

Extending the Sonic Flashlight to Real Time Tomographic Holography

Andreas Nowatzky^{1,2}, Damion Shelton¹, John Galeotti¹, George Stetten^{1,3,4}

¹ The Robotics Institute, Carnegie Mellon University,

² Department of Electrical and Computer Engineering, Carnegie Mellon University

³ Department of Bioengineering, University of Pittsburgh,

⁴ University of Pittsburgh Medical Center

<http://www.vialab.org>

Abstract. Real Time Tomographic Reflection (RTTR) permits *in situ* visualization of tomographic images, enabling natural hand-eye coordination to guide invasive procedures. The method as originally developed uses a half-silvered mirror to merge the visual outer surface of the patient with a simultaneous scan of the patient's interior. A viewpoint-independent virtual image is reflected precisely into the correct location within the patient. When applied to ultrasound, we call the resulting device the *sonic flashlight*. In this paper we present the first published description of a derivative method that replaces the half-silvered mirror with a Holographic Optical Element (HOE). The new technique, called *Real Time Tomographic Holography* (RTTH), has the advantage over the mirror-based system that the virtual image is no longer required to match the size and shape of the actual display. We are currently in the first year of an NSF grant from the Robotics and Human Augmentation Program (CISE/IIS) to develop prototype RTTH systems. Although at the time of this writing our systems are not yet functional, we consider the concept itself of potential importance, and wish to report it, along with our current progress towards implementation, to the Augmented Reality community.

1 Introduction

In the current practice of medicine, images are routinely acquired by ultrasound, computerized tomography (CT), magnetic resonance imaging (MRI) or other modalities. These images are viewed on a film or screen, rather than by looking directly into the patient. A number of researchers have worked to develop more natural ways to merge images with the perceptual real world.¹⁻⁴ We have previously reported the concept of *Real Time Tomographic Reflection* (RTTR), and applied it successfully to ultrasound.^{5,6} Conventional ultrasound produces a *tomographic* slice within the patient representing a set of 3D locations that lie in a plane. The image of that tomographic slice, displayed on a flat panel monitor at its correct size, can be reflected 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 accurately superimposed within the patient, independent of viewer location. The reflected image is truly occupying its correct location and does not require any particular viewpoint to be correctly perceived.

To accomplish RTTR, certain geometric relationships must exist between the slice being scanned, the monitor displaying the slice, and the mirror. As shown in Figure 1, the mirror must bisect the angle between the slice and the monitor. On the

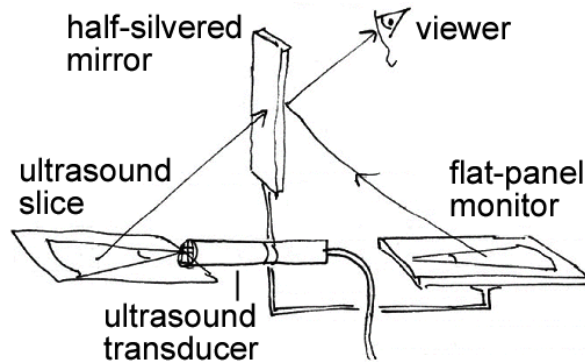


Fig. 1 Schematic representation of an ultrasound RTTR device, or *sonic flashlight*. A flat-panel monitor and an ultrasound transducer are placed on opposite sides of a half-silvered mirror to create an *in situ* virtual image.

monitor, the image must be correctly translated and rotated so that each point in the image is paired with its corresponding point in the slice. By fundamental laws of optics, the ultrasound image will appear at its correct physical location, independent of viewer location.

Figure 2 shows a recent prototype of the sonic flashlight, whose images come from a

standard ultrasound machine conducting a conventional scan along the axis of the ultrasound transducer (B-mode image). The flat panel monitor is likewise mounted along the axis of the transducer with the mirror bisecting the angle between the transducer and the flat-panel monitor. The mirror is mounted perpendicular to the axis, so that the image on the flat panel monitor is reflected to occupy the space being scanned by the ultrasound transducer. The device is designed to be held in one hand while performing an invasive procedure with the other.

In Figure 3, the sonic flashlight is shown in use. A human hand is seen with the transducer pressed against the soft tissue between the thumb and index finger. While not a common target for clinical ultrasound, the hand was chosen because it clearly

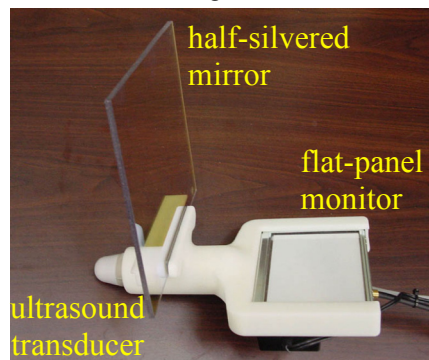


Fig. 2 Prototype of the sonic flashlight, showing the relationship of the ultrasound transducer, the half-silvered mirror and the flat-panel monitor.



Fig. 3 Photograph from the viewpoint of the operator, showing a scan of the hand using the sonic flashlight. The reflected ultrasound image is merged with the direct visual image.

demonstrates successful alignment. The external surfaces of the hand are located consistent with structures within the ultrasound image. The photograph cannot convey the strong sense, derived from stereoscopic vision, that the reflected image is located within the hand. This sense is intensified by head motion because the image remains properly aligned from different viewpoints. To one experiencing the technique in person, ultrasound targets within the hand are clearly accessible to direct percutaneous injection, biopsy or excision.

2 Extension to Holography

A significant limitation of the half-silvered mirror in the sonic flashlight is that the virtual image must be the same size as the actual display. For guiding procedures at relatively shallow depths, such as venous access in the arm, this is not a problem. A hand-held sonic flashlight with a relatively small flat-panel display can be easily manipulated. But for greater depths, a larger display is needed, making the device unwieldy. Replacing the mirror with a holographic optical element provides a solution. As shown in Figure 4, the basic configuration for a Real Time Tomographic Holography (RTTH) display consists of a source of monochromatic light (a laser), which passes through an LCD shutter similar to those found in most laptop computers. 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. With the holographic version, the virtual image produced by the holographic optical element can be much bigger than the actual display, allowing the use of miniature LCD shutter, instead of a full-size flat panel monitor. A hand-held device producing a large virtual image reaching deep within the patient could still be lightweight and

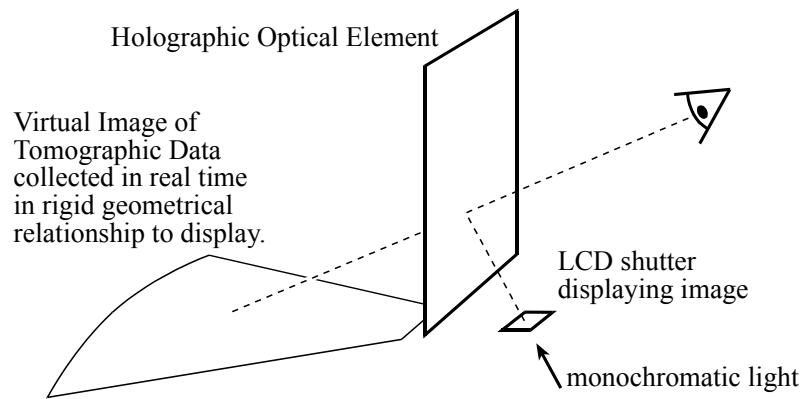


Fig. 4 Basic configuration of a display using Real Time Tomographic Holography. The LCD shutter creates a real time image from laser light. Each pixel is mapped by the Holographic Optical Element (HOE) into a virtual image whose size and shape is determined by the particular HOE. The virtual image is independent of viewer location.

manageable, facilitating ultrasound guidance for deeper procedures such as liver biopsy or amniocentesis.

Figure 5 illustrates the basic operation of a holographic optical element (HOE), created in this case by photographic means. During its creation, an interference pattern between spherical wavefronts from two point sources of in-phase monochromatic light (**A** and **B**) is captured photographically on the surface of the HOE 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.

Thus the HOE works like a mirror, creating a virtual image. The difference is that, unlike a mirror, the HOE can produce a virtual image whose size and shape differ from the actual image. The actual image, which was generated in the mirror-based system by a flat-panel monitor, is now generated by passing coherent light from a laser through an LCD shutter. The LCD shutter can be much smaller than the virtual image produced. Each pixel in the LCD translates to a pixel in the virtual image that, like the sonic flashlight, is viewpoint independent.

The HOE permits a wide assortment of mappings between pixels in the LCD and corresponding pixels in the virtual image, which present potential advantages over the mirror. The decreased size of the LCD relative to the virtual image can reduce the weight of a hand-held device, and avoid occlusion caused by the actual display in the line of sight. The avoidance of occlusion should be particularly useful in the C-Mode sonic flashlight, a variation on the sonic flashlight that we constructed based on a matrix-array real-time 3D ultrasound scanner.⁷ The design of the C-mode sonic

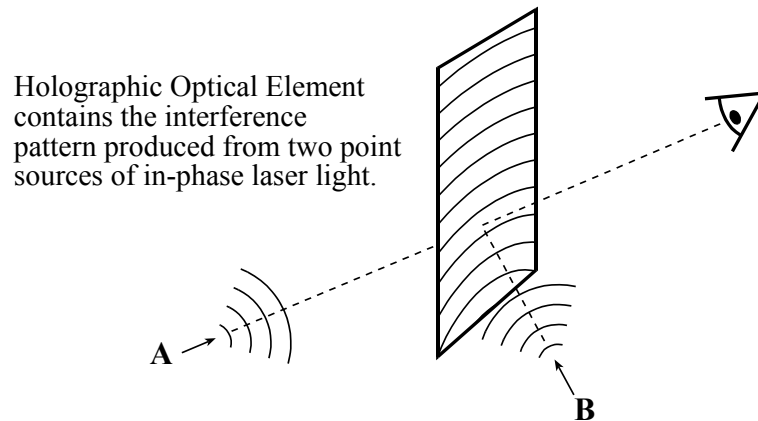


Fig. 5 The interference pattern produced by two sources of laser light, **A** and **B**, is stored in the Holographic Optical Element (HOE). Thereafter, a pixel of laser light emitted from **B** is reflected by that interference pattern in the HOE so as to appear to emanate from **A**.

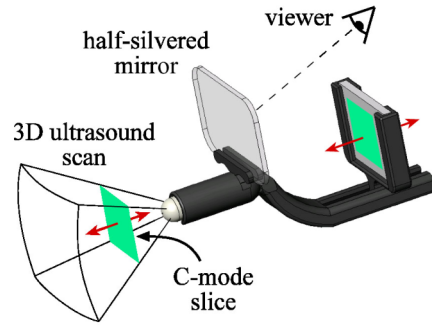


Fig. 6 C-Mode sonic flashlight, displaying a slice parallel to the transducer through a volume in real time. The display, mounted parallel to the mirror, may more easily block the viewer. This problem may be alleviated with an HOE.

flashlight is shown in Fig. 6. A matrix array 3D scanner is used capable of scanning an entire 3D pyramid, from which slices with various orientations may be extracted in real time. In this case, a C-mode slide (parallel to the face of the transducer) at a particular depth is shown. The flat-panel monitor is mounted parallel to the mirror, and is thus more apt to block the view of the operator. The use of the HOE based system with its miniature LCD image source may eliminate such occlusion.

The objective of the present research is to build and test the first prototype RTTH displays, while expanding the theoretical basis of their operation and optimizing their performance and accuracy. We do not expect this research to extend the theory of diffractive optics, but rather simply to apply that theory to the very narrow and specific purpose of building RTTH displays. The objective is to enable stereoscopic vision over a greater range of distance and shape than is possible with a half-silvered mirror. The criterion for success will be how well the virtual image can be located by a human observer. Existing RTTR displays based on the half-silvered mirror will serve as a reference benchmark. We are currently developing RTTH displays using two standard methods of HOE production: the *zone plate* and the *kinofom phase plate*. Sections 3 and 4 of this paper describe our progress using these two methods.

3 Design of a Zone Plate

Zone plates are a type of *binary diffractive element*, planar plates coated with a thin opaque or reflective layer that has either been completely removed or left intact according to the desired diffractive pattern. Because of their simplicity, the cost of binary diffractive elements is modest.

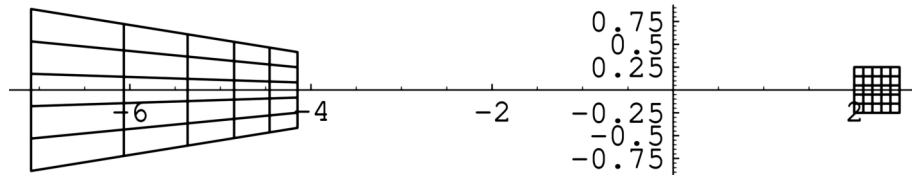


Fig. 7 Zone Plate Image Transformation.

The design of the RTTH display based on a zone plate is fairly straightforward. The dimensions of the primary display are given, and so is the desired range. The

classical lens-maker equation provides the starting point for the geometry: $1/x + 1/y = 1/f$, where f is the focal length, and x and y are the image distances. Figure 7 shows the mapping for a possible RTTH display configuration with a 0.5" x 0.5" primary display (left) and a virtual image (right) starting 4.25" behind the zone plate and extending 3.5". The keystone mapping is caused by the fact that both image planes are parallel to the optical axis, unlike conventional imaging where the image planes are generally perpendicular to the optical axis. The distortions of a zone plate operating at finite conjugates are analyzed in a paper by Young.⁸ Chromatic aberrations not a problem with monochrome imaging. Field curvature is corrected in software by controlling the mapping between the scanner data and the primary display. Of concern are astigmatism and coma because they affect the appearance of a single pixel in the virtual image. It is not possible to correct these in software.

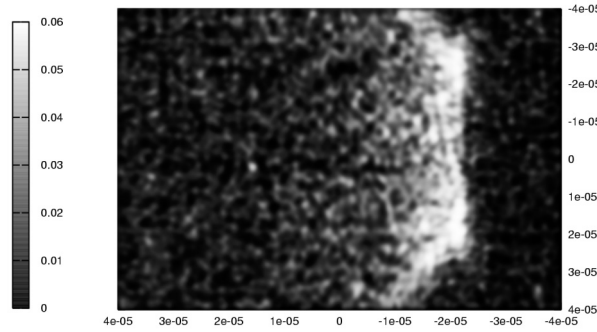


Fig. 8 Simulated point spread function from Zone Plate.

To get an idea of the severity of coma and astigmatism for a zone plate RTTH display, an optical simulation was performed. Figure 8 shows how an ideal point light source in the primary display will appear in the virtual image when observed through the zone plate. The specular nature of the image is due to the fact that a single wavelength was used in this simulation. The point has spread to about 40 μ m by 80 μ m, which would still appear as a point to a human observer. This spreading gets

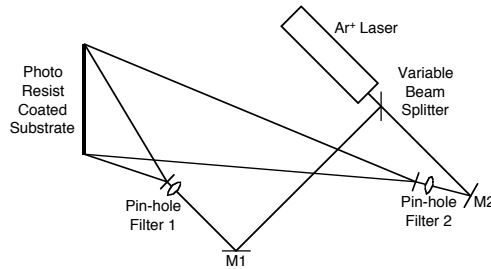


Fig. 9 Optical zone plate generation.

worse near the edge, but will be mitigated, somewhat, because the observer will look only through a small region of the zone plate at any given time. We intend to refine our simulations to take such effects into account, and to evaluate the actual RTTH displays once they are operational, using digital cameras at multiple viewpoints.

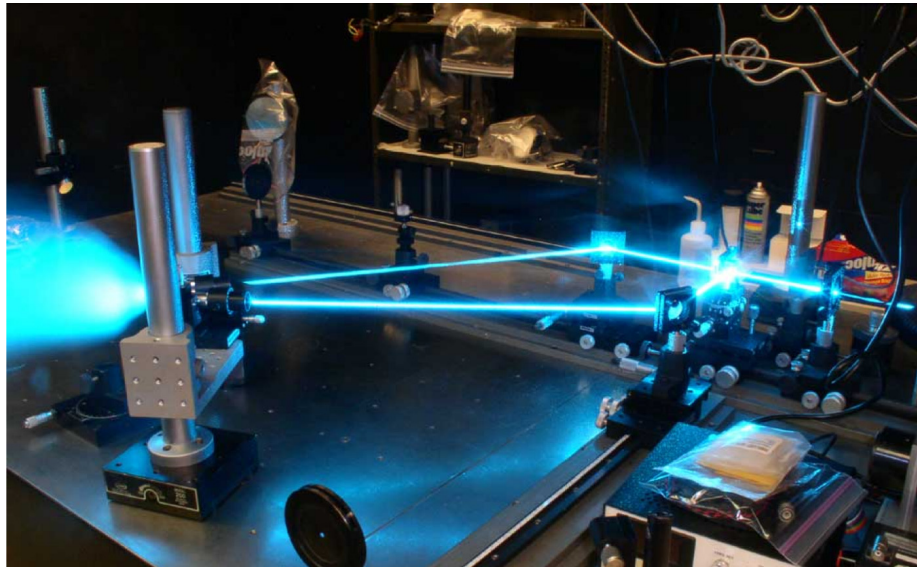


Fig. 10 Apparatus to produce a zone plate. Laser is split (right) and directed through 2 pinholes (left) to create an interference pattern, for projection onto a photographic plate (laser beams enhanced for visibility by long exposure and physically moving a target along the beams).

We have constructed an apparatus for creating zone plates photographically. As shown in Fig. 9, an argon ion laser (6 watt) is directed through a variable (polarizing) beam splitter to provide a pair of in-phase beams whose relative amplitudes can be adjusted. The beams are reflected by two mirrors (M1 and M2) to illuminate two pinhole spatial filters that are positioned along the desired optical axis of the RTTH display. The first pinhole is positioned at the center of where the LCD shutter will be located, while the second pinhole is positioned at the mirror-image of the center of the desired virtual image. The two optical path lengths are matched (not shown in Figure 8). The resulting spherical wavefronts create an interference pattern on a 5" square photographic plate. Figure 10 shows the apparatus in operation. In constructing it, we needed to overcome a number of problems specific to our application, e.g., manufacturing special holders for the pinholes to avoid silhouetting due to the particularly wide angle required of our spherical waves.

The photographic plate will be developed using a modified photolithography process. Unlike conventional lithography processes where the resist is developed under a UV floodlight, our apparatus relies on a comparatively low power laser at a lower (and presumably less-than-optimal) frequency. As a result, we have purchased glass plates coated with several different photoresists of varying absorption properties. Each plate consists of a 5" square soda-lime glass, coated first with a thin layer of chromium and then with one of the photoresists. We have begun to conduct a series of trials to determine the proper exposure time and photoresist type, with the

optimal combination producing the highest contrast interference pattern once the resist is developed. We will then etch the interference pattern into the chrome layer using a wet etch process, characterize the interference pattern using a surface profilometer to determine the spacing and amplitude between ridges of chrome, and examine the interference pattern using a scanning electron microscope to check for the correct duty cycle of the pattern.

A remaining challenge is to achieve uniform exposure, which is critical to the performance of the zone plate. Given the high numerical aperture, the only way realistically to achieve uniformity is by partially exposing the substrate with a mask using a computer-controlled x - y positioning system.

4 Design of a Kinoform Phase Plate

A *kinoform phase plate* differs from the zone plate in that, instead of a binary 0/1 pattern, the kinoform phase plates implement 16 to 256 levels of phase shifts. These are essentially thin surface relieves, where the thickness variation and its lateral distribution are comparable to the operational wavelength. For example the holographic security logos that are commonly embossed into credit cards use this method. By being able to implement arbitrary phase shifts, kinoform phase plates permit a much wider range of transfer functions. Like binary diffractive elements (e.g. zone plates), kinoform phase plates can be produced and replicated easily by lithographic means.

To produce a kinoform plate, we are refining a plan based on large-scale computation of parameters, to be sent out to a commercial house for production. The basic problem is to choose a surface height at an enormous number of discrete locations on the plate, so as to optimize the simultaneous production of many pairs of points in the display and the virtual image. The problem is somewhat simplified by the fact the viewer will only see each point in the virtual image through a small region of the plate at any given time, constrained by the aperture of the pupil. Still, optimization in a reasonable amount of time of such a large set of parameters (we expect tens of billions of parameters) requires some forethought, and we have established the following steps for our approach. (1) Average the zone plates that would correctly project each individual display pixel. This computation is very fast, and will allow us to optimize output image geometry to reduce higher-order terms. Some geometric distortion within the plane of the virtual image can be corrected by warping the display image, so we have a fair amount of tolerance in our choice of output geometry. (2) Decompose the kinoform into (a) the Fourier transform projection to infinity, (b) two 90 degree bends (defraction gratings), and (c) the reverse Fourier transform projection to infinity. These can all be added to produce a single kinoform. The Fourier transform projection to infinity was chosen because it is straightforward and well researched. (3) Optimize our initial parameters using

simulated annealing or other standard method. We have explored a number of possible error metrics to use in this optimization.

To speed optimization of the kinoform, the search space can be restricted to the quantized versions of parametrically defined polynomials. Then, treating the kinoform surface as a continuous polynomial of 2D location, we can evaluate the behavior of the kinoform by ray-tracing a relatively small (on the order of 50 x 50) number of rays from each pixel of the LCD through regularly spaced locations on the kinoform to their focal point in the virtual image. Because of the small number of rays involved, ray tracing is not impractical, and because of the smooth nature of the polynomial representation, a denser sampling of rays is unnecessary. If the optical axis is defined as the x -axis, the height on the kinoform as the y -axis, and position across the kinoform as the z -axis, then the polynomial would define the continuous thickness (along the x -axis) of the kinoform at each point in terms of that point's y - and z -coordinates. An example of such a polynomial is

$$P(y,z) = p_0\sqrt{p_1 + y^2 + z^2} + p_2\sqrt{p_3 + y^2 + z^2} + p_4y + p_5y^2 + p_6y^3 \\ + p_7z^2 + p_8z^4 + p_9yz^2 + p_{10}\left(\sqrt{y^2 + z^2}\right)^3$$

Note that this polynomial does not contain odd-order terms along the z -axis, to preserve the "left-to-right" symmetry of the kinoform. Using this polynomial, the kinoform would be optimized by adjusting the coefficients p_1 through p_{10} to minimize the ray-tracing "error," where error would be defined in terms of focus, image size, out-of-plane warp, etc. This is obviously a much smaller search space than adjusting every pixel of the kinoform. Although, a kinoform represented by such a polynomial is incapable of exhibiting all the complex optical transfer functions possible with a general kinoform, it may suffice for some applications, and serve as an excellent initialization for full pixel-by-pixel optimization of a kinoform.

5 Discussion

Visualizing scans of objects at their actual location 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* from within a object in real time.

Real Time Tomographic Reflection has received a US patent,⁹ and Real Time Tomographic *Holography* is patent pending.¹⁰

This research is supported by NSF Grant # 0308096.

References

1. State, A., Livingston, M., Garret, W., Hirota, G., Whitton, M., Pisano, E. & Fuchs, H., Technologies for Augmented Reality Systems: Realizing Ultrasound-Guided Needle Biopsies. ACM SIGGRAPH, New Orleans, LA, 439-446 (1996).
2. Rosenthal, M., State, A., Lee, J., Hirota, G., Ackerman, J., Keller, K., Pisano, E., Jiroutek, M., Muller, K. & Fuchs, H., Augmented Reality Guidance for Needle Biopsies: A Randomized, Controlled Trial in Phantoms. MICCAI 2001, Lecture Notes in Computer Science 2208, Springer-Verlag, 240-248.
3. Sauer, F., Khamene, A., Bascle, B., Schimmang, L., Wenzel, F. & Vogt, S., Augmented reality visualization of ultrasound images: System description, calibration, and features. International Symposium on Augmented Reality 2001, IEEE and ACM, New York City, 30-39.
4. Masamune, K., Fichtinger, G., Deguet, A., Matsuka, D. & Taylor, R., An Image Overlay System with Enhanced Reality for Percutaneous Therapy Performed Inside CT Scanner. MICCAI 2002, Lecture Notes in Computer Science, 2489, Springer-Verlag, 77-84.
5. Stetten, G., Chib, V. & Tamburo, R., System for Location-Merging Ultrasound Images with Human Vision. Applied Imagery Pattern Recognition (AIPR) 2000, Workshop, IEEE Computer Society, Washington, DC, 200-205.
6. Stetten, G. & Chib, V., Overlaying ultrasound images on direct vision. Journal of Ultrasound in Medicine 20, 235-240 (2001).
7. Stetten, G., Cois, W., Chang, D., Shelton, R., Tamburo, J., Castellucci, O. von Ramm, "C-mode Virtual Image Display for a Matrix Array Ultrasound Sonic Flashlight," Medical Image Computing and Computer-Assisted Intervention – MICCAI 2003. Lecture Notes in Computer Science, vol. 2879, pp. 336-343.
8. Young, M., Zone Plates and their Aberrations. Journal of the Optical Society of America 62, 972-976.
9. Stetten, G., System and Method for Location-Merging of Real-Time Tomographic Slice Images with Human Vision, U.S. Patent no. 6,599,247, July 29, 2003.
10. Stetten, G., Nowatzky, A., "Combining Tomographic Images in situ with Direct Vision Using a Holographic Optical Element," US Patent Application no. 20030199765, Continuation in Part to Patent no. 6,599,247.