure 5.5. Partial ray tracings are shown in Figure 5.6 and Figure 5.7. The reason that the optimization process was
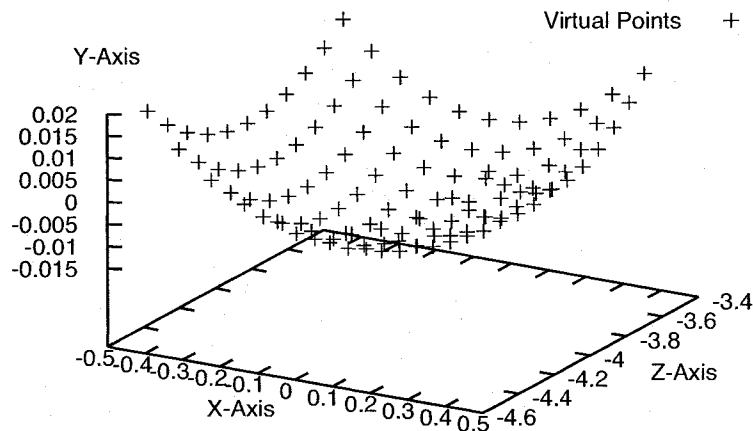

Figure 5.5: Sample points spanning a small virtual image which is centered approximately 100 mm away from the Sonic Flashlight body. All points are within 0.5 mm of the image plane, and the maximum RMS focal error for these virtual image points is approximately 1.5 mm. (The graphed dimensions are in inches.)

unable to produce a large virtual image is that all optical systems, whether based on lenses or an HOE, must follow the basic rule that magnification is equal to the ratio of image distance to object distance (based on the similar triangles from which the well known "lens equation" is commonly derived). Because the LCD must be kept out of the viewer's line-of-sight to the object being imaged, the LCD is constrained to be at least 75 mm (object distance) from the center of the HOE, and even further from the top of the HOE. On the other hand, the center of the HOE is confined within a small range of positions, approximately 300 mm (image distance) from the center of the virtual image, and the bottom of the HOE is only about 180 mm from the near edge of the virtual image. Accordingly, only low magnification powers can be achieved. It is worth pointing out that despite its small magnification, this result would produce a small in situ piece of the desired large virtual image, making it of potential use for some applications. However, for the full range of proposed applications we need considerably larger magnification from the LCD to the virtual image.

To achieve higher magnification, given the constraints of our initial geometry, we added a second, smaller HOE to the system, positioned close to the LCD. The small HOE could be placed
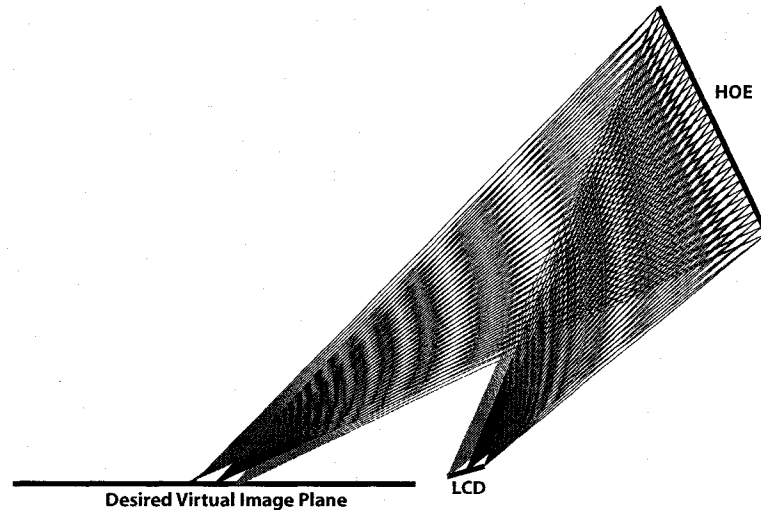
Figure 5.6: Partial ray tracing in the X=0 plane from 3 points on the LCD through the HOE. Rather than tracing the rays forward as they exit the HOE, they are back-traced to determine their apparent sources in the virtual image, as shown in Figure 4.1.
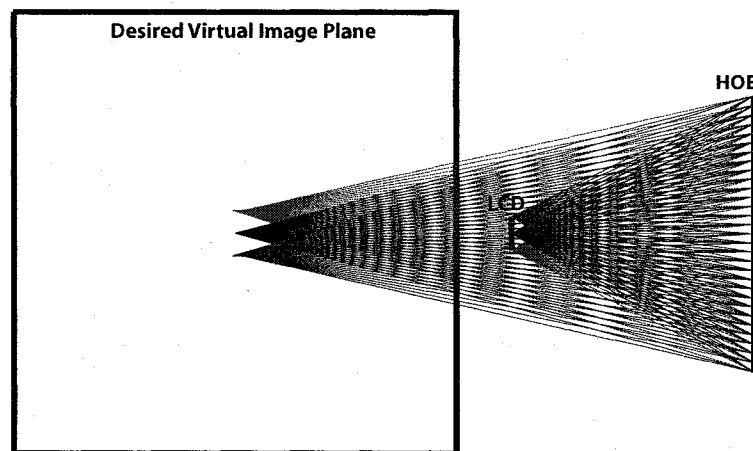
Figure 5.7: Partial ray tracing from 3 points on the x-axis of the LCD through the HOE. Again, rays are back-traced to determine their apparent sources in the virtual image.

arbitrarily close to the LCD, if needed, to help achieve the desired magnification. Such a layout is depicted in Figure 5.8. While it is possible that one or more simple lenses could be used instead of a second, smaller HOE, the second HOE could be smaller than any equivalent lens and would allow more flexibility to reduce possible optical distortions. In our optimization system the position and orientation of the second, "magnifying," HOE was defined relative to the optical axis from the LCD to the primary HOE so that the system retains a single primary optical axis. Therefore, we enhanced the optimization algorithm to recompute the position and orientation (as well as the size) of the small HOE each time the primary optical axis changes.
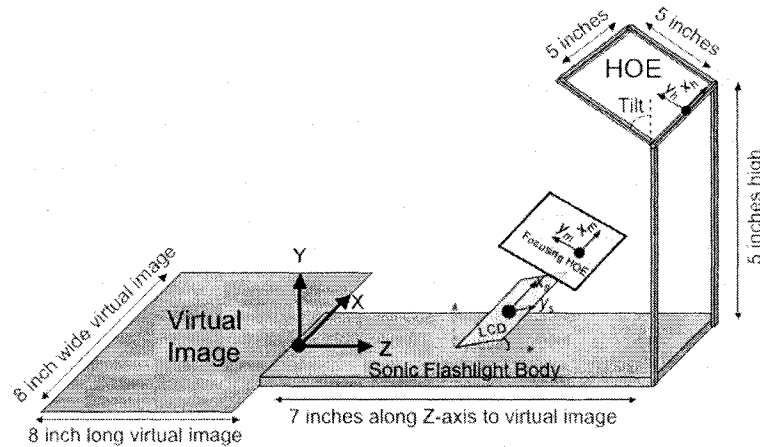
Figure 5.8: Modified layout for ultrasound guidance, with a second HOE close to the LCD to help achieve increased magnification

We did not proceed very far with our optimization of a two-HOE system before unresolvable legal issues between JPL and our universities forced us to find another HOE manufacturer. We ultimately settled on WP, where Richard Rallison helped the author learn how to use Zemax, a much more powerful tool than our in-house simulation and optimization software, and so we abandoned our software in favor of Zemax.

5.5.2 Conventional Zemax design

Zemax is presently the leading commercial optical engineering program. It enables the complete closed-loop optimization of virtually every aspect of complex optical systems containing virtually any type of optical element. When Dr. Rallison first heard of our needs, he provided us with advice, including some very basic initial Zemax designs. These were extremely useful in showing us how to circumvent certain limitations that we previously had thought precluded our using Zemax. One of the most important techniques Rallison showed us was that of tracing the rays in reverse, as discussed in sections 4.2 and 5.4. Tracing in reverse allows one to analyze the rays based upon how they intersect a surface, as is typically done in optics, rather than based upon their apparent focal point in 3D space surrounding the virtual image, which need not lie on any particular surface. He also advised that it would be much easier to create a virtual image further away and roughly orthogonal to the user's line of sight. Rallison's third critical piece of advice was to use fiber optic faceplates (discussed in subsection 5.1.3). By changing the effective surface shape of the LCD, a fiber optic faceplate can help compensate for field curvature distortion in the HOE. In addition, the fiber optic faceplate can function as the diffuser, preventing the additional reduction in optical efficiency (and possibly effective LCD resolution) that would result from using a separate diffuser. Sample screen captures of Rallison's starting-point Zemax designs are depicted in Figures 5.9-5.10.

Both of Rallison's designs provided a rough investigation of the complexity of our problem. His first design contained a paraxially approximated lens and a simple diffraction grating. This design separated the magnification and "off-axis" aspects of our problem. The theoretical aberration-free paraxial lens was very fast (1.3 aperture), indicative of the significant difficulty we would have de-
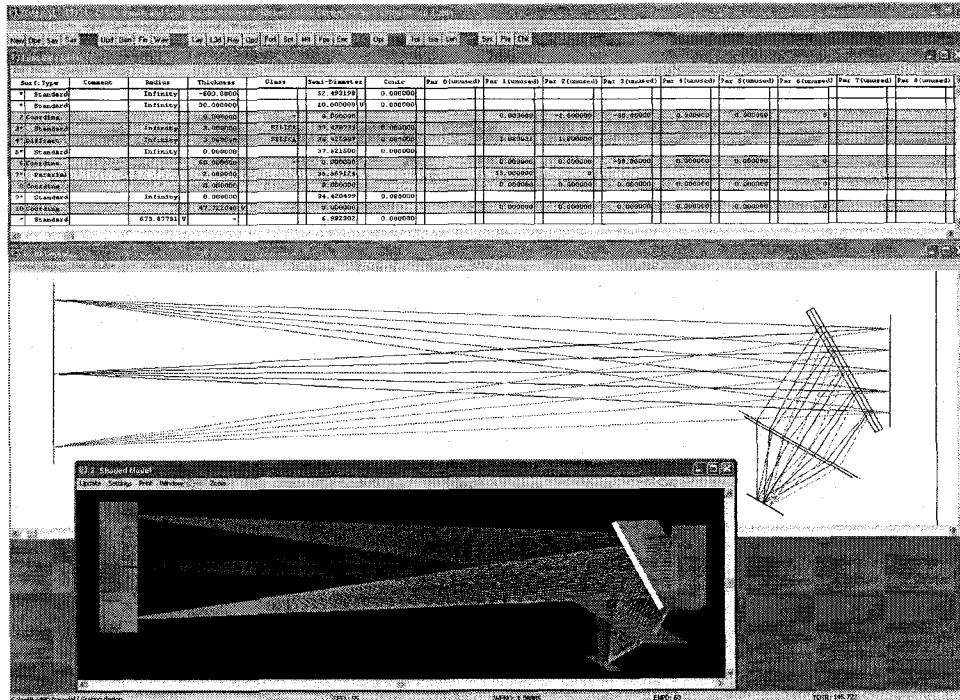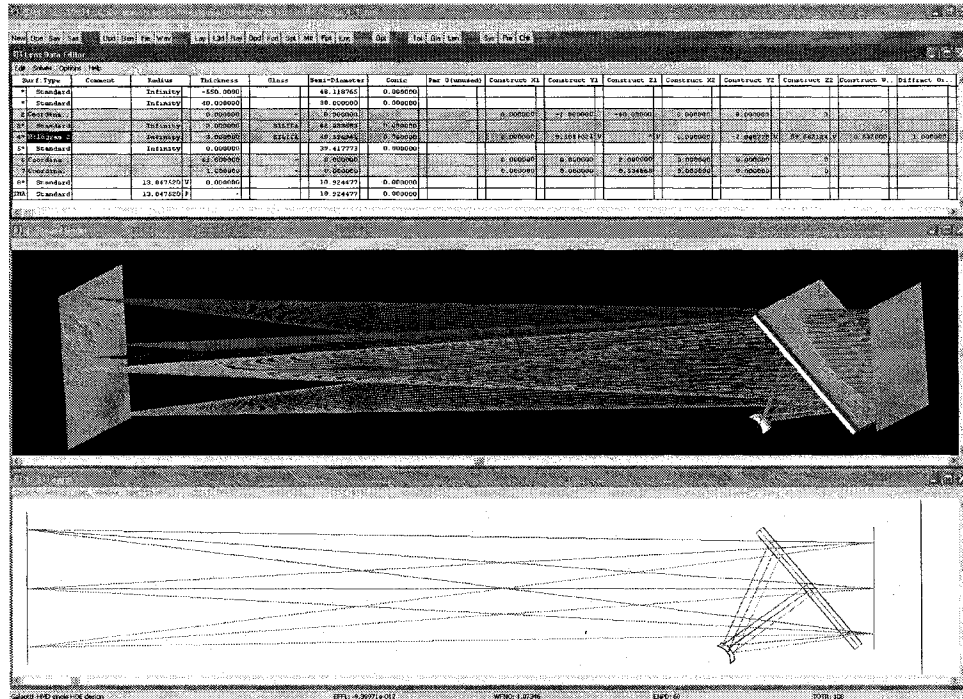
Figure 5.9: One of Rallison's design starting points in Zemax. This particular design utilizes two optical elements, one of which is an HOE and the other of which could be a conventional lens. This particular design is also set up as a paraxial approximation.

riving a suitable design. Furthermore, even with this theoretical aberration-free lens, the system was still not well focused (as can be seen by closely examining the focal points of the rays in the vicinity of the LCD in Figure 5.9). Rallison's other design contained no lenses, but it did contain both a zone plate DOE (modeled as a "Hologram1" surface in Zemax) and a fiber optic faceplate. Since the second design contained no rough approximations, its performance was much more indicative of expected performance. Unfortunately, the second design was very poorly focused, as can be seen by closely examining the focal points of the rays in the vicinity of the faceplate in Figure 5.10.

All of our design work during this relatively brief stage of design was done using the SE version of Zemax, modeling the HOE with Zemax's *Hologram1* and *Hologram2* surfaces, each of which is based only on the 3D locations of two construction points, and hence is only able to model zone plates (see section 4.1). We restricted ourselves in this manner because WP, our intended HOE manufacturer, specialized in optical recording of holograms, and because we hoped the addition of extra optics between the LCD and the HOE would enable us to achieve satisfactory results using this simpler class of hologram.

5.5.3 Optically fabricated hologram Zemax design

It quickly became apparent that, even with additional optics, the zone plate class of HOE was still insufficient for our needs. Rallison advised us that the EE version of Zemax contained a special capability to simulate and optimize arbitrary optically fabricated holograms, modeling an HOE by

Figure 5.10: Another of Rallison's design starting points in Zemax. This design utilizes one HOE and one fiber optic faceplate over the LCD, which effectively changes the shape of the LCD surface. This design is *not* set up as a paraxial approximation.

two construction points, plus intervening construction optics (see section 4.1). Zemax simulates such an HOE by first simulating its optical recording, to determine the final interference pattern of the construction beams, after they pass through all intervening construction optics. Zemax then uses the "recorded" interference pattern to simulate the behavior of the HOE. Because of the complex representation(s), Zemax uses two additional Zemax optical design files (one for each construction point) to model each such optically-fabricated HOE.

A representative example of one of our intermediate designs from this final design phase is jointly presented in Figure 5.4, which shows its construction optical layout, and in Figure 5.11, which shows its playback optical layout. A full field spot diagram analysis of this design is shown in Figure 5.12. This diagram was generated by tracing rays in reverse from 12 color-coded points in the virtual image (92 mm wide and 88 mm tall, located 1000 mm from the HOE) to their intersection with the LCD surface (19.3 mm wide and 15.5 mm tall). If the optical system were perfectly focused, all the points of a given color (i.e., those traced in reverse from a single point in the virtual image) would be coincident. The most significant defocusing optical distortions present in this system are *coma* and *astigmatism*, both of which are naturally problematic for optical systems, such as this one, that contain magnifying, off-axis components. Coma can be thought of as a combined variation in focal position and blur of rays that results when an optic images a point not on its optical axis; in a real image, coma appears as an "ice cream cone" shaped blur. Astigmatism results when, for a given imaged point, the optic functions at a different effective radial and tangental focal lengths; in a real image, astigmatism can appear as an oval, linear, or circular blur. More information

on coma and astigmatism is available in [34] and [35].
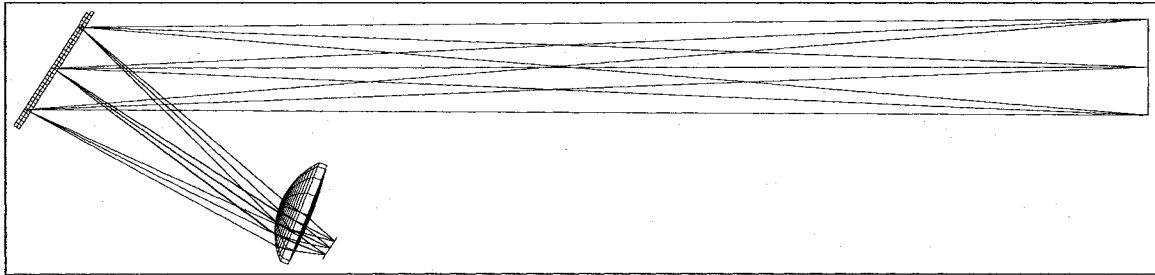
Figure 5.11: In this example optically fabricated HOE design, the HOE operates in conjunction with a custom lens located near the LCD. The geometry of this playback lens was optimized in conjunction with that of the HOE's two construction lenses, shown in Figure 5.4.
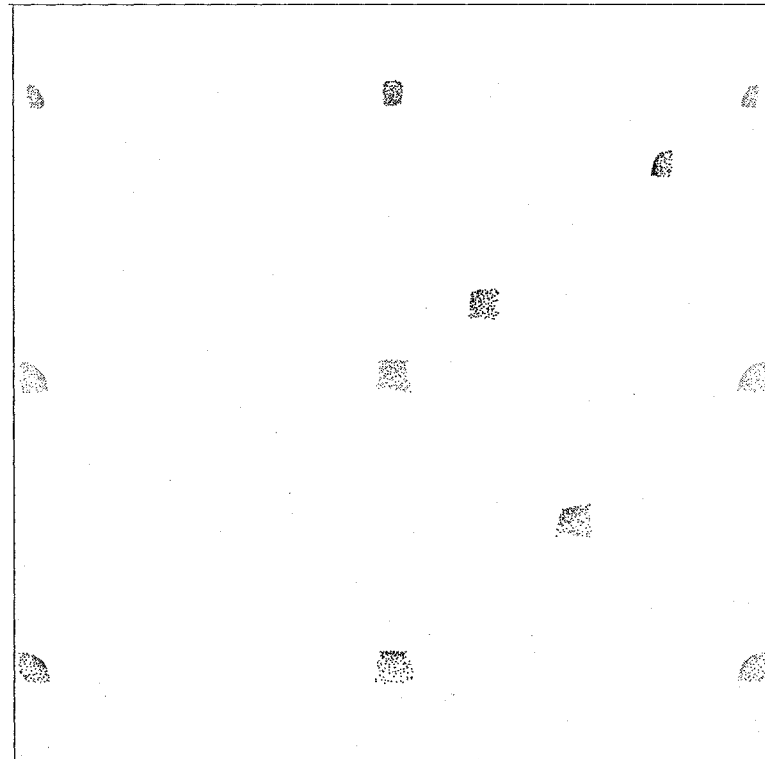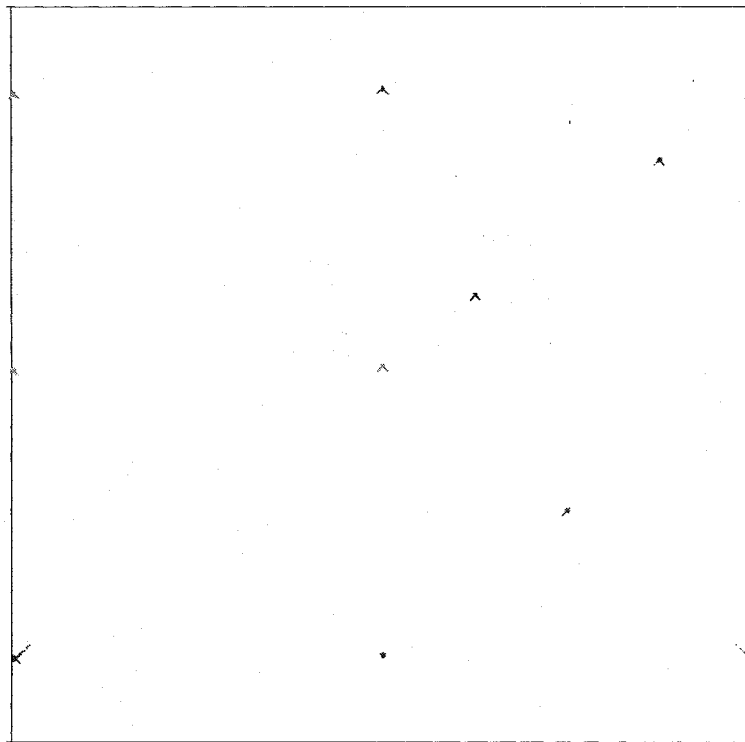

Figure 5.12: Full field spot diagram of the design shown in Figure 5.4 and Figure 5.11. Ideally, all of the points of a given color would be coincident. The square box drawn around the spot diagram is 19.3 mm on each side, equal to the length of the LCD.

5.6 Final steps

5.6.1 First manufactured design

After many optimization runs, including the example above, *Final Design 1* (FD-1) was derived. FD-1 contained only one construction lens, and it simulated well. FD-1's playback optics included two commercially readily available lenses, each with one flat surface, and a custom fiber optic faceplate, which Rallison agreed to make himself. FD-1's spot diagram, shown in Figure 5.13, is much better focused than Figure 5.12, and FD-1's full field spot diagram was generated from a 104 mm wide by 112 mm tall virtual image, larger than that used for Figure 5.12 (the distance to the virtual image and the size of the LCD remained the same). *Accordingly, FD-1 achieved a horizontal magnification factor of 5.39 and a vertical magnification factor of 7.23.* The playback optical layout of FD-1 is shown in Figure 5.14, and FD-1's construction optical layout is shown in Figure 5.15. The most problematic optical distortions present in FD-1 are, again, coma and astigmatism.

Figure 5.13: Full field spot diagram for FD-1. All the points of a given color are much more coincident than in Figure 5.12. The square box drawn around the spot diagram is 19.3 mm on each side, equal to the length of the LCD.

FD-1 was sent to WP, where Rallison examined it and confirmed that the design looked good. WP then attempted to record the VHOE for FD-1. FD-1's VHOE was difficult to manufacture because the more distant construction point (the one without a construction lens) was a *converging point*, rather than a more typical *diverging point*. Exposing the VHOE using a converging point on one side of the VHOE requires focusing a large-diameter beam, such that the beam starts on the
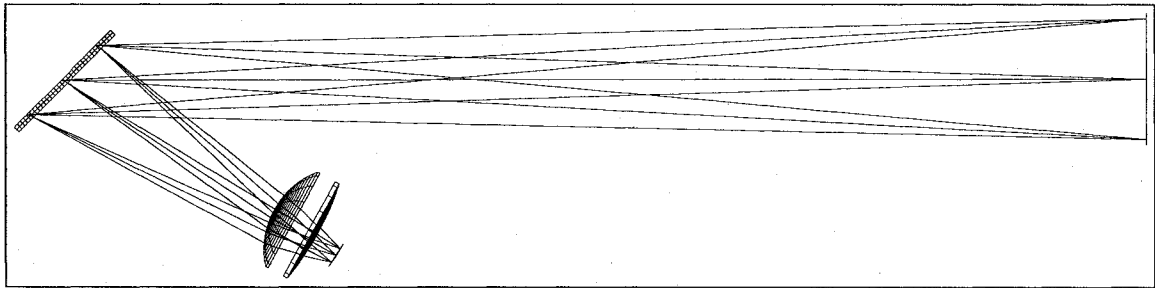
Figure 5.14: Playback optical layout of FD-1. The geometry of the playback lenses were optimized in conjunction with that of the HOE's construction lens and construction points, shown in Figure 5.15
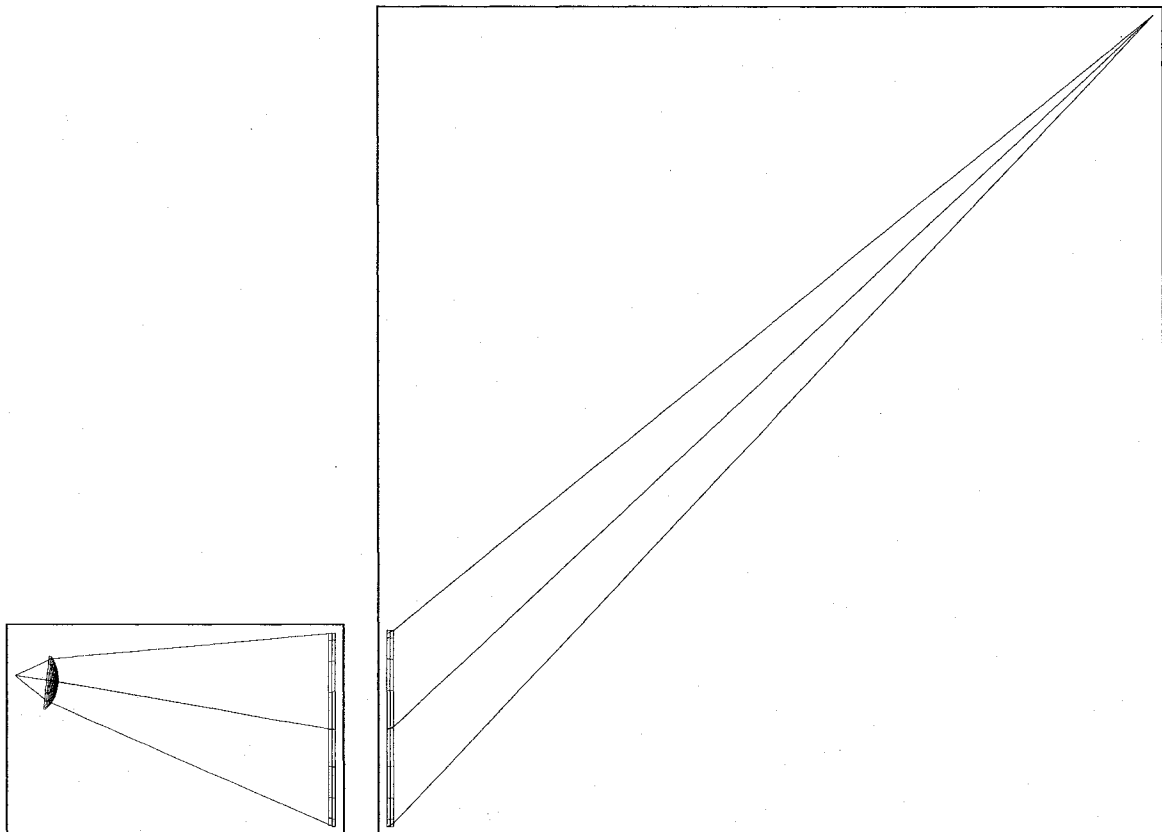
Figure 5.15: Construction optical layout of FD-1. One construction point directly exposes the surface of the HOE, while the light-rays from the other construction point pass through a construction optic before striking the surface of the HOE.

other side of the VHOE, passes through the entire surface of the VHOE, and then focuses at the converging construction point, as shown in Figure 5.16. WP had significant difficulties recording a VHOE with such a converging construction point, trying unsuccessfully to use both conventional lenses and large parabolic mirrors to focus a construction beam through a VHOE to the construction point. WP finally succeeded by first creating a large (203 mm square) axial focusing VHOE, and then using that VHOE to focus the construction beam through FD-1's VHOE. Rallison also ground the fiber optic faceplate for FD-1, which when finished had a slightly different radius of curvature than FD-1 specified, but was not different enough to significantly affect the optical performance of the system.
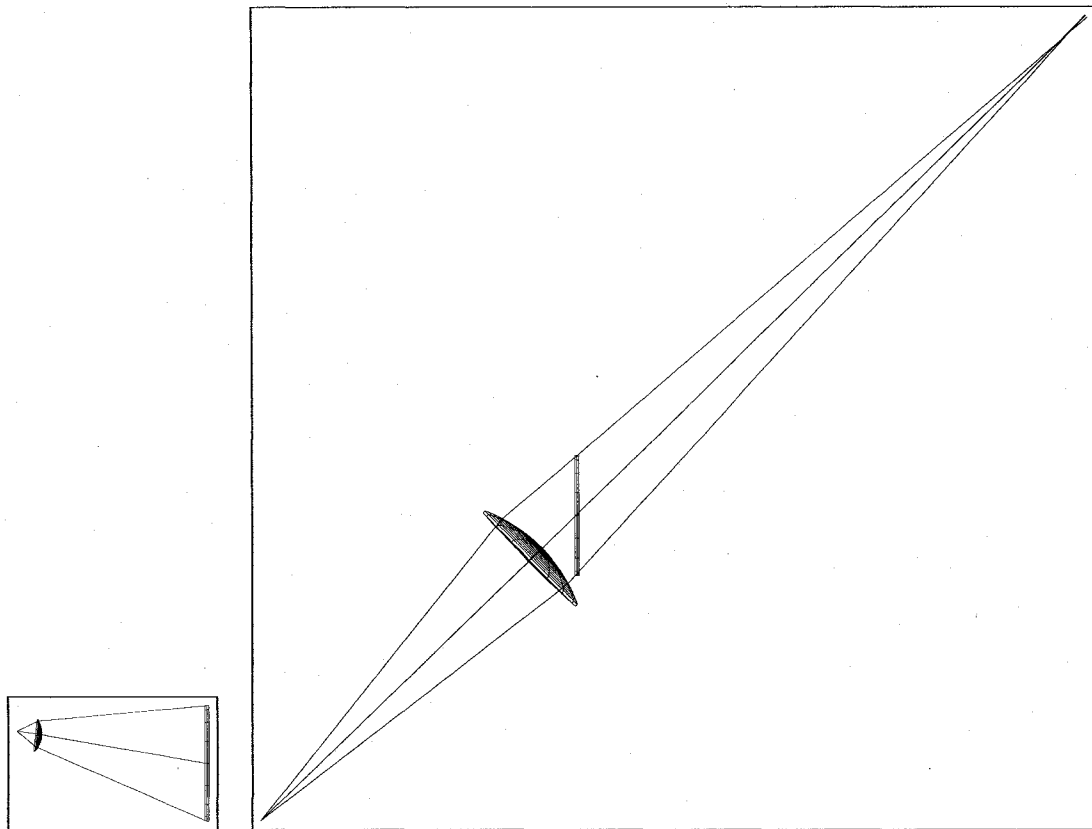
Figure 5.16: Actual construction geometry used to attempt recording of FD-1's VHOE. Note that the converging construction point on the right is achieved by focusing a beam from the other side of the VHOE. In this figure, a lens is used to focus the construction beam. Compare to Figure 5.15

When the VHOE for FD-1 arrived, it did not perform as expected. The author worked with Rallison and Elroy Pearson (of WP) to determine that the design was faulty due to a bug (or bugs) in Zemax's simulation of optically fabricated holograms. Reverse engineering the bug in Zemax, the author was able to achieve simulation results identical to those for FD-1 by mirroring the diverging construction optics to the other side of the HOE, and then making both construction points diverging. This new design is optically different from FD-1, and the two designs should not have simulated

identically. FD-1 was incorrectly simulated by Zemax, as evidenced by the incorrect performance of our VHOE. Apparently, Zemax was disregarding part of the HOE's construction geometry. After careful consideration of the layout and performance of the new design, it was assumed that Zemax was simulating it correctly.

5.6.2 Second manufactured design

Before having WP manufacture a revised design, the author took the opportunity to introduce a few small changes, as follows:

- The design was re-optimized for the exact curvature of the existing fiber optic faceplate manufactured by Rallison.

- The construction angle between the HOE and the longer construction axis (from the HOE's center to the distant construction point) was changed to exactly 45°, at Rallison's suggestion, to improve the efficiency of the final VHOE.

- The type of glass used to simulate the substrate for the VHOE was changed to exactly match the substrate used by WP (glass type B270).

To simplify manufacturing, Rallison then further modified the design by slightly reducing the angle between the two primary optical axes of the playback system, and then re-optimizing according to the author's merit function to derive *Final Design 2* (FD-2).

FD-2's reverse-traced full field spot diagram, shown in Figure 5.17, is virtually identical to Figure 5.13. As with FD-1, the most problematic optical distortions present in FD-2 are coma and astigmatism. Some simple analysis of the best and worst performing image points in the full field spot diagram were used to predict the stability of FD-2's virtual image. The best-focused image point traced in Figure 5.17 is the center point. The image points in the bottom corners of Figure 5.17 are the worst-focused of the 12 image points that were traced; by the symmetry of the optical system, these two image points perform identically. The spot diagram for each of these image points shows, for a single point in space in the virtual image, the locus of points on the faceplate that would be seen from all possible viewpoints. By measuring the horizontal and vertical point spreads across the surface of the faceplate/LCD as shown in Figure 5.18, and then multiplying by FD-2's magnification factors of 5.39 horizontally and 7.23 vertically, it is possible to predict the stability of these points in the virtual image. In Figure 5.18, two measurements per dimension were taken, one containing the vast majority of the traced rays (excluding only some of the rays from the more extreme viewpoints), and the other containing all of the traced rays. Figure 5.19 and Figure 5.20 illustrate the degree to which rays traced through most viewpoints are clustered in the center of their spot diagrams.

As can be seen in the figure, *over most viewpoints the center of the virtual image is predicted to be stable to within 0.8 mm (1.5mm over all viewpoints), and the least stable regions of the virtual image are predicted to be stable to within 1.5mm over most viewpoints (4.2mm over all viewpoints). If the worst-offending corners and/or edges of the HOE were masked, then the the entire system should be stable to within 1.5mm!* For an implementation of FD-2 properly calibrated to its scanner, the greatest anticipated error would then be 0.75mm in any direction.

The playback optical layout of FD-2 is shown in both Figure 5.2 and Figure 5.21, and FD-2's construction optical layout is shown in Figure 5.22 and Figure 5.23.
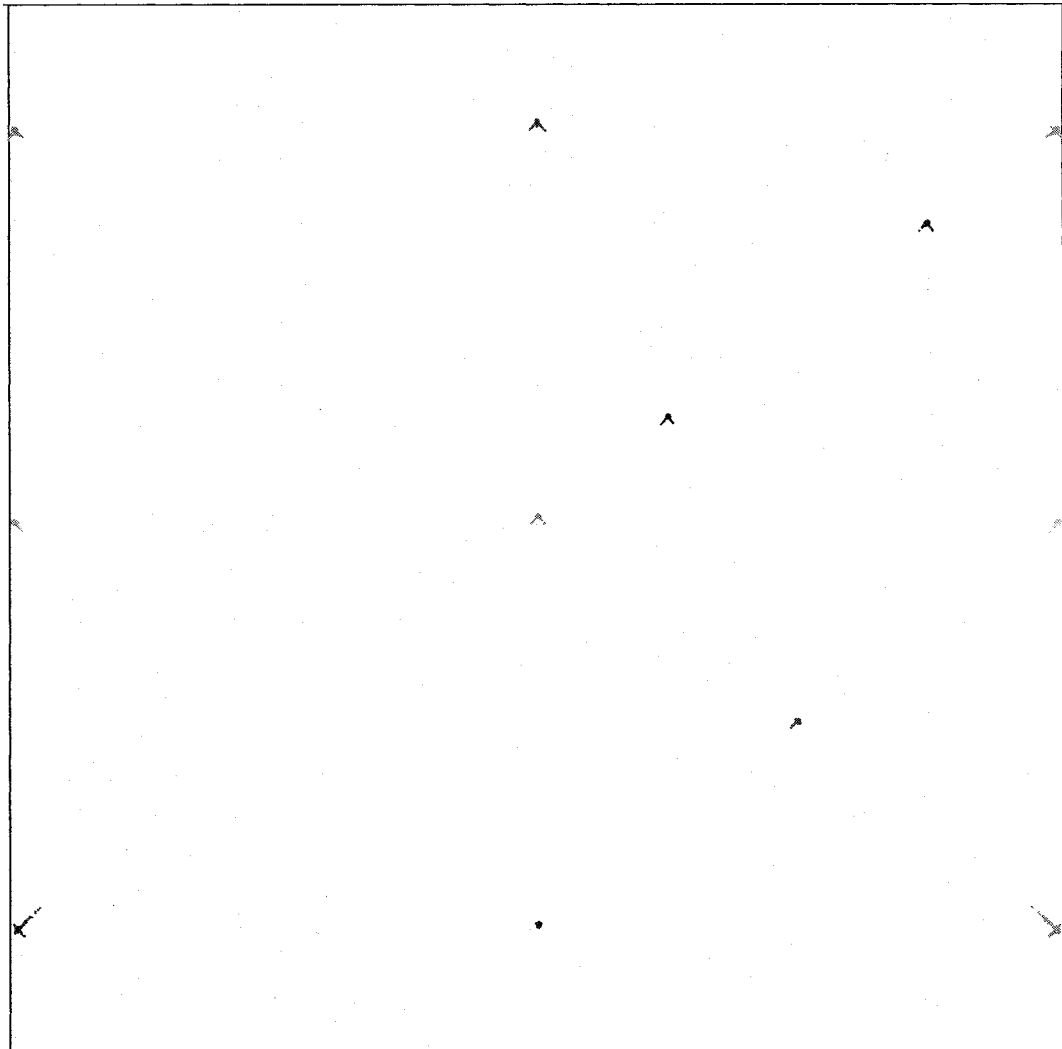
Figure 5.17: Full field spot diagram for FD-2, virtually identical to Figure 5.13. The square box drawn around the spot diagram is 19.3 mm on each side, equal to the length of the LCD. Of the 12 points traced, the center point is the best focused, and the two points in the bottom corners are the worst focused (because of symmetry, their performance is identical).
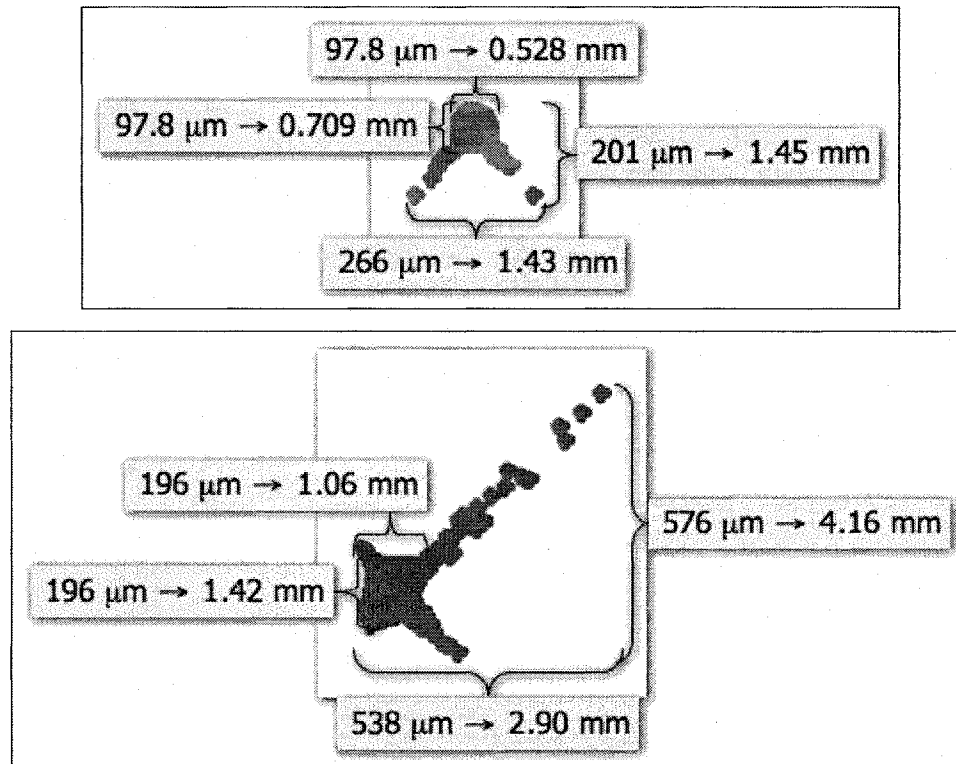
Figure 5.18: Analysis of the best (shown on top) and worst (shown on bottom) image points in FD-2's reverse-traced full field spot diagram (Figure 5.17). By measuring the horizontal and vertical point spreads across the surface of the faceplate, and then multiplying by the optical system's magnification factors of 5.39 horizontally and 7.23 vertically, it is possible to predict the stability of these points in the virtual image. Each ray's intersection point is shown using a relatively large dot to ensure that the most extreme points are easily visible. Most of the dots in this figure are actually very densely spaced and overlapping in the center of each cluster, and so two measurements per dimension were taken, one containing the vast majority of the traced rays, and the other containing all of the traced rays. Figure 5.19 and Figure 5.20 show the same two spot diagrams, rendered using much smaller dots to better reveal the densities of the rays' intersection points.
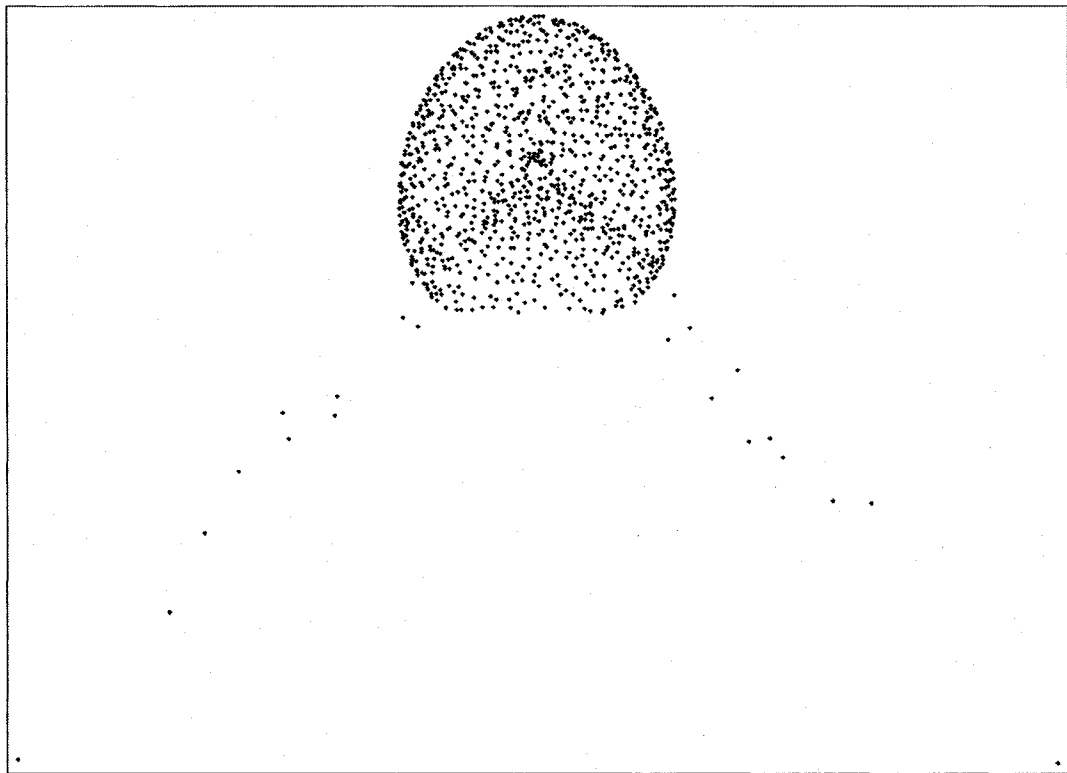
Figure 5.19: Enlarged view of the spot diagram for FD-2's best-performing image point, shown on the left in Figure 5.18. Notice that most of the rays intersect the faceplate in a densely packed area, while a few rays, traced through the most extreme viewpoints/corners of the HOE, extend to the far edges of the figure.
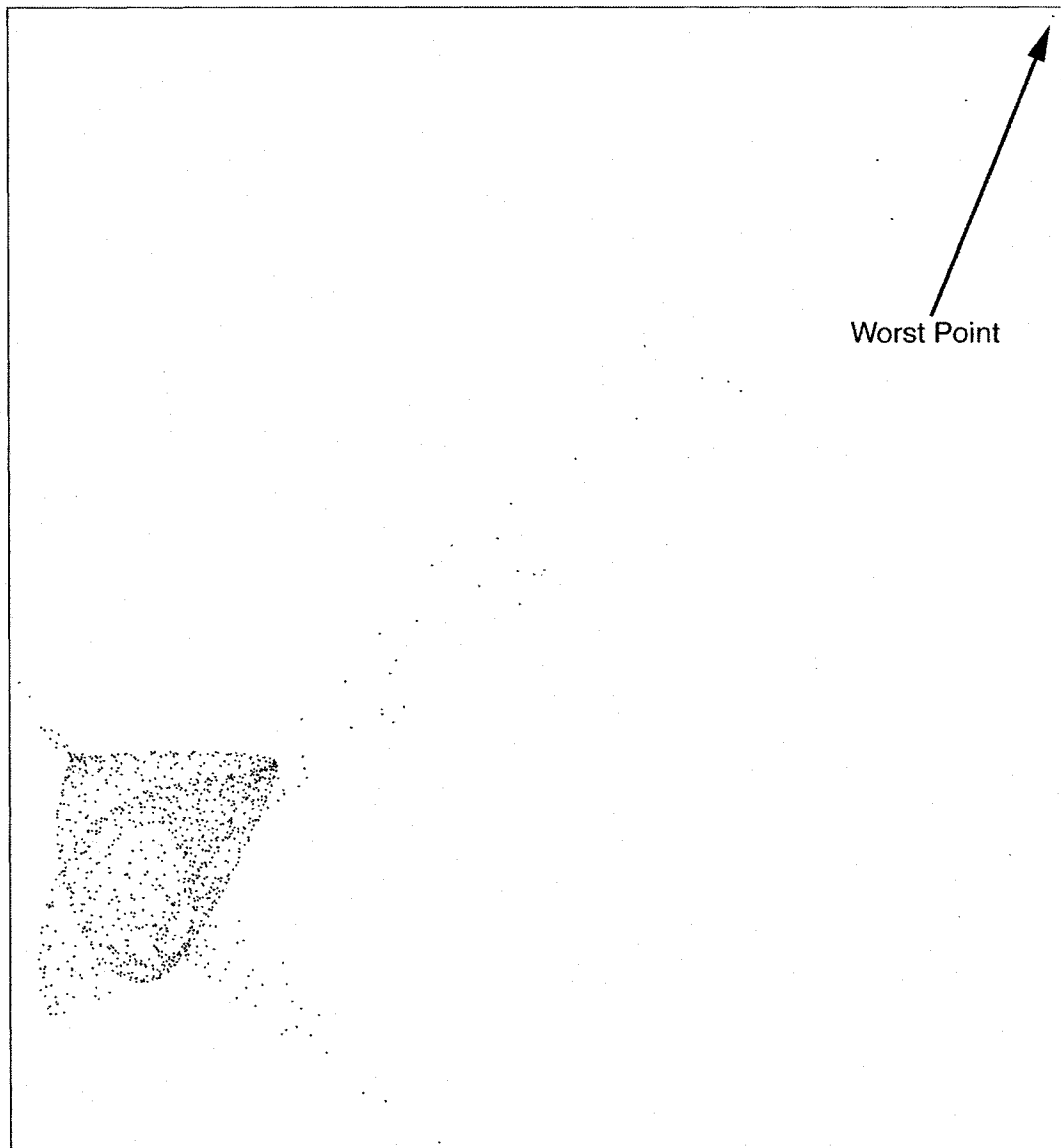
Figure 5.20: Enlarged view of the spot diagram for FD-2's worst-performing image point, shown on the right in Figure 5.18. Notice that most of the rays intersect the faceplate in a densely packed area, while a few rays, traced through the most extreme viewpoints/corners of the HOE, extend to the far edges of the figure.

VIRTUAL IMAGE IS 1M AWAY AT 45.003 DEG FROM HOE VIEWING SURFACE
VIRTUAL IMAGE CENTER IS AT Z = 536.81, Y = 544.08
ANGLE = 33.096 DEG

46.47

ANGLE =
11.907 DEG

16.67

ANGLE =
4.224 DEG

-71.98. 20.8
3.93

-25.63.
7.35

0.0

-300.95. -295.07.
-1.96   -3.20

-298.01
-2.58

ALL DIMENSIONS ARE IN MM.

LARGE LENS: JML PART CPX10345
SMALL LENS: JML PART CPX10495
FIBER OPTIC FACEPLATE: RADIUS OF CURVATURE = -117 MM
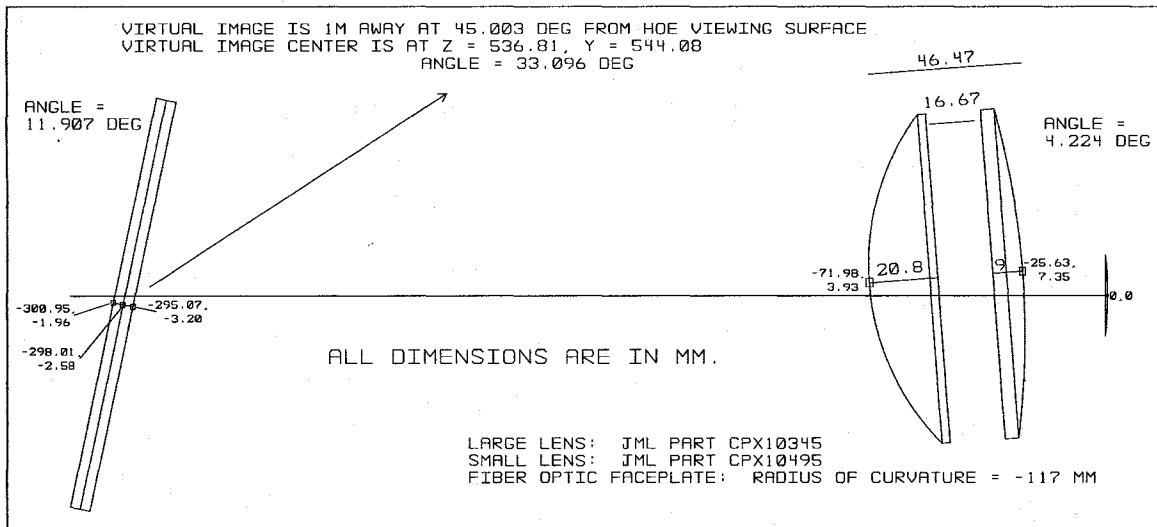
Figure 5.21: Playback optical layout of FD-2. The basic geometry of the playback lenses were optimized in conjunction with that of the HOE's construction lens and construction points, shown in Figure 5.22 and Figure 5.23, but the design was re-optimized to use two very similar commercially off-the-shelf available lenses from JML, parts CPX10345 and CPX10495, which are BK7 PCX lenses with respective radii of 77.2 mm and 257.5 mm.
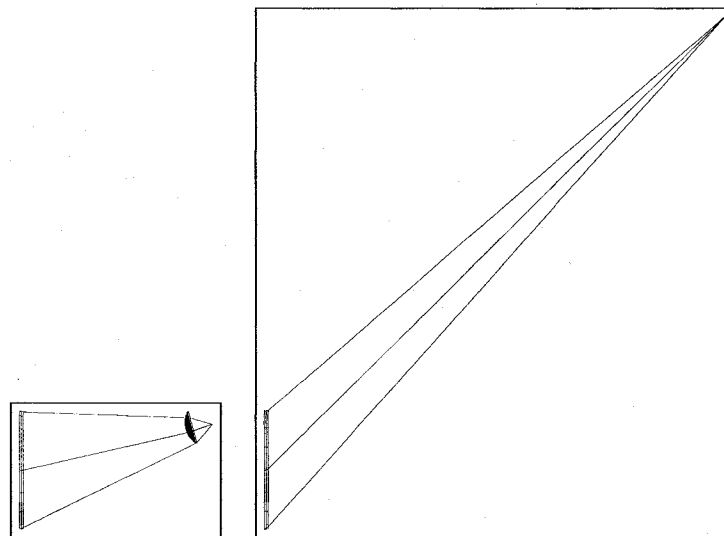
Figure 5.22: Construction optical layout of FD-2. Both construction points are diverging, and the light-rays from the closer construction point pass through a construction optic before striking the surface of the HOE.
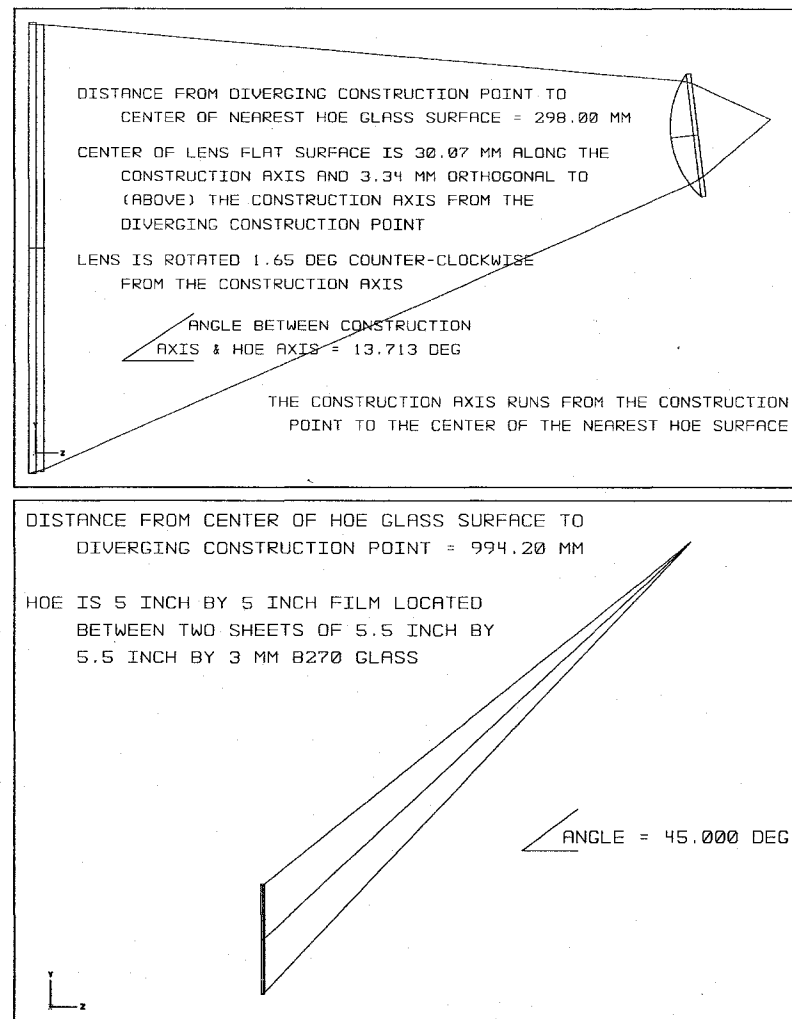
Figure 5.23: Construction optical layout of FD-2, annotated with design specifications, as sent to WP for manufacture. Both construction points are diverging. The construction lens is Edmund Scientific part 45716, which is a 75 mm focal length BK7 PCX lens with a radius of 38.76 mm. The top and bottom diagrams are scaled differently to maintain readability of the annotations; in both diagrams the HOE is 127 mm (5 inch) square.

Chapter 6

Experimental Apparatus

A completed RTTH system requires methods for device calibration and validation. The present work, focused on the development of an HOE suitable for RTTH, utilized an optical table rather than a hand-held plastic housing to implement and test our first complete RTTH optical system design.

There are several advantages to using an optical table and associated equipment to hold and position each of the optical elements specified by FD-2 (see subsection 5.6.2), rather than creating a "production" housing suitable for portable use. In particular, by using high-precision optical positioners on a rigid optical table, the following advantages are obtained:

- It is relatively trivial to experimentally interchange the components and procedures used to produce the LCD image on the fiber optic faceplate.

- Other experimental/validation hardware can easily be precisely and rigidly attached to the RTTH system.

- Each optical element can be precisely positioned, with tolerances better than typical manufacturing.

- Groups of elements can easily be precisely positioned relative to one-another.

- It is possible to empirically determine which (if any) aspects of optical alignment have tolerances small enough to require calibration in a manufactured system.

Because our RTTH optical design did not specify the method of applying the LCD image to the fiber optic faceplate, full implementation of our RTTH system required that we try various types of source image formation equipment, such as different laser sources, beam spreaders, etc. Our experiments made use of a computer-controlled camera, and it was helpful to be able to rigidly attach and precisely position it relative to the RTTH system. In general, the precise positioning capability enabled us to truly test the performance of our RTTH design, and could also be used in the future to determine or verify the necessary manufacturing tolerances for a "production" RTTH system. Finally, *because we could position each optical element very precisely, and because we did not connect a real scanner to our RTTH system, our system did not require additional calibration procedures for the present work.*

67

6.1 Source image formation equipment

We experimented with different methods of applying the LCD image to the fiber optic faceplate. Ultimately, we made use of a laser module, a beam expander (to achieve a larger diameter collimated beam), and a mirror to direct a wide laser beam through the entire active area of our LCD display to the fiber optic faceplate. All of this equipment, including the LCD and faceplate, was held by optical mounts attached to precision positioning equipment.

6.1.1 Laser

We used a 532 nm green laser module (class IIIa) as our laser source. In particular, we used part AGLM2-5 from IMM Meßtechnologie GmbH (http://www.imm-laser.de/). This diode pumped solid state laser module satisfies our safety concerns, due to its low power (the output of a class IIIa laser is < 5mW). As specified in section 5.2, we use a green laser due to the human eye's peak sensitivity to green light. Indeed, the laser module is capable of producing a small dot visible at 200 meters. Because RTTH spreads the light output of the laser over a large virtual image, and because of inefficiencies in the optics and the LCD, it is critical that the laser source be as bright as possible within the limits of safety.

6.1.2 Beam expander and mirror

The laser beam was directed through a 15x beam expander to increase the laser module's small beam diameter to about 25 mm before the beam was redirected via a front-surface mirror to the LCD. The beam expander was Edmund Optics (http://www.edmundoptics.com/) part NT58-209, which features a 2 mm entrance aperture, a focus lock, anti-reflection coated optics, and a focus mechanism that does not rotate the optics. The mirror was a high quality "right angle prism" broadband dielectric mirror (part MRA25-E02) from Thor Labs, Inc. (http://www.thorlabs.com/).

6.1.3 LCD

We used a small, high resolution monochrome LCD from Kopin Inc., Taunton, MA (http://www.kopin.com/). The LCD, model number 1280M, measures 24.6 mm diagonally, and has a specified resolution of 1280x1024 pixels. The LCD was controlled by its evaluation kit from Kopin, Inc., which connects the LCD as an additional display to a desktop computer, allowing simple and direct control of the display's image. We built a custom extension cable to increase the cable length between the LCD and the evaluation kit from about 40 mm to about 200 mm.

Ideally, it would be possible to use the LCD in front of the expanded laser beam as an image projector, suitable for projecting the LCD image onto the back surface of the fiber optic face plate. Unfortunately, the small pixel size of the LCD causes it to act as a diffraction grating for the laser light, each pixel turning its tiny incoming fragment of the laser beam into an extensive array of diverging laser beams, each separated by an angle of approximately 1.2°. Accordingly, if a "projection screen" is placed a short distance, such as 5 cm, from the LCD, multiple offset and overlapping images of the LCD will appear on the screen. If the screen is moved a sufficient distance from the LCD to keep the projected images from overlapping, then each of the projected images is very blurred, presumably due to interaction of the various transverse cavity modes of the laser. A laser

does not uniformly distribute light across its beam's cross section. Rather, the cross-sectional intensity of a laser beam is the combination of many modes, with each of the higher modes containing multiple lobes that are out of phase with one another, as shown in Figure 6.1.

TEM$_{00}$ TEM$_{01}$ TEM$_{10}$ TEM$_{11}$ TEM$_{02}$
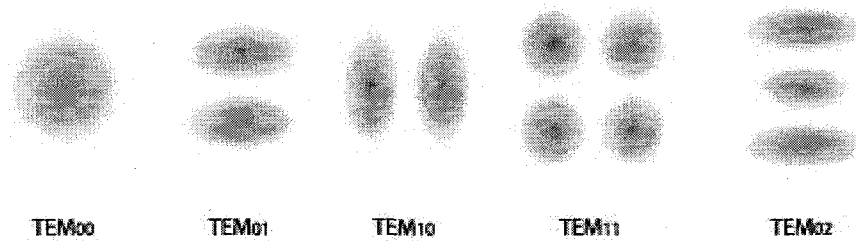
Figure 6.1: The cross-sectional intensity profile of a laser beam is formed by the combination of many modes, the lowest-order of which are depicted here. TEM$_{00}$ is typically the "desired" mode, since each of the other modes contains multiple lobes, each of which is 180 °out of phase with its neighboring lobe(s) [40].

Ultimately, we placed the faceplate as close as possible to the LCD, to minimize the spread of the "ghosted"/diffracted images. Even this arrangement was not without fault, however, because even the small thickness of the LCD glass provided sufficient distance for the unwanted first order diffraction images to shift approximately one pixel width in each direction relative to the desired zeroth order image, effectively applying a one pixel radius blur to the LCD's image.[1] We considered this acceptable in light of the high resolution of the LCD compared to the ultimate resolution of the RTTH system.

6.1.4 Fiber optic faceplate

To minimize the diffraction effects discussed above, the LCD was held directly against the fiber optic faceplate provided by Rallison. Both devices were held by a single clamp with rubber jaws, allowing them to be effectively positioned as a unit. The use of a separate diffuser, in addition to the faceplate, was considered to help spread the laser light more evenly across the HOE, but the use of a single clamp did not allow sufficient room to easily insert additional diffusing material between the LCD and the faceplate. It was trivial, though, to add additional diffusing material between the LCD and the laser. Theoretically, this should have diffused the light entering each LCD pixel, resulting in a corresponding diffusion of light exiting each fiber of the faceplate. Empirically, however, adding such diffusing material only blurred the projected virtual image, presumably because the thickness of the glass between the LCD and the faceplate allowed the light from each pixel to diffuse not only into the appropriate fibers of the faceplate, but also into the neighboring fibers associated with neighboring pixels. Adding a diffuser immediately after the faceplate was impractical due to the faceplate's curved surface. In light of these difficulties, and because the faceplate itself sufficiently illuminated the HOE, no additional diffusing material was used in the experiments.

---

[1]Just to be clear, we want the zeroth-order image from the LCD, but the first-order image from the HOE.

6.2 Computer-controlled camera

Our experimental data were gathered by a computer-controlled digital SLR camera made by Canon Inc., Tokyo, Japan. The camera, model *EOS 5D*, utilizes a 12.8 megapixel full frame (36 mm x 24 mm) CMOS image sensor that incorporates special hardware to minimize both fixed-pattern and random noise [41]. The primary lens used with the camera was Canon's model *EF 180mm f/3.5L Macro USM*, which produces 1x life-size magnification at its minimum focusing distance of 0.48 meters. Canon's MTF chart for the lens is shown in Figure 6.2.
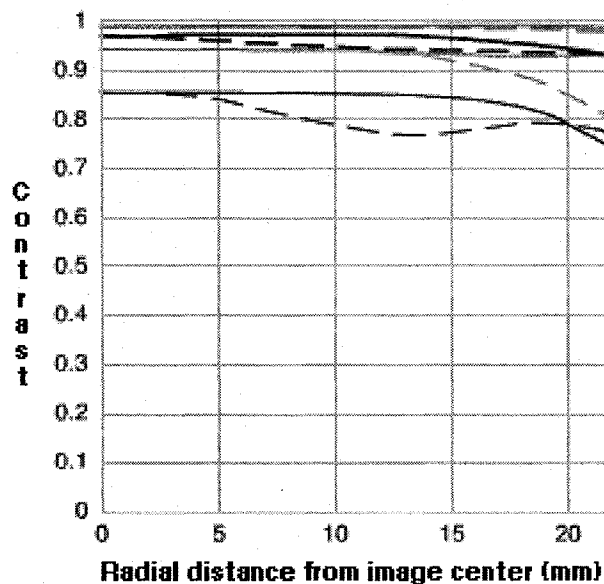
Figure 6.2: MTF Chart for Canon's 180 mm Macro lens [42]. An ideal lens would be graphed as horizontal lines at a constant value of 1.0, indicating an aberration-free lens (which is impossible to achieve). The black lines correspond to lens operation at f/3.5, the maximum relative aperture of the lens, while the blue lines correspond to a relative aperture of f/8. "The bold curves are for 10 lines/mm spatial frequency and the hair-line curves are for 30 lines/mm spatial frequency." The line pairs were oriented radially (*sagittally*) for the solid curves, and orthogonal to that (*tangentially* or *meridionally*) for the dashed curves. Differences between sagittal and tangential performance are indicative of third-order astigmatism. "Normally, if the MTF curves for 10 lines/mm are 0.8 or higher, the lens is superior" [43].

Canon provided us with their Eos Digital Software Development Kit (EDSDK), which allowed us to write software capable of controlling all electronically-controlled camera functions. As long as the camera was in "manual" mode, the computer attached to the LCD was also able to control virtually all aspects of camera shooting, other than focal length and focal distance. Accordingly, our experiment-control software was capable of displaying arbitrary virtual images, and then photographing each of those virtual images multiple times, using a range of aperture and shutter-speed settings.

The camera and lens were mounted to the optical table via a custom-assembled six degree-of-freedom positioning stage. The stage allowed horizontal translational movement over a one inch by one inch area, allowing the camera to be positioned anywhere above the optical table by attaching the stage to the appropriate mounting hole(s) on the optical table's surface (the mounting holes were spaced on a one inch grid).

6.3 Optical table positioners

As already discussed, all of the optics were held and positioned using precision translation and rotation stages. The arrangement of the stages allowed for both the individual movement of optical elements, as well as the movement of optical groups. The layout of the optics and their positioners on the table is shown in Figures 6.3-6.6.
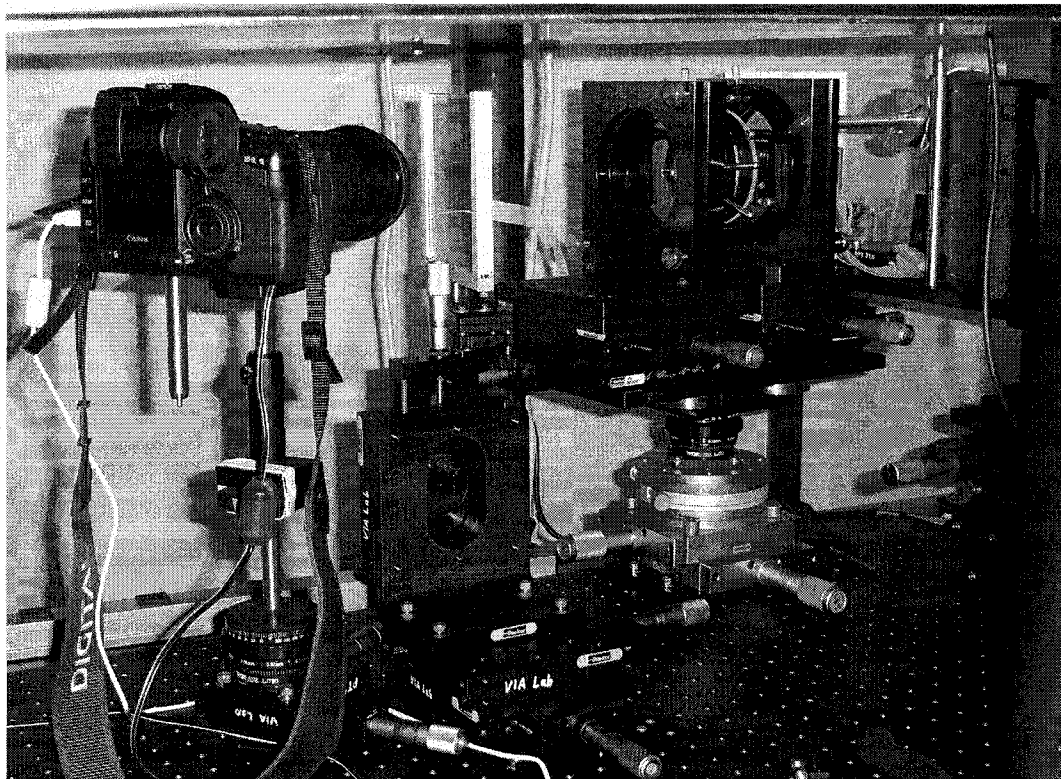
Figure 6.3: Positioners holding the camera, HOE, lenses, and part of the source image formation equipment. The HOE, located in the center, is shown from a different perspective in Figure 6.5.
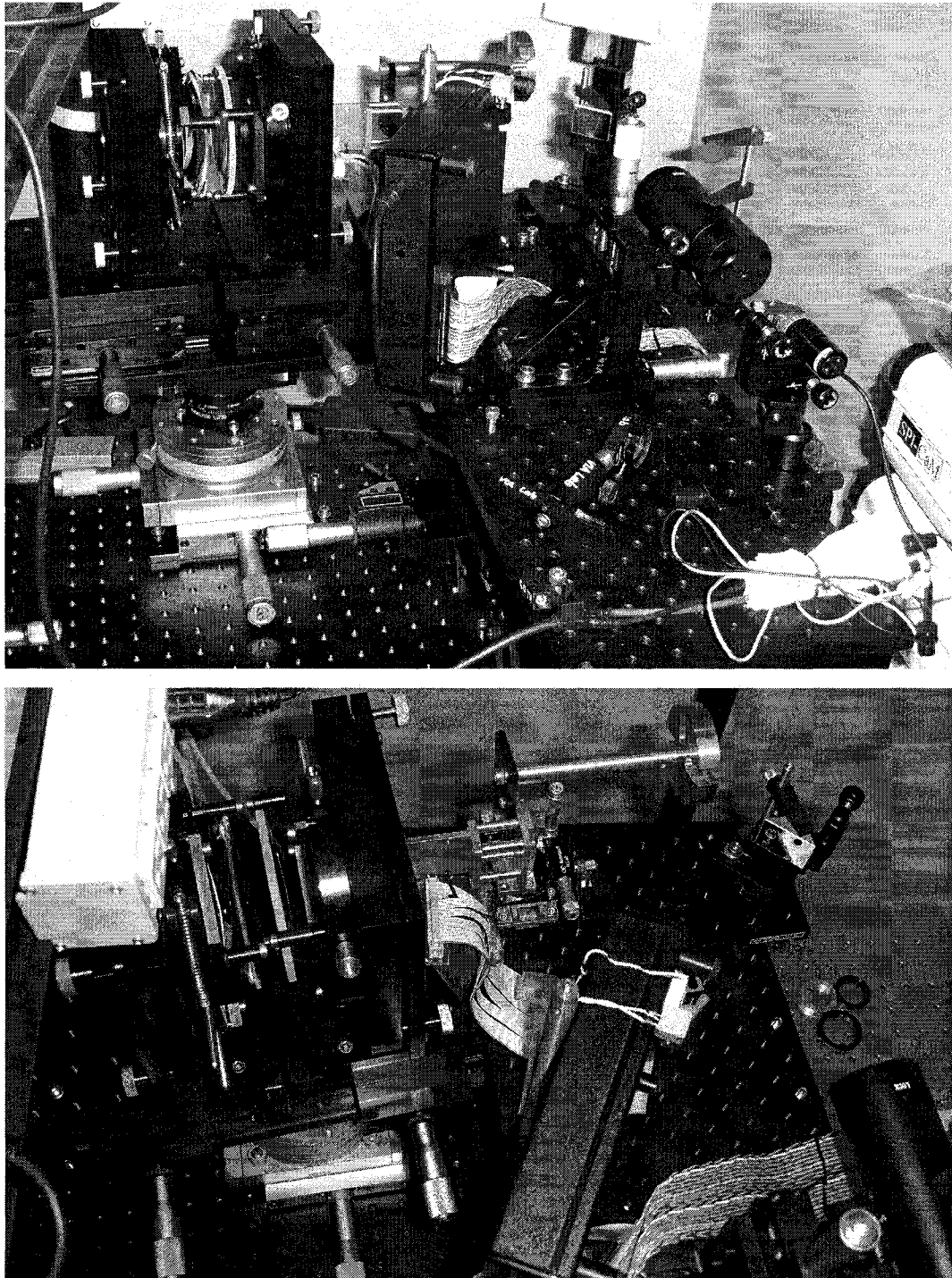
Figure 6.4: Positioners holding the lenses and all of the source image formation equipment, viewed from two different perspectives.
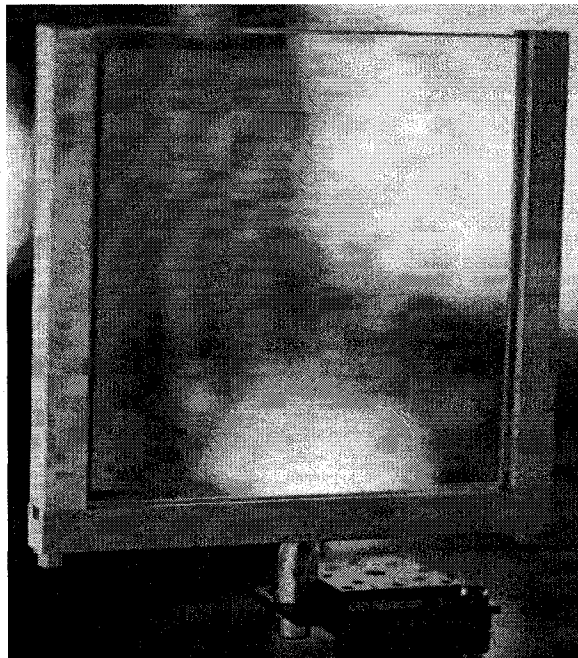
Figure 6.5: The HOE in its holder
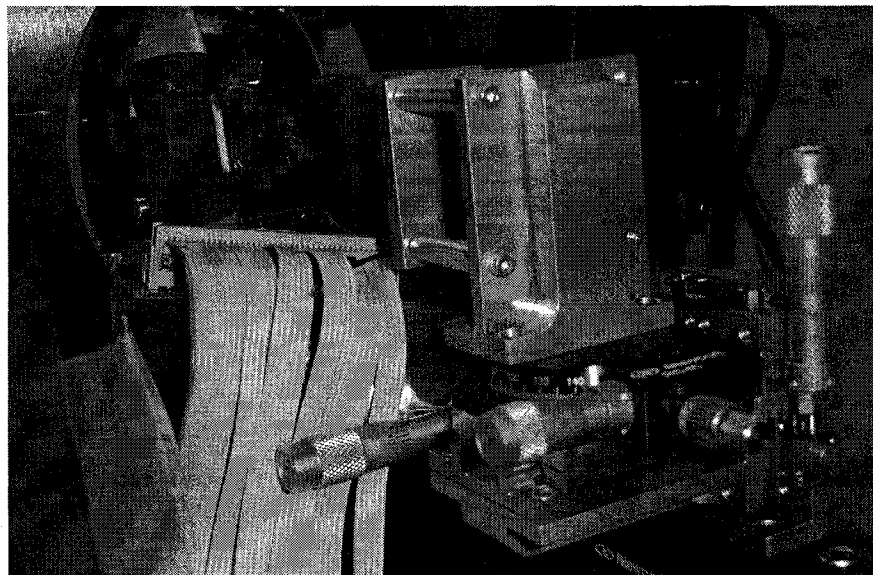
Figure 6.6: Positioners holding the LCD and fiber optic faceplate

Chapter 7

Experiments

7.1 Alignment

Before the optical characteristics of our RTTH system could be analyzed, it was necessary to properly align each of the optical components.[1] Unlike the anticipated typical handheld arrangement, in which the optical plane (the plane formed by the centers of all the optical elements) is oriented vertically, for the present experiments the RTTH system was setup such that the optical plane was oriented horizontally, parallel to the top of the optical table. Alignment was accomplished by a multistep process consisting of the following stages:

- Axial alignment

- Axial spacing

- Off-axis adjustments

7.1.1 Axial alignment

The initial alignment step was to temporarily position each optical element such that it was centered on and fully orthogonal to the particular "optical axis" depicted in Figure 5.21. This was accomplished by aligning two coaxial laser beams down the optical axis (one from each direction), and then carefully positioning each of the lenses such that as each of the laser beams passed through the lenses all of the reflections and refractions of the laser beam were uniformly centered around the laser beam itself. Because each lens significantly diverges any laser beam passing through, it can be difficult (and inaccurate) to align a series of optical elements using a single laser. Accordingly, we made use of a second co-axial laser positioned at the opposite end of the optical system, so that all of the optical elements received a sufficiently narrow beam for alignment. After aligning the lenses, the HOE was then made orthogonal to the optical axis by positioning it such that its reflection of the laser beam was also centered on the beam path, and the HOE was then centered on the optical axis by carefully translating the HOE until the laser beam's observable "bright spot" was measured to be at the center of the HOE's surface. Finally, the fiber optic faceplate (with attached LCD) was

---

[1]In a fully-implemented RTTH system that contains an actual scanning device, it would also be necessary at this stage to register the alignment and mapping between the virtual image and the scanning device.

74

likewise made orthogonal to the optical axis by aligning the LCD's zeroth-order reflection of the laser beam, and was then centered on the optical axis by translating it until the laser's "bright spot" was in the measured center of the faceplate.

7.1.2 Axial spacing

Once all of the optical elements were temporarily coaxially aligned as just described, their spacings along the temporary axis were then adjusted to be at specific values. This was accomplished using a procedure developed by the author, making use of telescoping gauges. A telescoping gauge is a tool that assists in precisely measuring the inside diameter of pipe. The tool is "T" shaped, as shown in Figure 7.1, with the top of the "T" being adjustable in width and the bottom of the "T" being the handle. The adjustable width of the tool is implemented using two sliding spring-loaded rods. At the bottom of the handle is a knob that can be turned to adjust the holding tension against the sliding rods.
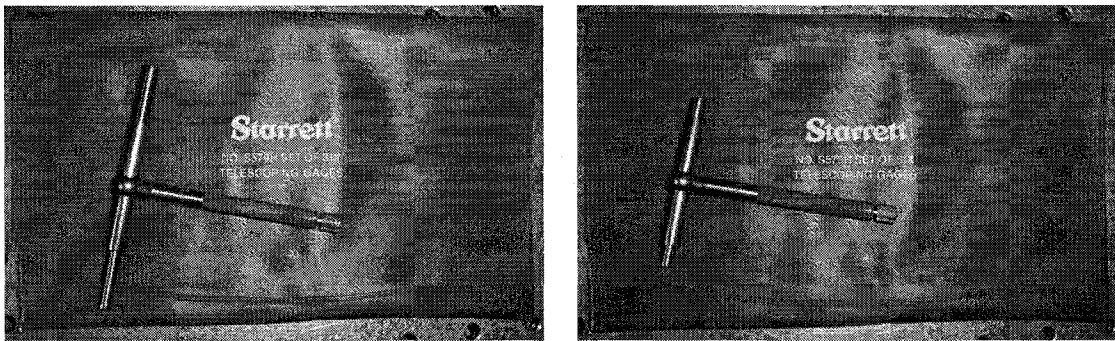
Figure 7.1: Telescoping Gauge, the tool used to adjust optical spacing. Shown with the top fully extended (left) and retracted (right).

The author would precisely set the width of the tool by adjusting locking calipers to the desired width, placing the top of the tool between the calipers, reducing the holding tension until the sliding rods expanded to firm contact against the caliper's measuring points, and then increasing the tool's holding tension to lock the tool at the exact width of the caliper. The tool would then be placed along the optical axis between two optical elements, and one of those optical elements would then be translated until the tool barely made contact with the centers of both optical elements. The translated optical element's current position along the optical axis would then be read directly from the relevant translation stage. The optical element would then be backed off to allow the removal of the tool before being returned to the previously observed position.

7.1.3 Off-axis adjustments

Upon completion of the axial alignment and spacing, it was then straightforward to appropriately position the optical elements off-axis by adjusting the translation and rotation stages holding those elements. Because each stage allowed the precise read-out of its current position, it was trivial to adjust each stage to accomplish the desired offset in translation or rotation relative to the prior known axial positioning.

7.2 Initial attempt at validation

Upon completion of alignment, a quick visual check confirmed that the virtual image appeared to be stable, well focused, and properly positioned. We then set up the camera to perform a stereo/multi-camera 3D reconstruction of the virtual image's position. This approach proved difficult and of dubious reliability, however, because the virtual image did not appear well-focused when viewed through the large-aperture lens of the camera. Rather, depending on the focal distance to which the lens was set, the virtual image appeared to be well-focused only in either of the horizontal or vertical directions, but not both at the same focal distance.

The observed optical distortion was a type of *first-order astigmatism*, such as is commonly corrected for in human vision by use of a cylindrical surface. (Because most optical systems are designed to be rotationally invariant about a single optical axis, such systems generally do not suffer from first-order astigmatism. Instead, rotationally invariant systems suffer from third-order astigmatism, whereby the system functions at a different effective radial and tangental focal lengths, as mentioned in subsection 5.5.3. Compared to pure first-order astigmatism, which is uniform across an optic, pure third-order astigmatism is not present on the optical axis, and it increases monotonically along the radius of an optic.) Because the first-order astigmatism was not readily viewed with the naked eye, it was assumed to be very small in the virtual image but "amplified" by the narrow depth-of-field of the large-aperture camera lens.

Attempts were made to reduce the effect of the astigmatism by shrinking the camera's aperture. While somewhat effective, this approach did not totally remove the astigmatism, making it difficult to derive meaningful results from multi-camera stereo reconstruction. Furthermore, as the aperture was reduced, speckle noise from the fiber-optic faceplate's diffusion of the laser light became more noticeable. (As aperture increases, such speckle is quickly averaged away over the surface of the lens, but at small aperture it is quite visible.) We acquired and tested many different diffusing materials, all of which introduced significant speckle, and we were unable to reduce the speckle to levels lower than that obtained when the faceplate was the only diffusing material in the system.

During this stage, it was also discovered that either the LCD or its evaluation kit had ceased proper functioning. Half of the columns in the LCD apparently ignored their input, but instead duplicated the value of their neighboring column, in effect halving the resolution of the LCD in one direction. This effect is particularly noticeable when rendering a line one pixel thick; if the line lies in one of the "ignored" columns, then it will not appear at all on the LCD.

7.3 Validation

In light of the astigmatism problem discovered in the first attempts at validation, we developed a different validation method, one based on the use of a movable physical target. The physical target, shown in Figure 7.2, consisted of two orthogonal pairs of parallel lines, which intersected to form four vertices, spaced approximately 80 mm horizontally and 70 mm vertically. (As laid out "sideways" on the optical table, the simulated virtual image size is 112 mm wide by 104 mm tall.)

Software capable of controlling both the camera and the LCD was written which allowed an observer to move five virtual crosshairs within the plane of the virtual image. Four of the crosshairs could be aligned with the four vertices of the physical target, while the fifth crosshair existed only to provide a central target on which the camera could be focused. An example photo of the crosshairs
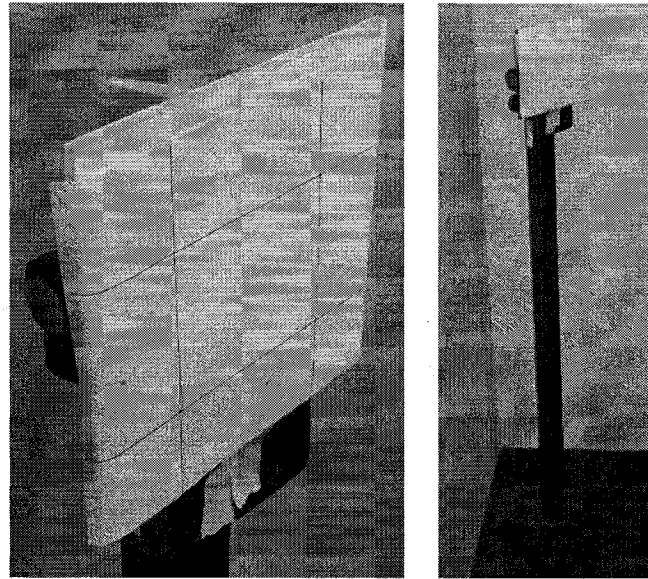
Figure 7.2: Physical target used for validation (left), shown with its mounting (right)

aligned with the physical target is shown in Figure 7.3.
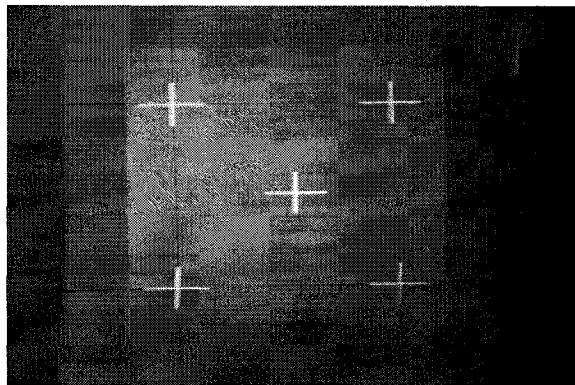
Figure 7.3: Virtual crosshairs aligned with the vertices of the physical target

By aligning the virtual crosshairs with the physical target's vertices, it was possible to visually check whether or not the virtual image and the physical target were coplanar. The check coplanarity, one needed only to move a viewing aperture (whether an eye or a camera lens, etc.) back and forth; assuming that the virtual image was stable in space, the crosshairs and physical vertices would remain aligned only if they were coplanar, otherwise motion parallax would indicate misalignment. Instability of the virtual image could be measured by minimizing motion parallax (to make the virtual image and the target coplanar), and then measuring how much the virtual crosshairs moved as viewpoint varied. Likewise, blur in the virtual image could be determined by measuring the thickness of the crosshairs, which were rendered on the LCD as 1 pixel thick. Vertical blur was measured using horizontal lines, and horizontal blur was measured using vertical lines. All mea-

surements were taken from camera-acquired images in units of pixels, which for each image were then converted to mm, with physical pixel spacing in the image determined by counting the pixel spacing between the physical vertices, whose actual physical spacing was known.

Because of the astigmatism present in the system, the virtual image plane existed at two different depths, one for vertical focus and the other for horizontal focus. When properly aligned with the physical target, each of these two depths would be free of motion parallax in its focused direction, but would exhibit motion parallax in its unfocused direction. Accordingly, by aligning the target twice, once for horizontal focus (by adjusting the physical target to eliminate perceptible horizontal parallax) and once for vertical focus (adjusted likewise, but eliminating vertical parallax), it was possible to independently measure horizontal blur and image stability in the horizontally focused virtual image plane, and likewise for vertical blur and image stability in the vertically focused virtual image plane. Furthermore, it was also possible to measure the astigmatism-induced distance between the two focal planes.

Chapter 8

Results and Analysis

Following the validation procedure described in section 7.3, we were able to collect data regarding several aspects of the virtual image, including its:

- Position and orientation

- Stability

- Minimum feature size (representative of blur)

All of the measurements were taken with the camera's image plane located 137 cm from the physical target, wherever the physical target was positioned. Due to the relatively large distance between the relatively small HOE and the virtual image, the camera was only able to view the entire virtual image over a range of 2.5 °horizontally and 2.7 °vertically.

All of the images used for this analysis were taken with a 180 mm lens at a relative aperture of f/16, which amounts to an aperture diameter of 11.25 mm. The human eye has a maximum relative aperture of about 2.1 corresponding to an aperture diameter of about 8 mm [44, 45]. Accordingly, at any given distance from the HOE, the camera was configured to image each point through a larger area of the HOE than would a human viewer, insuring that the blur results obtained were a "worst case" upper bound.

8.1 Position and orientation

The vertical focal plane of the virtual image was found to be located 100 cm from the center of the viewing surface of the HOE, in agreement with simulation, but the horizontal focal plane was found to be located 3 cm closer to the HOE.

The 33 ° angle between the virtual image and the optical axis depicted in Figure 5.21 was verified by approximately measuring the orthogonal distance from Figure 5.21's optical axis to the center of the physical target. That distance, measured at 559mm, can be combined with the known 1m spacing between the target and the HOE to calculate the angle as follows:

$$\theta = \arcsin \frac{559mm}{1000mm} = 34°$$

(8.1)

79

Accordingly the angle between the virtual image and the optical axis depicted in Figure 5.21 was verified to within 1°.

We believe that the astigmatism-induced 3 cm gap between horizontal and vertical focus is due to our method of measuring focus as the distance of light rays from their ideal focal point *within the plane of the image surface* (whether traced forward to the virtual image or in reverse to the faceplate's source image). If one of the light rays forming the virtual image is at an angle of 45°relative to the surface-normal of the virtual image, then that ray's measured focal error will be the same regardless of whether its blur is orthogonal to the virtual image or within the plane of the virtual image. If the blur is orthogonal to the virtual image, the light ray will intersect the plane of the virtual image at a distance from its ideal focal point that is equal to its actual orthogonal blur. However, as the above angle decreases from 45°, blur orthogonal to the virtual image registers as an increasingly smaller error compared to blur within the plane of the virtual image. Because the virtual image is located one meter from the HOE, its light rays have an angle substantially less than 45°, and so a 3 cm blur orthogonal to the virtual image results in rays that pass *very close* to their ideal focal point, resulting in a very small measured in-plane focal error.

If this hypothesis is correct, then future RTTH systems, if designed according to the methods and criteria used for this work, would be expected to have less "unanticipated" astigmatism if they project a virtual image closer to the HOE. Likewise, future RTTH systems that project the virtual image farther away would be expected to have more "unanticipated" astigmatism, unless additional design steps were taken to separately measure and penalize astigmatism during design optimization.

8.2 Stability

Across a series of three vertical viewpoints, the four crosshairs were found to drift over a vertical range of not more than 0.83 mm. In theory, the left crosshairs should perform identically and the right crosshairs should perform identically, and the observed deviations from such are likely due to either improper manual alignment of the physical target before beginning the experiment or to shifting of the physical target during the experiment. Examining Table 8.1, if the one outlier (the bottom-right crosshair) is excluded, then the maximum observed vertical range is 0.34 mm.

Over a similar series of three horizontal viewpoints, the four crosshairs had a maximum observed horizontal drift range of 0.99 mm. It is understandable that horizontal drift is greater than vertical drift, since the optical system is co-planar rather than co-axial, and only vertical drift is orthogonal to the optical plane. (Axial systems inherently have fewer aberrations than non-axial systems, in part because in an axial system any perceived drift is orthogonal to the optical axis.)

Because the virtual image was less stable horizontally, horizontal performance was taken as a "worst case" scenario, and it was further investigated by positioning the camera at two additional, more extreme viewpoints. The camera was positioned at each horizontal extreme such that only two of the crosshairs were visible at f/16, resulting in an increased viewing angle of 6.1° horizontally. Over this more extreme range of viewpoint, the maximum observed drift range for the crosshairs was 2.51 mm.

All of the stability measurements are well within the limits predicted by our simulation, as described in Figure 5.18 in subsection 5.6.2. (Recall that for the experiments the system was rotated on its side, and so experimental horizontal drift correlates with simulated vertical stability, and likewise experimental vertical drift correlates with simulated horizontal stability.)

Table 8.1: Virtual image stability: perceived range of each virtual crosshair

Camera Position	Cross-Hair Vertical Position in mm (Relative to Coresponding Physical Vertices)				Max Range
	Bottom-Left	Top-Left	Top-Right	Bottom-Right	
Center	0.02	0.20	-0.33	0.09	
Top	0.07	0.32	0.01	0.29	
Bottom	0.02	0.02	-0.31	-0.54	
Range	0.05	0.29	0.34	0.83	0.83
Camera Position	Cross-Hair Horizontal Position in mm (Relative to Coresponding Physical Vertices)				Max Range
	Bottom-Left	Top-Left	Top-Right	Bottom-Right	
Center	-0.15	-0.09	0.10	-0.12	
Left	-0.92	-0.67	-0.39	-0.72	
Right	0.07	0.11	-0.85	-1.09	
Range	0.99	0.78	0.95	0.97	0.99
Far-Left			-2.09	-2.63	
Far-Right	-1.80	-1.95			
Range	1.88	2.06	2.19	2.51	2.51

In an actual system, the center of each virtual point's drift range would be the location aligned with the virtual point's appropriate sensor data, resulting in a maximum position error of half the drift range. Accordingly, ignoring the astigmatism, as long as the viewer kept the entire virtual image in their field of view, virtual objects should be perceived as being *located within 0.5 mm of their actual scanned location*. It is currently unknown how the astigmatism problem would affect perceived 3D location in space. Optically it acts as a bimodal blur in depth, rather than a source of "instability" in the virtual image, but the resultant motion parallax effects could affect its perceived location in all three dimensions.

8.3 Minimum feature size

The graphical crosses used to analyze the stability of the virtual image were also used to measure the minimum projectable feature size in the virtual image, which corresponds to the size of one LCD pixel, blurred by the LCD/faceplate interface (see section 6.1) and then magnified and blurred by the optics of FD-2. Minimum feature size was measured using a one-pixel thick line, oriented either horizontally or vertically. Figure 8.1 shows how line thickness was measured. Within the basic viewing range, the maximum observed vertical blur rendered a one-pixel thick horizontal line 0.79 mm thick, and the horizontal blur likewise rendered a one-pixel thick line 0.78 mm thick. Considering that these measurements are less than the observed drift ranges above, blur is unlikely to be a limiting factor in any RTTH implementation based on FD-2. At the extreme far-left and far-right viewpoints, where only two of the crosshairs were barely visible, the maximum observed horizontal blur increased modestly to 0.97 mm, significantly less than the horizon drift range at these extremes. Details are provided in Table 8.2.
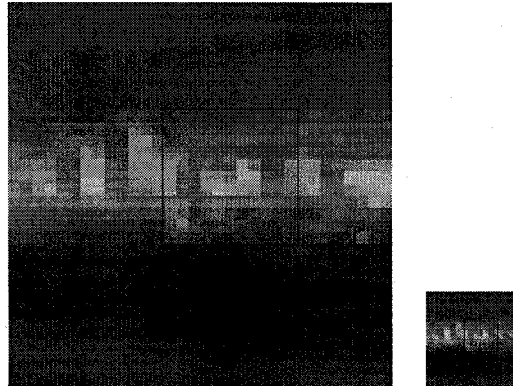
Figure 8.1: The blue box indicates the pixels that were included as part of the holographic line for purposes of measuring its thickness. The figure on the left shows a highly magnified view of a portion of a holographic line, while the figure on the right shows a less magnified view, for purposes of better showing how the line will normally be perceived.

Table 8.2: Virtual image blur: directional blurring of each virtual crosshair

Camera Position	Cross-Hair Horizontal Thickness (Vertical Blur) in mm				Max
	Bottom-Left	Top-Left	Top-Right	Bottom-Right	
Center	0.54	0.54	0.54	0.44	
Top	0.54	0.54	0.39	0.44	
Bottom	0.54	0.79	0.59	0.54	
Max	0.54	0.79	0.59	0.54	0.79
Camera Position	Cross-Hair Vertical Thickness (Horizontal Blur) in mm				Max
	Bottom-Left	Top-Left	Top-Right	Bottom-Right	
Center	0.68	0.73	0.49	0.54	
Left	0.58	0.58	0.77	0.77	
Right	0.63	0.78	0.73	0.78	
Max	0.68	0.78	0.77	0.78	0.78
Far-Left			0.88	0.83	
Far-Right	0.73	0.97			
Max	0.73	0.97	0.88	0.83	0.97

Chapter 9

Conclusions

Visualizing real-time scans of objects at their actual location merged with direct vision facilitates interpretation and interaction based on natural psychophysical skills. Extending the original mirror-based concept to holography yields a wide variety of possible mappings between the display device and the virtual image, permitting a smaller display device and a larger virtual image. As such, it may represent an important new method of displaying information *in situ* within an object in real time.

This dissertation presented the design, alignment, and initial validation of the first RTTH optical system. Although not designed for any particular RTTH application, the HOE was designed to meet most of the critical and unique design challenges that RTTH imposes, *successfully demonstrating the projection of a stable, off-axis magnified virtual image.* Accordingly, this dissertation has demonstrated the chief advantages of RTTH:

- Essentially viewpoint-independent direct viewing of an in-situ virtual image, with the virtual image physically larger than the source image used to project it

- A clear, optically unobstructed view of reality in the region surrounding the virtual image

- Lack of any imposed physical objects between the HOE "viewport" and the virtual image

Our initial validation of the RTTH system revealed an unexpected significant astigmatism error, which in retrospect our design-optimization's merit function was not well-equipped to prevent. Although not readily apparent to normal human vision, because the effect of virtual-image astigmatism on human spatial perception is not known, it is possible that the astigmatism could negatively impact real-world use of this RTTH system. The astigmatism also greatly hampered our initial attempts at system validation, motivating us to develop a custom method for analyzing our RTTH system in spite of its astigmatism.

Aside from the aforementioned astigmatism, our RTTH system performed well, with resolution and stability on the order of one mm when viewed from "normal" viewpoints, and with the geometric measurements of the virtual image in good agreement with our model. Encouraged by our results, we intend to move forward, conducting psychophysical experiments with our present system and ultimately producing a new RTTH system suitable for visualizing clinical ultrasound for the guidance of invasive deep procedures, as will be described in the next chapter.

83

Chapter 10

Future Work

As an exploratory proof-of-concept, the present work has led to many opportunities for further research. The present system could be fit into a custom-built hand-held physical housing, attached to either a real or a virtual scanner, and used as a basis for conducting psychophysical experiments to better understand the engineering trade-offs involved in designing an RTTH system. Furthermore, new and better RTTH designs are now more feasible, each targeted toward a specific use and particular scanning technology.

10.1 Hand-held form-factor

Building a hand-held physical housing for the present RTTH design would enable additional experimentation with RTTH, in particular with regard to psychophysical experimentation. Before construction of such a hand-held device, additional experimentation with the present system would be necessary to determine the allowable manufacturing tolerances for the physical housing. Any required tolerances that are not readily attainable would require implementation of physical adjustments and appropriate calibration procedures.

10.2 Virtual tomographic holography

Damion Shelton, a former student in our lab, has constructed a virtual ultrasound system that is able to generate simulated image data for a tracked Sonic Flashlight [27, 31]. He has termed the process *virtual tomographic reflection* (VTR). The advantage of such a system is that it can be used to generate data for any desired in situ objects and then track user interaction with such objects, enabling psychophysical studies to produce accurate, quantified data.

The virtual ultrasound system would make an ideal "scanner" for the present RTTH implementation, enabling *virtual tomographic holography* (VTH). VTH would demonstrate RTTH applied to various application areas, and it would set the stage for psychophysical studies. Such studies could help determine better optics-optimization criteria for future RTTH systems.

It should be pointed out that the calibration procedures for the VTR system are more complex than those for the normal Sonic Flashlight, due to the need to calibrate components of both the visualization and image data simulation systems. Likewise, new calibration procedures would be necessary for VTH.

<center>84</center>

10.3 Psychophysical experiments

Psychophysical experiments and human-factors studies hold the potential of significantly increasing our understanding of the design trade-offs involved in RTTH systems. These experiments could elucidate the effects of various optical distortions on human-performance of various tasks in the context of RTTH systems. Because physical interaction between a human and a virtual image is a primary goal of RTTH, the knowledge gained from such psychophysical experiments could serve as a valuable guide when creating the merit functions that will guide optimization of future RTTH systems.

10.4 Real ultrasound transducer

Our lab has recently received funding from the NIH to continue this research by implementing an RTTH system designed for the specific purpose of using clinical ultrasound to guide invasive medical procedures. Designing such a system will be considerably more difficult than was the design of FD-2, because such a system will require a significantly shorter distance from the HOE to the virtual image, amplifying the effects of optical distortions. Our current plan is to make use of at least one secondary HOE to help minimize the relevant distortions to acceptable levels. Ultimately, a full implementation (with a real scanning device) will be required to reap the potential benefits of RTTH.

Bibliography

[1] M. Rosenthal, A. State, J. Lee, and e. al., "Augmented reality guidance for needle biopsies: a randomized, controlled trial in phantoms," in *MICCAI 2001*, ser. Lecture Notes in Computer Science, vol. 2208. Springer-Verlag, 2001, pp. 240–248.

[2] F. Sauer, A. Khamene, B. Bascle, L. Schimmang, F. Wenzel, and S. Vogt, "Augmented reality visualization of ultrasound images: system description, calibration, and features." in *International Symposium on Augmented Reality*. New York City: IEEE and ACM, 2001, pp. 30–39.

[3] H. Fuchs, A. State, M. Livingston, W. Garrett, G. Hirota, M. Whitton, and E. Pisano, "Virtual environments technology to aid needle biopsies of the breast - an example of real-time data fusion," *Stud Health Technol Inform*, vol. 29, pp. 60–1, 1996.

[4] H. Fuchs, A. State, E. Pisano, W. Garret, G. Hirota, M. Livingston, M. Whitton, and S. Pizer, "Towards performing ultrasound-guided needle biopsies from within a head-mounted display," in *Visualization in Biomedical Computing*, Hamburg, Germany, 1996, pp. 591–600.

[5] A. State, J. Ackerman, G. Hirota, J. Lee, and H. Fuchs, "Dynamic virtual convergence for video see-through head-mounted displays: Maintaining maximum stereo overlap throughout a close-range work space." New York City: International Symposium on Augmented Reality (ISAR), 2001.

[6] A. State, M. Livingston, W. Garret, G. Hirota, M. Whitton, E. Pisano, and H. Fuchs, "Technologies for augmented reality systems: realizing ultrasound-guided needle biopsies," in *ACM SIGGRAPH*, New Orleans, LA,, 1996, pp. 439–446.

[7] K. Masamune, Y. Masutani, S. Nakajima, I. Sakuma, T. Dohi, H. Iseki, and K. Takakura, "Three-dimensional slice image overlay system with accurate depth perception for surgery," in *Medial Image Computing and Computer-Assisted Intervention (MICCAI)*, vol. 1935. Pittsburgh: Springer, 2000, pp. 395–402.

[8] G. Stetten, V. Chib, and R. Tamburo, "Tomographic reflection to merge ultrasound images with direct vision," in *Applied Imagery Pattern Recognition Workshop*, J. Aanstoos, Ed., 2000, pp. 200–205.

[9] G. Stetten and V. Chib, "Overlaying ultrasound images on direct vision," *Journal of Ultrasound in Medicine*, vol. 20, no. 1, pp. 235–240, 2001.

[10] D. Drascic and P. Milgram, "Perceptual issues in augmented reality," in *Stereoscopic Displays and Virtual Reality Systems III*, ser. Proceedings of the SPIE, S. S. Fisher, J. O. Merrit, and M. T. Bolas, Eds., vol. 2653. SPIE, 1996, pp. 123–124.

[11] R. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. Julier, and B. MacIntyre, "Recent advances in augmented reality," *IEEE Computer Graphics and Applications*, vol. 21, no. 6, pp. 34–47, 2001.

[12] (2007) Nomad display systems. [Online]. Available: http://www.microvision.com/nomad.html

[13] S. Hofstein, "Ultrasound scope," U.S. Patent 4 200 885, April 29, 1980.

[14] J. C. Maxwell, "On the elementary theory of optical instruments," in *Camb. Phil. Soc. Proc.*, vol. 1, 1856 (1866), pp. 238–240.

[15] J. Maxwell, "On the general laws of optical instruments," *Quart Jourrnal of Math*, vol. 2, pp. 233–246, 1858.

[16] M. Born and E. Wolf, *Principles of Optics*, 7th ed. Cambridge University Press, 1999, ch. 4.2, pp. 153–157.

[17] H. Taylor. (2007, August) Parallax panoramagram displays. [Online]. Available: http://www.macs.hw.ac.uk/~hamish/9ig2/topic64.html

[18] (2007, March) Multiview lenticular display. [Online]. Available: http://www.business-sites.philips.com/3dsolutions/3dtechnology/multiviewlenticulardisplay/index.html

[19] G. Stetten, "System and method for location-merging of real-time tomographic slice images with human vision," U.S. Patent 6 599 247, July 29, 2003.

[20] B. Wu, R. L. Klatzky, D. Shelton, and G. Stetten, "Interaction of visual and haptic cues in the image-based perception of depth." Sarasota, FL: Vision Sciences Society, May 2006.

[21] B. Wu, R. Klatzky, D. Shelton, and G. Stetten, "Learning in image-guided reaching changes the representation-to-action mapping." Sarasota, FL: Vision Sciences Society, May 2007.

[22] R. Klatzky, B. Wu, D. Shelton, and G. Stetten, "Learning to perform actions in near space under perceptual vs. cognitive control," *ACM Transactions on Applied Perception*, (in press). [Online]. Available: http://www.vialab.org/main/Publications/pdf_protected/Klatzky_ACMTAP_2007.pdf

[23] W. Chang, M. Horowitz, and G. Stetten, "Intuitive intra-operative ultrasound guidance using the sonic flashlight, a novel ultrasound display system," *Neurosurgery*, vol. 56, pp. ONS 434–437, April 2005.

[24] W. Chang, N. Amesur, R. Klatzky, A. Zajko, and G. Stetten, "Vascular access: Comparison of us guidance with the sonic flashlight and conventional us in phantoms," *Radiology*, vol. 241, pp. 771–779, December 2006.

[25] W. Chang, G. Stetten, L. Lobes, D. Shelton, and R. Tamburo, "Guidance of retrobulbar injection with real time tomographic reflection," *Journal of Ultrasound in Medicine*, vol. 21, pp. 1131–1135, 2002.

[26] W. Chang, N. Amesur, D. Wang, A. Zajko, and G. Stetten, "First clinical trial of the sonic flashlight - guiding placement of peripherally inserted central catheters." Chicago, Illinois: Radiological Society of North America, November 2005, paper Number SSJ03-02 (in press).

[27] D. Shelton, R. Klatzky, and G. Stetten, "Method for assessing augmented reality needle guidance using a virtual biopsy task," in *IEEE International Symposium on Biomedical Imaging*. Arlington: IEEE, 2004, p. in press.

[28] D. Shelton, B. Wu, R. Klatzky, and G. Stetten, "Design and calibration of a virtual tomographic reflection system." International Symposium on Biomedical Imaging (ISBI), April 2007.

[29] K. Masamune, G. Fichtinger, A. Deguet, D. Matsuka, and R. Taylor, "An image overlay system with enhanced reality for percuta-neous therapy performed inside ct scanner," in *Medial Image Computing and Computer-Assisted Intervention -MICCAI 2002*, ser. Lecture Notes in Computer Science, vol. 2489. Pittsburgh: Springer-Verlag, 2002, pp. 77–84.

[30] W. Chang, "Guiding vascular access with the sonic flashlight - preclinical development and validation," Ph.D., University of Pittsburgh, 2004.

[31] D. Shelton, "A virtual ultrasound system for human factors & accuracy analysis of the sonic flashlight," November 2004, thesis Proposal at the Robotics Institute of Carnegie Mellon University.

[32] D. Wilson, R. Muller, P. Echternach, and J. Backlund, "Electron-beam lithography for micro- and nano-optical applications," in *Micromachining Technology for Micro-Optics and Nano-Optics III*, ser. Proceedings of the SPIE, E. Johnson, G. Nordin, and T. Suleski, Eds., vol. 5720. SPIE, 2005, pp. 68–77.

[33] A. Nowatzyk, D. Shelton, J. Galeotti, and G. Stetten. (2004) Extending the sonic flashlight to real time tomographic holography. Workshop AMI-ARCS 2004, held in conjunction with MICCAI 2004 September 30th, 2004, Rennes (France). [Online]. Available: http://ami2004.loria.fr/

[34] J. M. Geary, *Introduction to Lens Design: With Practical Zemax Examples*. Willmann-Bell, Inc., 2002.

[35] W. J. Smith, *Modern Optical Engineering*, 3rd ed. McGraw-Hill, July 2000.

[36] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, no. 4598, pp. 671–680, May 1983.

[37] S. Kirkpatrick, "Optimization by simulated annealing: Quantitative studies," *Journal of Statistical Physics*, vol. 34, no. 5/6, pp. 975–986, March 1984.

[38] D. Geman and S. Geman, "Stochastic relaxation, gibbs distributions, and the bayesian restoration of images," *IEEE Transactions on PAMI*, no. 6, pp. 721–741, 1984.

[39] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed.   Cambridge, United Kingdom: Cambridge University Press, 1992, pp. 444–455.

[40] (2007, August) Melles Griot optics guide: Laser modes. [Online]. Available: http: //www.mellesgriot.com/products/optics/gb_3_1.htm

[41] (2007, August) Canon CMOS sensor technology: Noise reduction technology. [Online]. Available: http://web.canon.jp/imaging/cmos/technology-e/noise_reduction.html

[42] (2007, August) EF 180mm f/3.5l macro USM. [Online]. Available: http://www.usa.canon. com/consumer/controller?act=ModelInfoAct&fcategoryid=155&modelid=7324

[43] (2007, August) Understanding the MTF graph. [Online]. Available: http://www.photodo.com/ topic_108.html

[44] J. E. Boff, Kenneth R. ; Lincoln, *Engineering Data Compendium: Human Perception and Performance*.   HARRY G ARMSTRONG AEROSPACE MEDICAL RESEARCH LAB WRIGHT-PATTERSON AFB OH, 1988.

[45] E. Hecht, *Optics*, 2nd ed.   Addison Wesley, 1987.