Design of the first optical system for real-time tomographic holography (RTTH)

John M. Galeotti[a], Mel Siegel[a], Richard D. Rallison[b], and George Stetten[a,c]

[a]The Robotics Institute, Carnegie Mellon University, Pittsburgh, PA 15213, USA;
[b]Wasatch Photonics, 1305 North 1000 West #120 Logan, Utah 84321, USA;
[c]Department of Bioengineering, University of Pittsburgh, Pittsburgh, PA 15261, USA

ABSTRACT

The design of the first Real-Time-Tomographic-Holography (RTTH) optical system for augmented-reality applications is presented. RTTH places a viewpoint-independent real-time (RT) virtual image (VI) of an object into its actual location, enabling natural hand-eye coordination to guide invasive procedures, without requiring tracking or a head-mounted device. The VI is viewed through a narrow-band Holographic Optical Element (HOE) with built-in power that generates the largest possible near-field, in-situ VI from a small display chip without noticeable parallax error or obscuring direct view of the physical world. Rigidly fixed upon a medical-ultrasound probe, RTTH could show the scan in its actual location inside the patient, because the VI would move with the probe. We designed the image source along with the system-optics, allowing us to ignore both planer geometric distortions and field curvature, respectively compensated by using RT pre-processing software and attaching a custom-surfaced fiber-optic-faceplate (FOFP) to our image source. Focus in our fast, non-axial system was achieved by placing correcting lenses near the FOFP and custom-optically-fabricating our volume-phase HOE using a recording beam that was specially shaped by extra lenses. By simultaneously simulating and optimizing the system's playback performance across variations in both the total playback and HOE-recording optical systems, we derived and built a design that projects a 104x112 mm planar VI 1 m from the HOE using a laser-illuminated 19x16 mm LCD+FOFP image-source. The VI appeared fixed in space and well focused. Viewpoint-induced location errors were <3 mm, and unexpected first-order astigmatism produced 3 cm (3% of 1 m) ambiguity in depth, typically unnoticed by human observers.

Keywords: RTTH, RTTR, HOE, AR, holographic optical element, augmented reality, display, virtual image, in situ, tomographic

1. INTRODUCTION

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 (RT) guidance to mentally integrate two disparate frames of reference. The difficulty of achieving 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.

1.1 Previous Augmented Reality Techniques

All of these application areas require human interaction with an environment that must be scanned and visualized in RT. 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.[1–8] These techniques fall into the broad category of *augmented reality* (AR). AR enhances what is predominantly a real scene with virtual objects. For example, in the case of guiding needle biopsy, both the patient and needle are real. The "virtual" addition to the scene may consist of a RT ultrasound image being used to guide the biopsy procedure or an MR or CT image acquired previously but appropriately registered to the patient in RT. AR seeks to project such images *in situ* (in the location from which they were scanned).

The most common method for displaying an in-situ object in current commercial and research AR systems is to present a separate viewpoint-dependent rendering to each eye, such that stereoscopic vision may determine depth.[3,4,9] Doing so requires both RT computation to generate an appropriate rendering for each eye, as well as some sort of tracking to determine each eye's 3D position. Many such systems make use of either a head-mounted display (HMD) or special, e.g. polarized, glasses. All of these systems have a number of difficulties, some of which are technical hurdles but others of which are intrinsic, such as conflicting depth cues (e.g. accommodation and ocular convergence).[10]

Optical systems that generate in-situ images without tracking or a head-mounted display offer a solution to many of these difficulties. The perceived 3D location of each in-situ point in these systems is essentially independent of viewpoint, allowing natural depth perception of the 3D scene. These systems are termed *autostereoscopic*.[11] Such in-situ systems may either project a *holographic image* or they may project a true optical virtual image (VI), in the latter case either by means of a semi-transparent mirror or, as presented here, by means of a *holographic optical element* (HOE). In contrast to a generally static holographic image, an HOE is a hologram of an optical *system*, rather than of a static object. As an optical system, a HOE can produce a desired VI projection from a dynamic image source, such as an LCD or other spatial light modulator (SLM). A true optical VI occurs is the result of the *apparent* in-focus convergence of light rays, such as a reflection in a mirror or the magnified image produced by a magnifying glass. True optical VIs are naturally autostereoscopic, and a perfectly focused true optical VI, such as a reflection from a flat mirror, is optically indistinguishable from a real object. (Hereafter, "VI," when not prefixed by "pseudo," will mean a true optical VI.)

The aforementioned application areas' requirement for physical interaction has two important consequences for any relevant VI-based AR system. First, the VI must be located close to the human observer, optically in the *near field*. Second, the VI must be *stable* in space. That is, the perceived location of each point in the VI must be essentially independent of viewpoint. Near-field visualization is much more difficult than far-field visualization "at infinity." Far-field visualization, as employed in existing "head-up" holographic displays (e.g., for fighter pilots), is significantly simplified because translational movement by the observer does not change the angles at which the image points' light rays strike the observer.

The stability of a VI is actually just a measure of focus. If a VI is perfectly focused, then it will appear stable in space because all light rays from a single point (i.e. a "pixel") in the VI will appear to come from a single point in space (by virtue of being in focus). On the other hand, if a VI is not perfectly focused, then the degree of perceived blur will depend on the size of the viewing aperture used. The image may appear only slightly blurred when viewed over the aperture of an eye's pupil, but it would appear more blurred when viewed over a larger aperture, such as one encompassing a range of viewpoints. Thus, any instability in space is actually just lack of focus when observed across a viewing aperture that encompasses the range of all possible viewpoints. Regrettably, optical aberrations are inherent to any non-trivial optical component (e.g., something other than a perfectly flat front-surface mirror), whether refractive such as a lens or diffractive such as an HOE. Thus, no HOE is capable of projecting a perfect VI, and there will always be some degree of optical aberration that results in lack of focus (but in a well designed system the lack of focus will be minimized to an acceptable level).

There are several existing AR techniques for projecting a near-field in-situ autostereoscopic VI in RT, with Stetten's sonic flashlight being the most recent.[6,12,13] All of these preexisting VI-based AR techniques utilize RT tomographic reflection (RTTR). As shown in Figure 1, RTTR utilizes a half-silvered (semi-transparent) mirror to project an in-situ VI from a RT image source (such as an LCD or OLED display). By rigidly connecting

the autostereoscopic visualization device to the scanning device, the need for cumbersome tracking equipment is altogether avoided. However, RTTR is typically restricted to displaying a 2D manifold (albeit corectly located and perceived in 3D space) of the same size, shape, and mirrored position as the display source. Accordingly, it can be difficult or impossible to construct a hand-held RTTR system that is capable of guiding deeper procedures such as liver biopsy or amniocentesis but is not unwieldy and that does not block physical access to the patient by placing a mirror in the way of surgical tools. Given the availability of portable scanning technologies, such as ultrasound, it is especially desirable to have an AR system with a hand-held form factor.

1.2 Real-Time Tomographic Holography

There is a need for an AR method that is not only capable of RT, in-situ autostereoscopic visualization of "large" objects in the near field, but is also capable of doing so from a hand-held or wearable device without using head tracking, blocking the operator's vision, or placing physical equipment in the way of tools (e.g., weapons or surgical tools). Here we present the first implementation of our alternative approach, originally proposed (but unimplemented) in Ref. 14.

We have named our approach RT tomographic holography (RTTH). RTTH is derived from, and is similar to, RTTR, but RTTH has the added capability of generating a large in-situ VI from a small RT image source (e.g., an LCD back-lit by a laser), and RTTH imposes fewer design constraints than RTTR on the position of the image source relative to the VI. The goal of RTTH is to holographically place a viewpoint-independent, RT, 2D-manifold VI precisely into the actual scanning location (and orientation) in the 3D world, enabling natural hand-eye coordination to guide invasive procedures without requiring tracking or a head-mounted device. In an RTTH system (as depicted in Figure 2 and Figure 3), the VI is viewed through a narrow-band Holographic Optical Element (HOE) with built-in power that generates the largest possible near-field, in-situ VI from a small display chip without noticeable parallax error or obscuring direct view of the physical world. Rigidly fixed upon, e.g., a medical-ultrasound probe, RTTH could show the RT, i.e. live, ultrasound scan in its actual location inside the patient, because the VI would move with the probe. As with RTTR, the use of a 2D image source unfortunately restricts an HOE-based system such that, at any given moment, it can only visualize data lying on a 2D manifold; still, the projected tomographic data can be correctly located and perceived in 3D space, and different manifolds of data may be examined in temporal sequence by physically moving the device (and thus the location of the in-situ visualization).

In general, because RTTH uses an HOE in lieu of a semi-transparent mirror, the VI is neither constrained to match the size nor shape of the actual display, nor to lie in a specific location or orientation. The latter lack of constraint is because an HOE, in addition to having built-in power, can also bend light rays like a diffraction grating, allowing the LCD to be placed out of the way. Thus, a hand-held RTTH device could project a large VI while leaving room for long tools without being unwieldy. The use of a narrow-band HOE allows the HOE to appear very transparent, with both the patient and the VI clearly visible. This is important because many invasive procedures require a clear view of the exterior of the subject under examination, as well as the virtual objects within it. Because an HOE is the only known optical element to possess all of these desirable capabilities, RTTH may be ideal for guiding deep clinical procedures such as liver biopsy and amniocentesis.
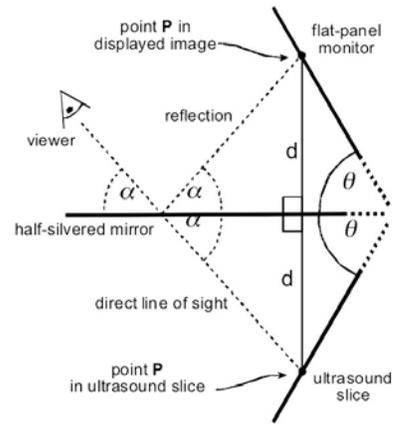

Figure 1. RTTR Configuration: A half-silvered mirror bisects the angle between the in-situ virtual image (VI) and the flat-panel monitor. The VI is coincident with the scanned data, e.g. ultrasound slice, within the patient. Point P in the VI and its corresponding location on the monitor are equidistant from the mirror along a line perpendicular to the mirror (distance = d). Because the angle of incidence equals the angle of reflection (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.
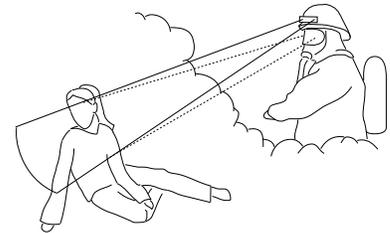
Figure 2. An RTTH display projects a nearly stable virtual image at the actual location of the data in real time, as shown here displaying infrared laser rangefinder data through smoke using an HOE in a fireman's visor.
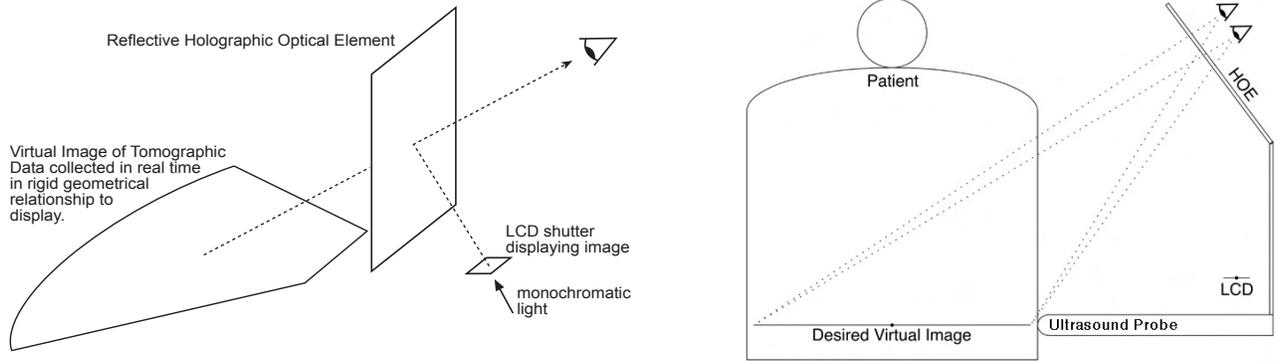
Figure 3. An HOE can be used to project a nearly stable autostereoscopic virtual image (VI) at the actual location of the data in RT. The size, shape, and position VI need not be identical to that of the display source. The diagram on the left utilizes a reflective HOE, while the diagram on the right utilizes a transmissive HOE.

2. METHODS

2.1 Design Problem

For the present work our primary goal was to successfully demonstrate each of the key advantages of RTTH by designing, building, and testing an RTTH system to project an unoccluded VI in the *near-field*, *conveniently positioned* relative to the source LCD, *larger* than the source LCD (magnified), and well *focused*, not only appearing sharp, but also being *stable* in space (i.e., perceived location essentially independent of viewpoint).

Designing such a system is non-trivial. Producing a near-field VI larger than the source LCD requires that the RTTH system have a relatively short effective focal length (EFL), and yet the requirement for stereoscopic human viewing imposes a hard limit on the minimum size of the HOE viewing aperture. Furthermore, conveniently positioning the VI relative to the source LCD to allow room for operating tools, e.g. as depicted in Figure 3, requires a *non-axial* optical design. Finally, because the HOE is large and the VI is near, the resultant range of viewpoints is relatively large as well. Thus, we were tasked with designing a well-focused, fast, non-axial optical system with a wide viewing angle.

2.2 Design Strategy

To increase our chance of success, we designed the image source along with the system-optics, allowing us to ignore both planer geometric distortions and field curvature. The former were compensated by using *RT pre-processing software*, and the latter by attaching a custom-surfaced fiber-optic-faceplate (FOFP) to our image source. The ability to ignore these two optical aberrations allowed us to further minimize the remaining aberrations.

To further simplify our design task, we minimized the required EFL by opting to project a planar VI of modest size (104x112mm) at a distance of 1 m from the HOE. Using a 1 m distance also simplified our design by reduced the range of viewpoints for seeing the VI through the HOE. We also minimized the aperture of our system by specifying a relatively small (for stereoscopic viewing) 5x5 in square HOE. Due to the human eye's peak sensitivity to green light, we specified a 532 nm operating wavelength for our HOE to allow the projection of a bright VI when illuminating our 19x16 mm LCD image source with a relatively safe (i.e., low power) class IIIA laser module.

There are many methods available for the specification, simulation, and manufacturing of an HOE. Ultimately, we specified the use of a volume-phase HOE (VHOE) in order to make the HOE appear as transparent as possible. A VHOE operates by having a different refractive index (and thus optical path length) at every microscopic point on a very thin layer of a special material, which is sandwiched between two glass sheets. A large VHOE is customarily produced photographically, by recording the interference pattern from the simultaneous exposure of two different wavefronts from laser source(s); typically these wavefronts are spherical, and the resulting VHOE thus records a zone plate. Wasatch Photonics, Inc. (WP) manufactured our VHOE, making use of proprietary techniques and a dichromated gelatin (DCG) recording medium to produce an evenly exposed VHOE with high

diffraction efficiency (allowing for a bright VI) and low loss (high transparency), all of which are important for an RTTH system.

We were able to gain additional design freedoms by using non-spherical wavefronts when recording the VHOE. Doing so required that we optically fabricate our VHOE using recording beam(s) that were specially shaped by extra lenses. These extra lenses are known as *construction optics*, since they are only used to construct the HOE, and are not used during HOE playback. Construction optics are located between the spherical wavefront emitters (known as *construction points*) and the VHOE recording surface. The ability to include arbitrary construction optics is a double-edged sword; on one hand they provide additional design flexibility beyond the relatively limited positions of the two construction points, but on the other hand systems containing them can be complex to design and optimize, because there is no a-priori constraint on the number, type, or arrangement of construction optics. We simultaneously simulated and optimized the system's playback performance across variations in both the RTTH device (the "playback" system) and the HOE-recording system. The additional design freedoms enabled by such simultaneous optimization of multiple optical systems proved critical in achieving an acceptably focused system.

2.3 Design Procedure

Following standard practice,[15] we designed our HOE-based system using optical simulation (ray tracing) and optimization software (specifically, Zemax EE). It is perfectly valid to trace rays starting at a point in the VI rather than at a point in the LCD, and this was desirable for us since the LCD presented a single surface as opposed to the VI which existed in 3-space when not perfectly focused. When tracing from a point in the VI through the HOE, the simulation needed 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 VI, likewise all the rays from a point in the VI back-trace to converge at a single source point in the LCD, and the focus of an optical system can be measured by examining the blur at either end of the system.

We jointly optimized the HOE's phase function (which fully describes the refractive index across the HOE's surface) with the remainder of the device's optical components, including their physical layout. A complete RTTH system requires, at minimum, an HOE and a display source, such as an LCD. As previously stated, our display source consisted of an LCD paired with a fiber-optic face plate. Each pixel of the display source must emanate diffused light at the HOE's operating wavelength, which may be accomplished by use of an expanded laser beam for illumination paired with an optical diffuser (the FOFP served as our diffuser).

In order to help correct for the HOE's optical aberrations, it proved necessary to place additional corrective optics between the display source and the HOE. To avoid either blocking the operator's vision or placing physical equipment in the way of surgical tools, the corrective optics had to be kept out of the way. The properties of these corrective optics were jointly optimized with the rest of the optical system. We ultimately utilized two additional lenses, not only to correct for aberrations, but also to off-load some of the requisite optical power from the HOE.

2.4 Final Design

Our final design contained only one construction lens, and it simulated well. Its playback optics included two commercially readily available lenses, each with one flat surface, and a custom FOFP. The playback optical layout of our design is shown in Figure 4, and our design's construction optical layout is shown in Figure 5.

The basic geometry of the playback lenses was optimized in conjunction with that of the HOE's construction lens and construction points, but the playback 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. These two playback lenses are separated by 16.67 mm and are at a small angle and offset relative to the "local" optical axis between the center of the LCD and the center of the HOE. Likewise, the construction design was also re-optimized to use a very similar commercially off-the-shelf available lens from Edmund Scientific, part 45716, which is a 75 mm focal length BK7 PCX lens with a radius of 38.76 mm.
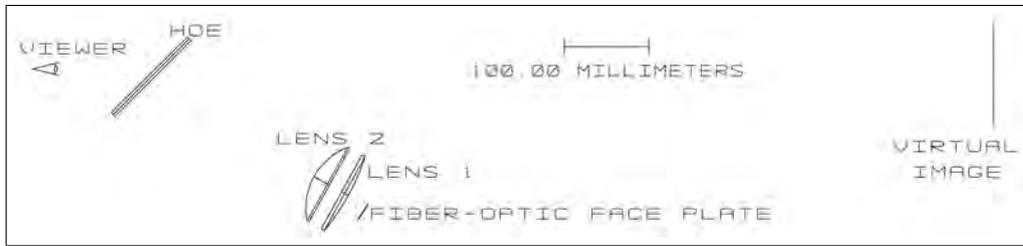
Figure 4. Final playback layout of our RTTH system
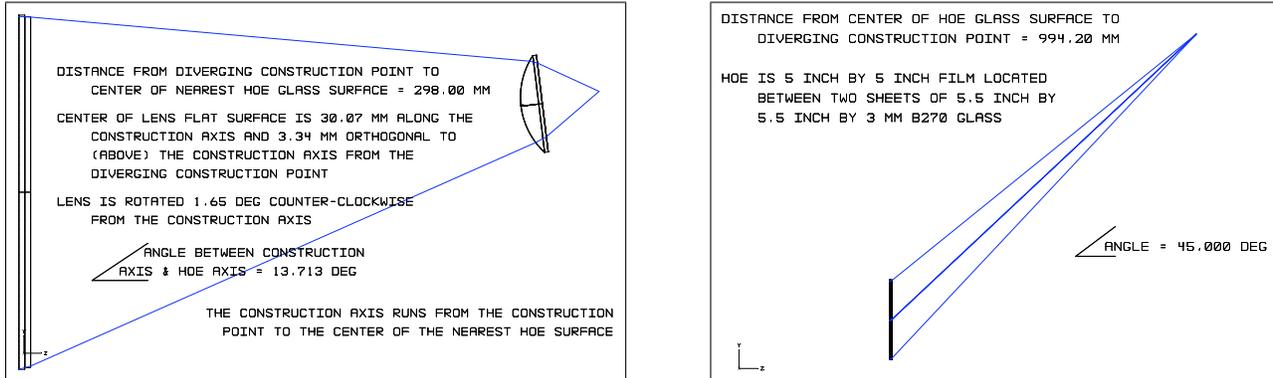
Figure 5. Construction optical layout of our design, annotated with design specifications. Both construction points are diverging. The left and right diagrams are scaled differently to maintain readability of the annotations; in both diagrams the HOE is 5 inch (127 mm) square.

2.5 Design Analysis

A reverse-traced, full field spot diagram analysis of our design is shown in Figure 6. This diagram was generated by tracing rays in reverse from 12 color-coded points in the VI (104 mm wide by 112 mm tall, located 1000 mm from the HOE) to their intersection with the LCD surface (19.3 mm wide and 15.5 mm tall). As can be (approximately) observed, our design achieved *a horizontal magnification factor of 5.39 and a vertical magnification factor of 7.23*. 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 VI) would be coincident. The most significant defocusing optical aberrations readily visible 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.

Some simple analysis of the best and worst performing image points in the full field spot diagram were used to predict the stability of the VI. The best-focused image point traced in Figure 6 is the point at the center of the LCD. The image points for the bottom corners of the LCD 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 VI, the locus of points on the FOFP that would be seen from all possible viewpoints. By measuring the horizontal and vertical point spreads across the surface of the FOFP/LCD as shown in Figure 6, 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 VI. Two such 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 6 also illustrates the degree to which rays traced through most viewpoints are clustered in the center of their spot diagrams.

Over most viewpoints, the center of the VI was predicted to be stable to within 0.8 mm (1.5mm over all viewpoints), and the least stable regions of the VI were predicted to be stable to within 1.5mm over most viewpoints (4.2 mm over all viewpoints). Thus, if the worst-offending corners and/or edges of the HOE were masked, then the entire system would be expected to be stable to within 1.5 mm. For any future implementation
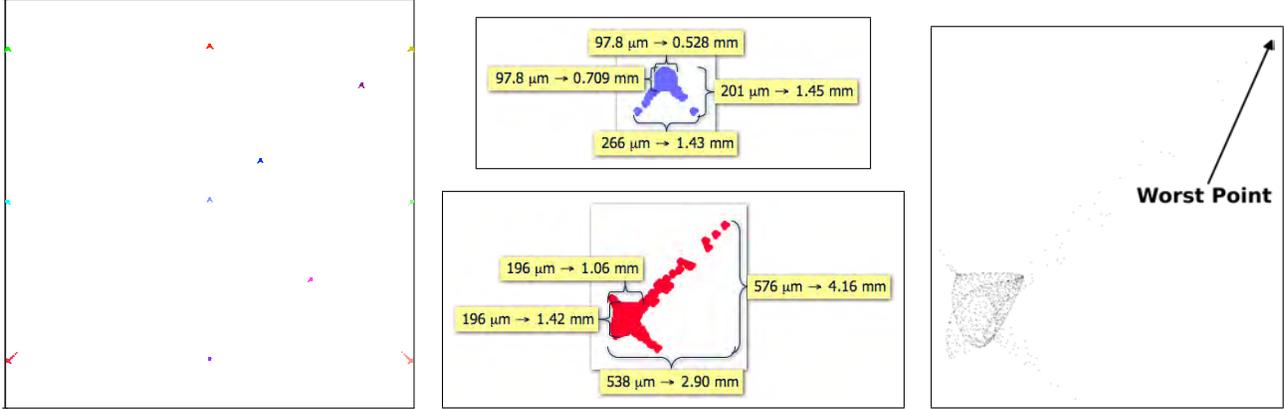
Figure 6. **Left:** Full field spot diagram for our design. The square box drawn around the spot diagram is 19.3 mm on each side, equal to the length of the LCD. Of the 12 points traced, the center point is the best focused, and the two symmetric points in the bottom corners are the worst focused. **Center:** Analysis of the best (shown on top) and worst (shown on bottom) image points. By measuring the horizontal and vertical point spreads across the surface of the FOFP, 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 VI. 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. **Right:** The worst spot diagram, rendered using much smaller dots to reveal the densities of the rays' intersection points.

of our design properly calibrated to its scanner, the greatest anticipated error would then be 0.75 mm in any direction.

3. EXPERIMENTS

3.1 Experimental Apparatus

3.1.1 Source image formation equipment

Because our RTTH optical design neither simulated nor specified a method of applying the LCD image to the FOFP, full implementation of our RTTH system required that we experiment with various types of source image formation equipment, such as different laser sources, beam spreaders, etc. Ultimately, we made use of a laser module and a beam expander (to achieve a larger diameter collimated beam) to direct a wide laser beam through the entire active area of our LCD display to the FOFP. All of this equipment, including the LCD and FOFP, was held by optical mounts attached to precision positioning equipment.

The laser module was 532 nm (green), class IIIa ($< 5\text{mW}$), and diode pumped solid state. The laser beam was directed through a 15x beam expander to increase its small diameter to about 25 mm before the beam was redirected to a small, high-resolution LCD. The active area of the LCD, from Kopin Inc., Taunton, MA, measured 24.6 mm diagonally, with a monochrome resolution of 1280x1024.

Ideally, it would have been possible to use the LCD in front of the expanded laser beam to project the LCD image onto the back surface of the fiber optic face plate. Unfortunately, the small pitch of the LCD pixelation caused it to also act as a two-dimensional diffraction grating when illuminated by the laser. The diffraction pattern maxima were calculated to be 1.2 ° apart, which matched the observed separation in the two-dimensional lattice of maxima. Accordingly, we placed the FOFP in direct contact with the LCD, to minimize the spread of the "ghosted"/diffracted images. Even so, the small thickness of the LCD glass resulted in a residual diffraction-induced one-pixel-radius blur, which we considered acceptable given the high resolution of the LCD compared to the ultimate resolution of the RTTH system. In addition to curving the LCD source image, the FOFP also provided necessary diffusion, allowing light from each pixel to reach the entire HOE's surface.
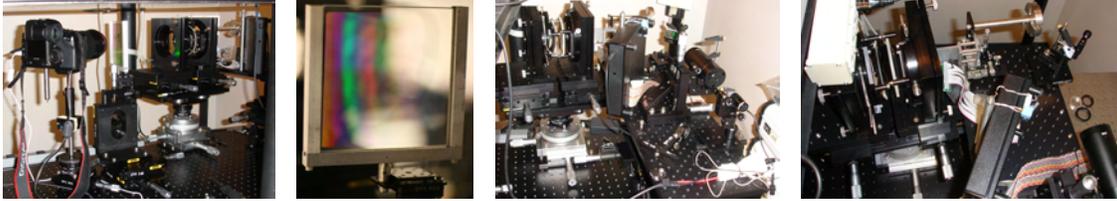
Figure 7. Left: Positioners holding the camera, HOE, lenses, and part of the source image formation equipment. The HOE, located in the center of the first figure, is shown from a different perspective in the center-left figure. The two right figures show the positioners holding the lenses and all of the source image formation equipment, viewed from two different perspectives.

3.1.2 Other equipment

Our experimental data were gathered by a computer-controlled, 12.8 megapixel digital SLR camera (Canon, Inc. model *EOS 5D*) which incorporates special hardware to minimize both fixed-pattern and random noise.[16] The primary lens used with the camera was Canon's model *EF 180mm f/3.5L Macro USM*. The camera and all of the optics were positioned using precision translation and rotation stages. The layout is shown in Figure 7. Note that unlike the eventual 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 following experiments the optical plane was oriented horizontally, parallel to the top of the optical table. Also, unlike a fully-implemented hand-held system, our experimental layout did not contain an actual scanning device, and so it was unnecessary to register (align the mapping of) a scanning device to the VI.

3.2 Validation

Upon completion of alignment, a quick visual check confirmed that the VI appeared to be stable, well focused, and properly positioned. The VI did not, however, 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 VI could appear well-focused only in either of the horizontal or vertical directions, but not both, at a given focal distance. The observed optical aberration was a type of *first-order astigmatism*, such as is commonly corrected for in human vision by use of a cylindrical surface. (Most optical systems are rotationally invariant about a single optical axis, and thus suffer from third-order but not first-order astigmatism.)

In light of the astigmatism problem, we developed a validation method based on the use of a movable physical target. The physical target, shown in Figure 8, 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 VI 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 VI. 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 aligned with the physical target is shown in Figure 8.

Figure 8. Physical target used for validation (top left), shown with its mounting (top right) and aligned virtual crosshairs (bottom)

By aligning the virtual crosshairs with the physical target's vertices, it was possible to visually check whether or not the VI 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.) side-to-side or up-and-down; assuming that the VI 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 VI was measured by minimizing motion parallax (to make the VI and the target coplanar), and then measuring how much the virtual crosshairs moved as viewpoint

varied. Likewise, blur in the VI was 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 measurements 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 VI 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 VI plane, and likewise for vertical blur and image stability in the vertically focused VI plane. Furthermore, it was also possible to measure the astigmatism-induced distance between the two focal planes.

All of the images used for blur analysis were taken with the 180 mm lens set to 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.[17, 18] 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.

4. RESULTS AND ANALYSIS

Our system's VI appeared focused and stable (i.e., its perceived location was essentially independent of viewpoint). We successfully designed, built, and demonstrated what we believe to be the first HOE designed for RTTH. Our system successfully qualitatively demonstrated the key advantages of RTTH, projecting a VI in the *near-field*, that appears *larger* than the source LCD, *off-axis* relative to the LCD, and appears both *focused* and *stable*.

Following the validation procedure described in subsection 3.2, we were able to collect quantitative data regarding several aspects of the VI, including its position and orientation, its stability, and its 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 Due to the relatively large distance between the relatively small HOE and the VI, the camera was only able to view the entire VI over a range of 2.5 °horizontally and 2.7 °vertically.

4.1 Position and Orientation

The vertical focal plane of the VI 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 VI and the LCD's "local" optical axis (the axis between the center of the LCD and the center of the HOE) was verified to within 1 °.

We believe that the astigmatism-induced 3 cm gap between horizontal and vertical focus was 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 VI or in reverse to the FOFP's source image). Because the VI is located one meter from the HOE, its light rays travel, to a crude approximation, orthogonal to the VI, and so a 3 cm blur orthogonal to the VI 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 would be expected to have less astigmatism if they project a VI closer to the HOE. Likewise, future RTTH systems that project the VI farther away would be expected to have more "unanticipated" astigmatism, unless additional design steps were taken to separately measure and penalize first-order astigmatism during design optimization.

4.2 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 (as should the right crosshairs), and the observed deviations from this 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. Accordingly, if 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.

Because the VI 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 subsection 2.5. (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.)

If coupled with a scanning device, 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 keeps the entire VI 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 astigmatism acts as a bimodal blur in depth, rather than a source of "instability" in the VI, but the resultant motion parallax could affect its perceived location in all three dimensions.

4.3 Minimum Feature Size

The graphical crosses used to analyze the stability of the VI were also used to measure the minimum projectable feature size in the VI, which corresponds to the size of one LCD pixel, blurred by the LCD/FOFP interface (see subsubsection 3.1.1) and then magnified and blurred by the optics of our designed system. Minimum feature size was measured using a one-pixel thick line, oriented either horizontally or vertically. Figure 9 shows how line thickness was measured. Within the basic range of viewing angles, 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. As would be expected, these measurements are less than the observed drift ranges above, and so blur is unlikely to be a limiting factor in any RTTH production implementation based on our present design. At the extreme far-left and far-right viewpoints, where only two of the crosshairs were barely visible, the maximum observed horizontal blur modestly increased the minimum feature size to 0.97 mm, significantly less than the horizontal drift range at these extremes.
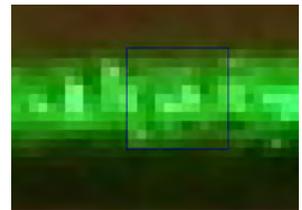
Figure 9. Minimum Displayable Line Width: The blue box indicates the pixels that were counted to measure line thickness.

4.4 Demonstration of Pre-Recorded Ultrasound Projected In Situ

Actual in-situ projection of fetal ultrasound is shown in Figure 10. For this demonstration, we played back previously recorded ultrasound. However, had we attached an actual ultrasound machine to our device we could have projected RT ultrasound data in situ. All that would have been necessary to assure correct registration between the VI and the patient's internal anatomy would have been to rigidly hold the ultrasound scanning probe in correct alignment with the stable VI, e.g. as depicted in Figure 3.
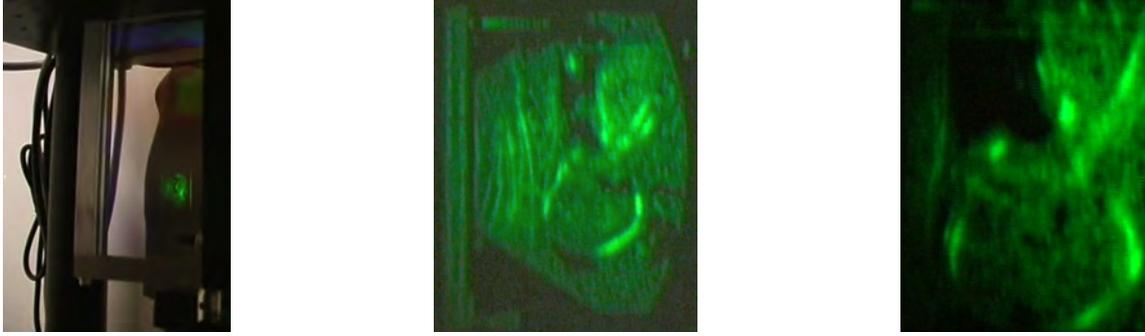
Figure 10. Photos showing fetal ultrasound projected inside the mother. The images appear to have higher resolution in person, due primarily to the larger dynamic range and smaller pupil size of the human visual system. The left photograph shows the ultrasound image floating inside the mother, as viewed through the HOE. The other photographs are "zoomed in," showing part of the fetus, oriented facing left with the head down. Some of these are still-shots from a video we recorded showing in-situ autostereoscopic visualization of fetal movement inside the mother.

5. CONCLUSIONS

We have successfully designed, built, and tested the first RTTH optical system. It displays a stable VI at 1 m. Although not designed for any particular RTTH application, the HOE met the most critical and unique design challenges that RTTH imposes, *successfully demonstrating the near-field projection of a stable, off-axis magnified VI.* Accordingly, we have demonstrated the chief advantages of RTTH:

- Essentially viewpoint-independent direct viewing of an in-situ VI

- Display of a VI physically larger than the source image used to project it

- Lack of any imposed physical objects between the HOE "viewport" and the VI, both allowing room for long tools to operate and providing an unobstructed view of reality in the region surrounding the VI

Our initial validation of the RTTH system revealed an unexpected first-order astigmatism that produced a 3 cm (3% of 1 m) ambiguity in depth, which was not perceived by human observers. Aside from the aforementioned astigmatism, our RTTH system performed well, with both resolution and stability on the order of 1 mm when viewed from "normal" viewpoints (viewpoint-induced location errors were <3mm over an extreme range of viewpoints), and with the geometric measurements of the VI 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, such as liver biopsy and amniocentesis. Such a new system would display a larger VI at a distance of significantly less than 1 m.

ACKNOWLEDGMENTS

REFERENCES

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

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

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

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

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

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

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

[8] Stetten, G. and Chib, V., "Overlaying ultrasound images on direct vision," *Journal of Ultrasound in Medicine* **20**(1), 235–240 (2001).

[9] Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., and MacIntyre, B., "Recent advances in augmented reality," *IEEE Computer Graphics and Applications* **21**(6), 34–47 (2001).

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

[11] Dodgson, N. A., "Autostereoscopic 3d displays," *Computer* **38**, 31–36 (August 2005).

[12] Hofstein, S., "Ultrasound scope," (April 1980). Patent 4200885, Filed Feb. 2 1978.

[13] Stetten, G., "System and method for location-merging of real-time tomographic slice images with human vision," (July 2003). Patent 6599247, Filed Oct. 11 2000.

[14] Nowatzyk, A., Shelton, D., Galeotti, J., and Stetten, G., "Extending the sonic flashlight to real time tomographic holography," (2004). Workshop AMI-ARCS 2004, held in conjunction with MICCAI 2004 September 30th, 2004, Rennes (France).

[15] Rallison, R. D. and Arnold, S. M., "Wavelength compensation at 1.064 microns using hybrid optics," in [*SPIE*], *Diffractive and Holographic Optics Technology III, San Jose CA* **2689**, 267–273 (Feb 1996).

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

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

[18] Hecht, E., [*Optics*], Addison Wesley, 2 ed. (1987).