High-Resolution Underwater Acoustic Imaging with Lens-Based Systems

B. Kamgar-Parsi,[1] B. Johnson,[2] D. L. Folds,[3] E. O. Belcher[4]

[1] Information Technology Division, Naval Research Laboratory, Washington, DC 20375

[2] Naval Explosive Ordnance Disposal Technology Division, Indian Head, MD 20640

[3] Ultra-Acoustics, Inc., Woodstock, GA 30188

[4] Applied Physics Laboratory, University of Washington, Seattle, WA 98105

Received 12 August 1996; revised 30 November 1996

ABSTRACT: In recent years, several sonars designed for high-resolution, short-range underwater imaging have been developed. These imaging systems use an acoustic lens to focus the incoming waves on an array of transducers. In this article we describe three prototype systems that use a line-focus or a point-focus lens and operate at a frequency of 300 kHz or 3 MHz. The line-focus lens produces two-dimensional (2D) intensity images, while the point-focus lens produces 3D intensity images. We present sample images taken from moving and stationary platforms, and discuss the techniques used for processing the acoustic backscatter data to reconstruct and visualize the scene. The images, particularly those taken with a point-focus lens, show a remarkable degree of detail. © 1997 John Wiley & Sons, Inc. Int J Imaging Syst Technol, **8,** 377–385, 1997

Key words: acoustic imaging; acoustic lens; underwater imaging

I. INTRODUCTION

In recent years, we have developed and tested several imaging sonar prototypes for high-resolution, closeup, underwater imaging. These imaging systems are designed to provide centimeter resolution from distances of a few meters for such applications as identifying objects, inspecting pipes and ship hulls for damage or barnacle buildup, etc. Although acoustic imaging has lower resolution than optical imaging, these sonars are particularly useful for imaging in turbid water, where cameras and lasers fail at short ranges. Also, these sonars are designed for use as diver-held imaging systems (or for mounting on small platforms); thus they must be compact and lightweight and consume little power. To satisfy these requirements, we have investigated the feasibility of imaging systems that use acoustic lens for beamforming. This eliminates extensive beamforming electronics used in the more conventional sonars.

The imaging systems have a focal plane, or retina, populated with a large array of transducer elements. There are two basic approaches to beamforming. The first is the phase-change-only method that assumes the signal is a sinusoidal pulse that is long enough to be present at all elements in the array simultaneously. This technique adjusts the phase of each element's signal based on an assumed direction and then adds the adjusted signals. If the acoustic wave is coming from the assumed direction, the adjusted signals will add coherently, and a large output will be present. If the signal is coming from another direction, the additions will not be coherent and the output will be small. The second is the delay-and-sum method, which makes no assumptions about the waveform's type or length. This technique simply delays and sums the outputs from each element. The computational requirements of the two approaches are similar, but the delay-and-sum approach requires additional memory for storing signal segments. The advantage of delay-and-sum approach is that it can obtain better range resolution, since the signals can be short or have a wide bandwidth not acceptable to the phase-change-only approach. The delay-and-sum method is generally preferred by those working in acoustic imaging with two-dimensional (2D) arrays, e.g., [17].

An acoustic lens is a delay-and-sum beamformer in which the delays and sums are accomplished by the refraction and concentration of acoustic rays induced by the curved surfaces of the lens. A transducer array is still required, but each array senses a preformed beam from a specified direction. If 10^4 beams are wanted in parallel, then the array would have 10^4 elements regardless of the resolution desired. Typically, images require that the echo intensity from each transducer element be mapped to the corresponding pixel. Since each transducer element senses a preformed beam, the electronics need to retain only the amplitude information before digitization. This allows, for example, for 1-cm range resolution, a sampling rate of about 75 kHz versus 40 MHz for digital beamformers. Also, while an acoustic lens uses no power for beamforming, an electronic

Correspondence to: B. Kamgar-Parsi
Contract grant sponsor: Office of Naval Research

CCC 0899–9457/97/040377-09

delay-and-sum beamformer must perform billions of operations (additions and multiplications) for a typical image. Thus, the acoustic lens eliminates much of the beamforming electronics and the associated power consumption. However, there are several challenging problems in building an acoustic lens. The focal plane must contain a large array of densely packed transducer elements. The design and fabrication of such a focal plane, while keeping side lobes down, is difficult. Moreover, the lens must be designed so as to have minimal off-axis aberrations, and be made of materials that absorb or reflect the acoustic rays as little as possible to yield a good signal-to-noise ratio. Also, the focal length must be short enough to keep the size of the imaging system small. In this article we discuss three of the lens-based imaging systems we have developed, and present sample images. This article is a summary of efforts over several years, and is therefore not comprehensive. For more details we refer to the following reports [2,3,9,16] and references therein.

The first system, described in Section II, has a liquid-filled lens with a 2D hemispherical array of 300-kHz transducers, which works in a manner similar to the eye. A single image from this sensor lacks sufficient resolution for our purposes; hence, to compensate for that we oversample the scene by mounting the sensor on a moving platform. Thus, many snapshots must be taken and combined. The second system, Section III, has a solid lens with a 2D flat array of 3-MHz transducers. (Actually, it has a single transducer, but the scanning mimics a 2D flat array.) These two systems are point-focus sensors, where transducers have conical beam patterns and the lenses are spherical. The collected data are in the form of a 3D grid of backscatter intensity values, and we discuss how they are processed and rendered to produce a 3D image of the scene. The third system, Section IV, also has a solid lens but a 1D array of 3-MHz transducer. (This also has a single transducer, but the scanning mimics a 1D array.) The transducer beam pattern is fan shaped and the lens is cylindrical. Together, they form a line-focus sensor that produces 2D intensity images or shadowgrams. Because the fabrication of 2D transducer arrays is a challenging task, we considered this sensor to investigate the tradeoff between image quality and ease of array fabrication. In Section V, we discuss our conclusions and future work.

II. POINT-FOCUS LIQUID-FILLED LENS

The first imaging system prototype designed in this work is a sensor that operates in a manner analogous to the human eye. It is a liquid-filled lens which consists of a thin, hemispherical shell and a focal plane, or retina, containing small ceramic transducers. Here, we give a brief description and refer to [2] for more details. The lens retina is populated with transducers in eight rows of 16 elements each. The transducers in each row are separated and the rows staggered to produce 1.5° resolution beams at the −3-dB points in both azimuth and elevation. The hemispherical retina and the acoustically transparent hemispherical lens window together form a cavity which is filled with a fluid. The fluid has a speed approximately 60% of the sound speed in water and refracts the sound entering the window to focus on the retina. The small number of transducers is due to the difficulty of fabricating a curved 2D array. This lens has a field of view of 48° in azimuth and 12° in elevation. Its operating frequency is 300 kHz, which corresponds to a wavelength of 0.5 cm in water. Each beam has a conical beam pattern for transmit and receive. The range bins, defined by the received time delays, are about 10 cm. The sensor maintains focus in an ambient temperature range that spans 13°C without having to change the lens fluid. Data are both recorded within the lens unit and sent topside to a PC for real-time graphical display.

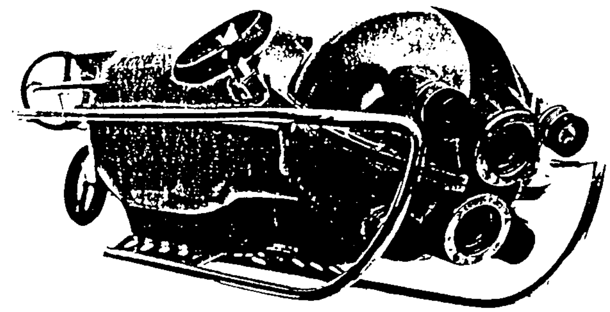

Figure 1. Picture of the 1.6-m-long ROV.

This system has a resolution of 25 cm at a range of 10 m; thus, a single scan produces a sparse sample that cannot resolve centimeter-size features. To obtain a higher-resolution image, overlapping samples are needed. Hence, the sensor is mounted on a platform and is passed slowly over the scene while it pings and acquires data. In the active mode, a row of transducers on the lens retina simultaneously emit a pulse and then receive and record the energy of the backscattered waves. Typically, one ping-cycle of the sensor is 1 s long. In a tow tank experiment, the sensor was mounted on a moving carriage and passed slowly along the tank over a 1.6-m-long remotely operated vehicle (ROV) (Fig. 1) sitting on the tow tank floor. This was repeated four times with the ROV placed at different angles (the nominal angles between the ROV axis and the sensor path were 0°, 30°, 60°, and 90°). Figure 2 shows 3D volume rendered images obtained from the four single passes. Images are rotated to correspond to the same pose as in Figure 1. Note that the closest viewing distance was 8 m. Also note that the sensor resolution at this distance is only about half the size of the ROV's head. Below, we describe how these image are generated.

Data acquired by this sensor are the acoustic backscatter energy from within volume elements, or range bins, shaped like a sliced cone. The range bins vary in size and may be irregularly positioned in space, because the sensor moves. To reconstruct the scene, we represent the space with a regular grid of small cubic volume elements, or voxels, and estimate the backscatter from each voxel. A high voxel value indicates the presence of a surface. This is done in two stages. First, since each range bin intersects several voxels, its backscatter must be distributed among them. Second, a voxel may be ensonified by several beams resulting in different estimates which must be combined to yield a single value. Algorithms for distributing returns involve modeling the beam patterns and computing range bin intersections with voxels. This is conceptually straightforward but can be computationally intensive. Algorithms for combining estimates are nontrivial because we have to deal with occlusions, partial illuminations, and differences in angle of incidence, particularly if the sensor or object is moving. See [10] for more details.

Having obtained a 3D gray-level image, we must find a way to see the object of interest through the fog of ambient noise. How best to enhance and visualize the object is also a topic of ongoing investigation. There are two main approaches to visualizing 3D scalar data: volume rendering and surface rendering [5,11]. In

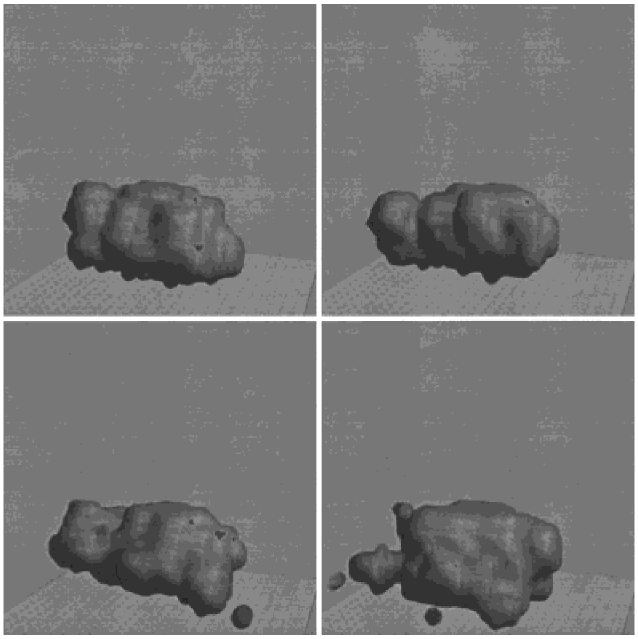

Figure 2. Volume rendered 3D images of the ROV taken in four different passes.

volume rendering, we assign to each voxel a color and a partial transparency, and then form images by blending together colored semitransparent voxels that project to the same pixel on the image plane. In surface rendering, we first apply a surface detector to the volume data, then tile the surface with polygons, and finally render the surface. The main drawback in surface rendering is that we have to classify each voxel as belonging to a surface or not. Its advantage is in rendering speed, particularly if we want to rotate the computer generated image of the scene and view it from different angles. This is because the surface voxels are determined and tiled with polygons only once. Rotating and rendering the reconstructed surfaces for subsequent views require little additional CPU time. In volume rendering, however, when the view direction is changed all the computations must be repeated anew. Nevertheless, more efficient and clever ways of volume rendering are being developed and tested [1]. Also, aliasing effects, which commonly appear in surface rendering of digitized surfaces [13], are less pronounced in volume rendering [1].

In our experience, when data are of high resolution and relatively clean, surfaces can be reliably detected and reconstructed, and surface rendering produces satisfactory 3D images. However, in cases where the resolution is not sufficiently high and the data are further blurred by sensor motion, volume rendering appears to produce better images. This is mainly because surface detection becomes less reliable. Images in Figure 2 were obtained by volume rendering. Before the scene is actually rendered, it is thresholded and filtered so as to remove isolated voxels cluttering the view [10].

Scene resolution may be further enhanced by combining several different data sets of the same scene. However, the data sets must first be registered; otherwise, the scene will become blurred and distorted. We have developed algorithms based on 3D moment matching that perform scene registration with enough accuracy to significantly enhance the resolution [10]. Figure 3 is the result of registering and combining the data taken in the four passes shown in Figure 2. Note that such details as the valve on the ROV side and the strobe lights and cameras on the head begin to emerge as bumps in the right positions.

In these experiments, we have seen that it is possible to obtain higher-resolution images (higher than the sensor resolution) by moving the sensor to collect overlapping samples and compounding them. We have also seen that data from multiple passes can be registered accurately and combined to further improve the scene resolution. However, data acquisition and processing in this manner is time-consuming. In addition, scene compounding, if it is to be done reliably, requires fairly accurate sensor position information—a luxury that may not exist in most situations. An acoustic snapshot taken by a stationary sensor with higher resolution would be more practical.

III. POINT-FOCUS SOLID LENS

To obtain images higher in resolution than those obtained by the liquid-filled lens, we have to use a focal plane more densely packed with higher-frequency transducers. In this design, we consider a 2D array of 3-MHz transducers in a flat focal plane (easier to fabricate than a curved retina) and a solid lens. Here, we give a brief description of the imaging system and refer to [8] and [16] for more details. In this imaging sonar, beamforming is done with a biconcave solid lens made of crystal polystyrene. This lens is lighter than the liquid-filled lens discussed in the previous section, and maintains focus over a wider temperature range. At the on-axis focal point behind the lens, there is a single 3-MHz transducer element 3 mm in diameter. Together, the element and the lens form a conical beam with 0.15° resolution between the −3-dB points at the target distance. The wavelength is about 0.05 cm in water; thus, for imaging centimeter-size features, one can use ray optics for lens design. To obtain an image, the system

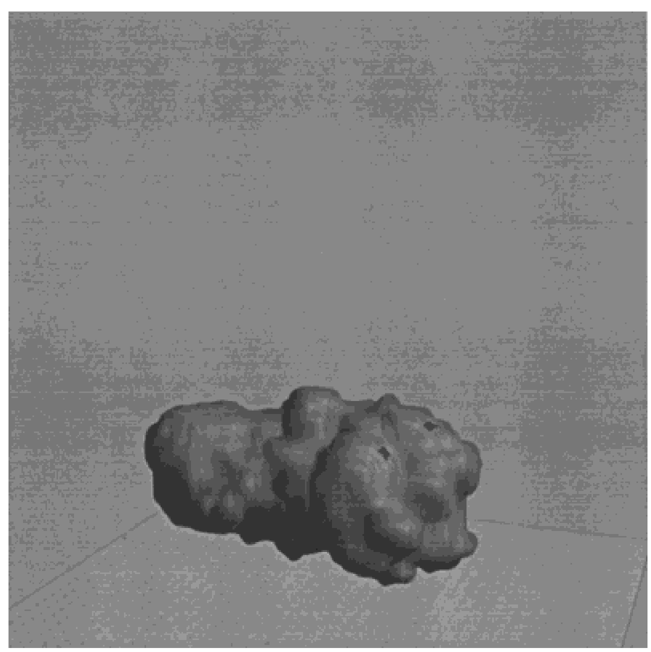

Figure 3. A 3D image of the ROV after registering and combining the four passes shown in Figure 2.

mechanically scans and interrogates the object at regular intervals (typically 0.5 cm) from left to right and from top to bottom. This scanning strategy requires a few hours to produce an image, but it mimics an ideal lens with a large flat focal plane populated with some 100×100 transducers, obviously with no aberration and no cross-talk among the elements. To speed up image acquisition, mechanical scanning must be replaced by a focal plane populated with many transducers. However, for the single element lens used here, this will result in significant off-axis image distortions. Techniques for reducing aberrations through compound lens design, together with innovative scanning strategies, are currently under development.

The cross-range resolution of this imaging system is about 0.5 cm at a distance of 2 m. The down-range resolution is determined by the sampling rate of the echoes, typically 75 or 150 kHz, which correspond to about 1 or 0.5 cm, respectively. Thus, an image taken with a stationary system from a distance of a few meters should be capable of resolving features smaller than 1 cm. The backscatter data from this imaging system are in the form of a 3D (cross-range, down-range) grid of values. The data are typically of high quality with fairly low noise. To reconstruct and visualize the scene, we perform the following operations. First, we transform the data into a 3D rectilinear grid of cubic voxels using linear interpolation (the grid is typically 100^3 large). Then we apply a surface detection algorithm to label voxels as surface or not. Then we construct a surface by tiling it with small polygons, and finally render the surface to obtain a 3D view of the object. There are a number of approaches to surface detection. These include traditional gradient-based techniques, e.g., [12], or more novel physics-based techniques based on surface particles that diffuse through the image field and/or interact with each other, e.g., [18]. The latter are relatively slow and are typically used for interactive surface detection. In this work, we used gradient-based surface detection algorithms. (Taking the first maximum return along each beam leads to similar results.) For rendering the polygon-tiled surfaces, we used Phong shading, which gives the objects a realistic appearance [11].

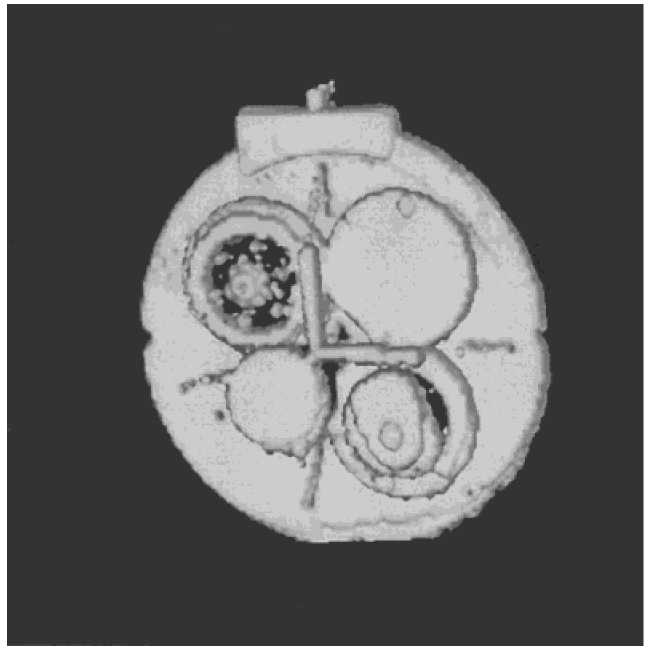

Figure 5. A 3D surface rendering of the plate, from point-focus lens data, taken as the plate faces the lens.

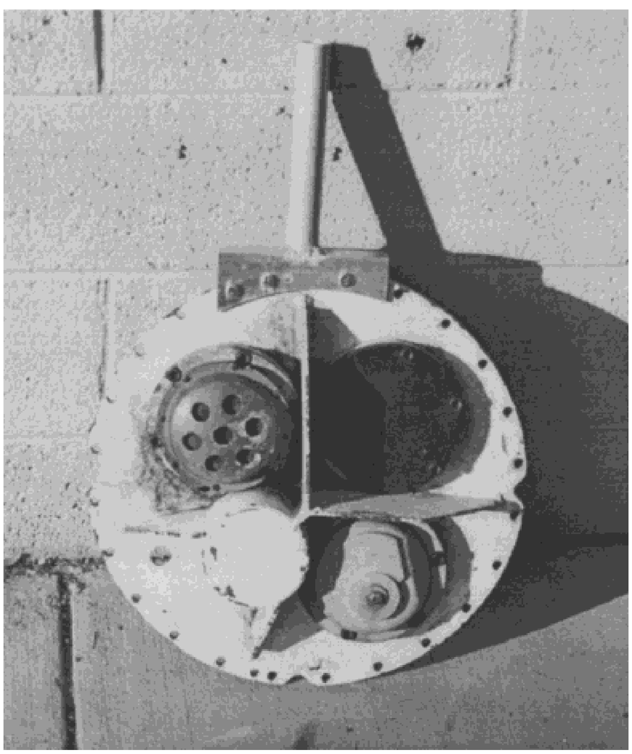

Figure 4. Photograph of the plate with four circular substructures.

We present a sample of images taken with a stationary point-focus lens in a test tank. The scenes include objects either suspended in water or resting on a sandy bottom. Figure 4 is a photograph of one of the objects used in the experiments. It is a circular plate (45 cm in diameter) with four circular substructures (ranging between 10 and 18 cm in diameter) as well as four ribs (0.6 cm wide by 8.5 cm high) perpendicular to the plate. Figure 5 is the surface-rendered acoustic image of the plate. The plate is suspended in water while directly facing the lens. Note the remarkable degree of detail that can be observed in this image: the small hole in the bottom-left quadrant, the rivets, and even a hint of the three bolt heads on the handle holding the plate.

Figure 6 is also a surface-rendered image of the plate suspended in water, while facing the lens at a 30° angle. Because the plate makes an angle with the lens axis, the angle of incidence is not favorable and the returns are weaker. This results in a less accurate surface detection, and hence a noisy appearance in which some of the details become difficult to discern. This noisy appearance can be improved by applying a set of filters to the 3D data before surface rendering (Fig. 7), where certain details such as rivets and bolt heads can now be discerned. The filtering consists of a sequence of filters: a low-pass filter, a high-frequency emphasis (HFE) filter, and a median filter. The low-pass filter removes some of the noise; the HFE filter enhances small protruding features while filling small holes; and the median filter cleans up certain artifacts produced by HFE filter, such as transforming a

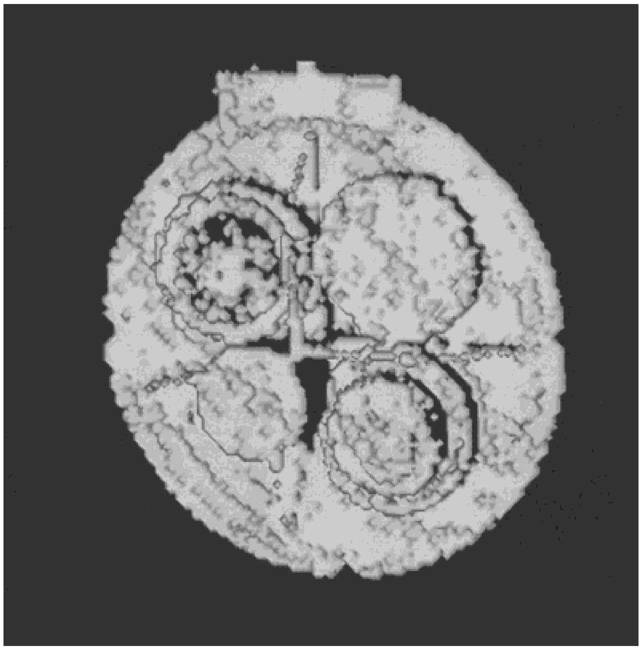

Figure 6. A 3D surface rendering of the plate, from point-focus lens data, taken as the plate is tilted with respect to the lens.

Figure 8. Photograph of the cylinder with a recessed well.

single surface to a doublet [9]. Note that small-scale features appear fatter than they really are, e.g., the ribs.

Figure 8 is a photograph of another object used in the experiments. It is a section of a cylinder (47 cm in diameter) with a recessed well (18 cm in diameter and 10 cm deep) that has eight hexagonal nuts (1.9 cm in diameter) and bolts. Figure 9 is the surface-rendered acoustic image of this object while suspended in water. The object was in a pose where the well bottom directly faced the lens. Note that not only all the bolt/nut combinations clearly stand out, but that even certain bolts and nuts can be distinguished from one another. The noisy appearance of the cylindrical part is because of weak returns due to unfavorable incidence angles, as well as aliasing due to scene quantization.

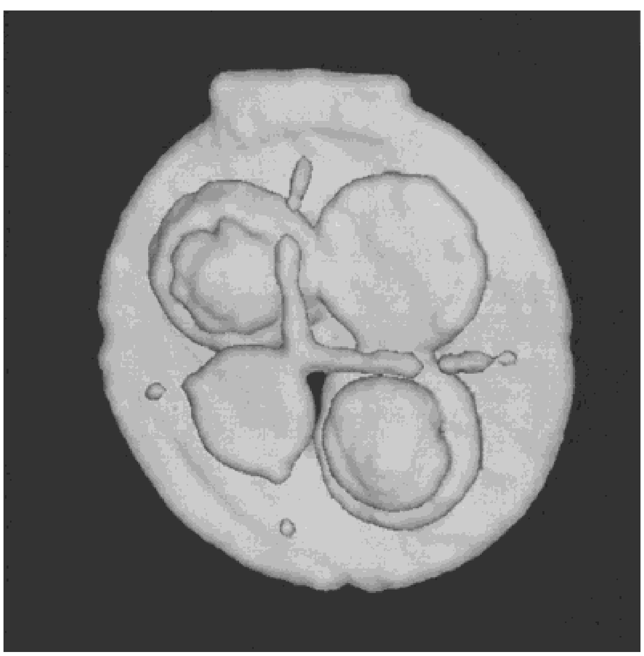

Figure 7. Results of filtering the reconstructed surface shown in Figure 6.

Figure 9, as well as others above, were obtained in a monostatic imaging mode, i.e., each transducer is used to transmit and receive, which requires the transducers to be relatively large to put out enough power. If the image plane transducers are used only to receive, then they can be made smaller, which makes the fabrication of the 2D transducer array less difficult. Figure 10 is a sample image that shows the quality of the bistatic (receive only) imaging mode. The original data are more noisy than its monostatic counterpart (Fig. 9) and is filtered prior to rendering. This image also shows the hollow metallic bar used to suspend the cylinder in the water tank. From our experiments, it appears that monostatic imaging mode provides higher-quality images.

In other experiments, we placed objects on a sandy bottom in a water tank. Figure 11 is the photograph of one of the objects used in these experiments. It is a 14-in-long cylinder with a diameter of 9 in, with one end open and the other end closed. The cylinder has two sets of six hexagonal nuts (diameter of $\frac{1}{2}$ in and height of $\frac{3}{8}$ in) and bolts, and a 2-in-tall fin. Figure 12 shows different visualizations of the acoustic backscatter data from the cylinder. The lens is 30 in above the floor and 60 in from the objects, and is looking downward at a 20° angle. The top panel of Figure 12 is a traditional 2D range image: bright

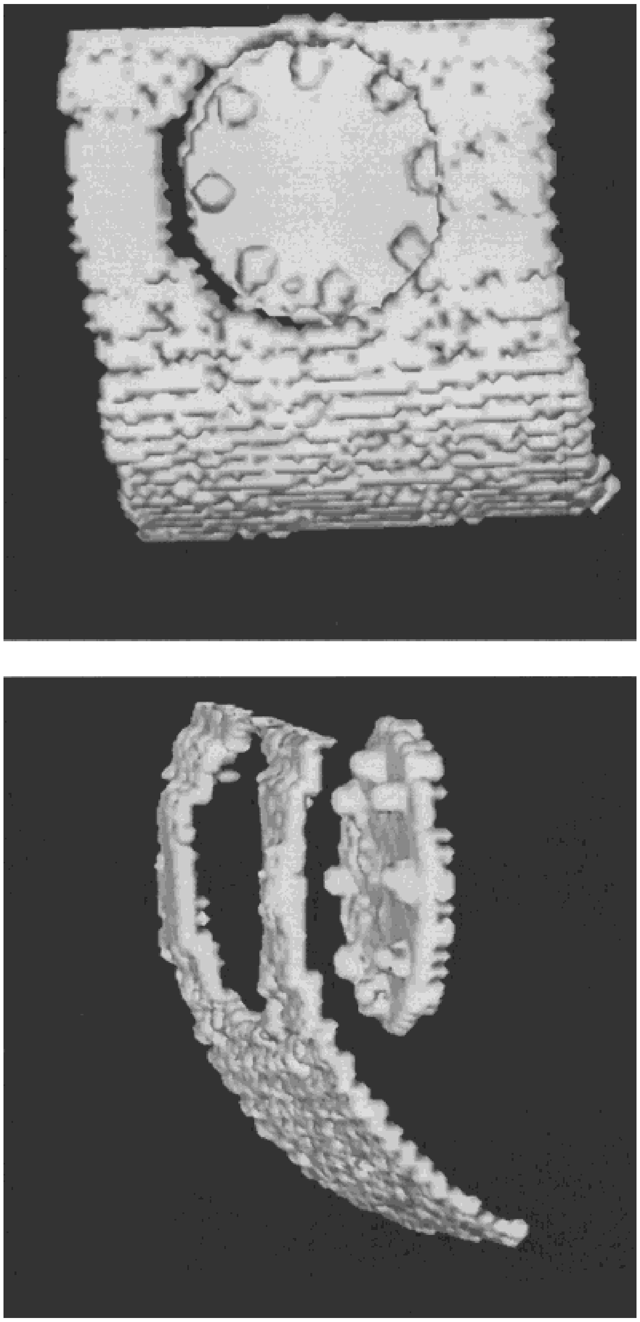

Figure 9. Two views of the same 3D acoustic image of the cylinder in Figure 8, taken with the point-focus lens as the recessed well was facing the lens. The object is suspended in the water tank.

indicates near, while dark indicates far. The range image is obtained by taking the first strong return (first maximum return above a threshold) along each beam. The body of the cylinder as well as the fin and the nuts on top can be seen, while the nuts on the side are not visible. The top portion of the cylinder is also not visible because the angle of incidence is nearly tangential. The middle panel is an acoustic photograph, where the backscatter intensities along each beam are integrated. In this image, the nuts are visible because they form concave corners which are strong

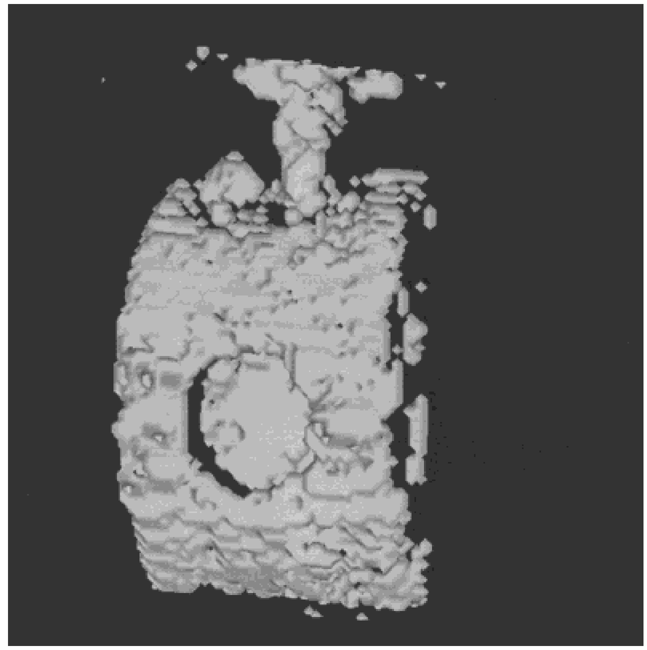

Figure 10. A 3D surface rendered image of the cylinder in Figure 8, taken with the point-focus lens (used in bistatic mode) with the recessed well facing the lens. The object is suspended in the water tank.

reflectors. Also note the image of the bottom in the cylinder and vice versa. The bottom panel of Figure 12 is a surface rendered 3D image of the scene filtered in a manner similar to Figure 7. Note that the nuts on the side of the cylinder are now clearly visible, albeit larger than they really are.

In the case of an object on the bottom, the surface reconstruction for 3D rendering is complicated by interreflections between the object and the bottom. This object/bottom mirror effect produces phantom surfaces that appear behind the true surfaces. The strength of these interreflections is typically comparable to direct

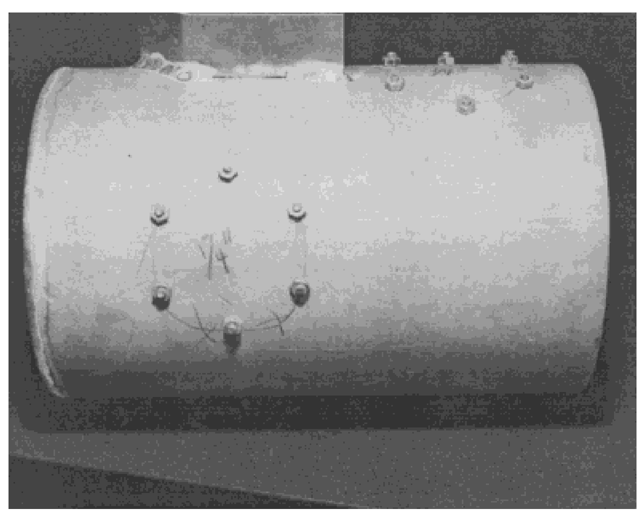

Figure 11. Photograph of the foot-long cylinder with two hexagonal nut patterns and a fin.

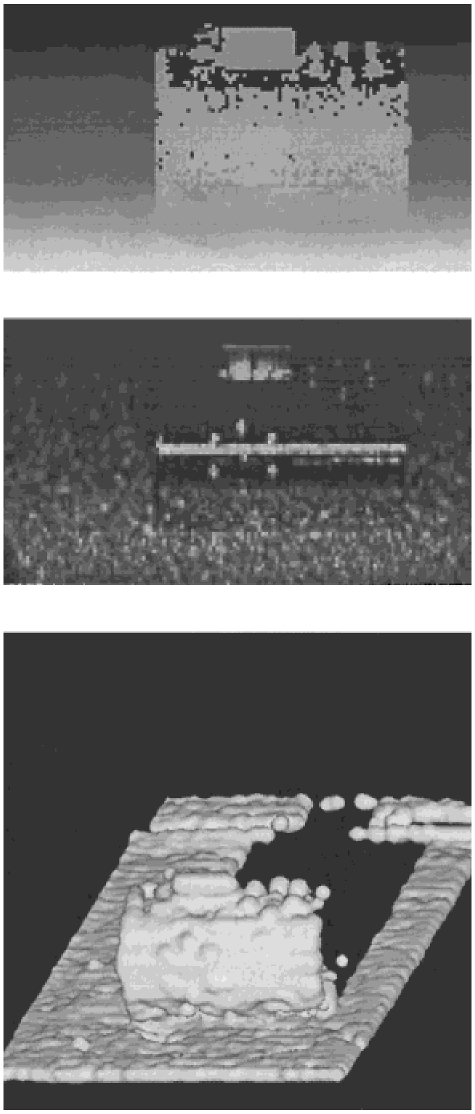

Figure 12. Acoustic images of the foot-long cylinder on a sandy bottom taken by the point-focus lens: range image (top); intensity image (middle); a surface rendering of range data filtered so as to enhance smaller features (bottom).

reflections from surface patches facing the lens, and is often stronger than reflections from surface patches at an angle with the lens axis. Therefore, a surface detection algorithm could detect these phantom surfaces while missing some of the true surfaces. We have found that taking the first maximum return (above a threshold) along each beam as the surface location and discarding the remainder of the beam can avoid this problem to some extent. However, it may happen that the return from a true surface patch with an unfavorable pose is not strong enough to be detected; hence, the range image will be a mixture of (mostly) true and (some) phantom surfaces, and thus appears noisy. The range images shown here were obtained in this manner.

Experiments were also conducted in a test tank, where objects with small bolt patterns and ribs, e.g., Figure 4, were imaged from a distance of about 2 m under various levels of turbidity. The results show that the images are not degraded by the mud and silt commonly found in shallow water. Some degradation occurs at high levels of suspended sand [8].

IV. LINE-FOCUS SOLID LENS

Fabrication of 2D transducer arrays, the associated electronics, and the interconnection between them is a challenge. Therefore, we also considered an imaging system consisting of a 1D array with a line-focus lens to investigate the tradeoff between image quality and fabrication ease. In this section, we give a brief description of the imaging system (see [3] for more details) and present sample images.

In this imaging configuration, beamforming is accomplished with a solid cylindrical lens made of crystal polystyrene. A curved transducer element 2 mm tall by 6 cm long is placed at the on-axis focal point behind the lens and is resonant at 3 MHz. The lens has an anechoic rubber coating with a 0.57-cm slit to block azimuthal side lobes. Together, the element and lens form a fan beam 0.17° in elevation and 4° in azimuth at the −3-dB points at the target distance. To obtain an image, the system mechanically scans and interrogates the object at 0.35-cm intervals top to bottom and 7.8-cm intervals left to right. (Scanning at 7.8-cm intervals was done because the 4° beam was not wide enough to ensonify objects of interest in their entirety.) This scanning process typically takes 15 min but simulates a lens with good off-axis performance coupled with a 150 × 7 element array with no interelement cross-talk. This system, in effect, interrogates a line (a thin stripe) on the object plane and records the backscatter intensities. The cross-range (elevation) resolution is a function of beamwidth, while the down-range (azimuth, time delay) resolution is a function of sampling rate of backscatters in time.

The line-focus imaging system mimics a vertical array of closely packed transducers with a beam that is narrow in elevation and wide in azimuth. Displaying the data taken with this system is straightforward. The fan-beam data are the backscatter energy as a function of azimuth and elevation, which can be naturally represented as a 2D gray-scale image. To do so, we form a regular rectilinear grid (or a polar grid, depending on the scanning strategy used), appropriately scaled so that down-range and cross-range have identical dimensions. The acoustic backscatter energies are digitized to the usual 8-bit gray-level range, i.e., 256 discrete values. For the data sets in which the cross-range and down-range are sampled at different resolutions, the grid size is selected to be the smaller sampling interval. We use a bilinear interpolation to fill in the grid sites from neighboring points.

Figure 13 shows the image of the plate (Fig. 4) suspended in a water tank, at three different angles of incidence. The angle between the ensonification direction and the normal to the plate is, respectively, from top to bottom, 30°, 45°, and 60°. In these images, the vertical axis is cross-range and the horizontal axis is down-range. The images are taken from a distance of about 1.2 m. These images show the general shape of the object as well as a significant amount of detail. Note, for example, the bolt heads used to attach the plate to the bar, and the rivets in the upper right quadrant of the plate. Also note that in the 30° image the details of the structure in the lower right quadrant are visible, while in the 60° image the hemispherical feature in the upper left quadrant, including the holes, can be discerned. Those patches with surface normals parallel to the illumination direction produce the strongest returns and appear brightest in the image. These include corner reflectors, bolt heads, rivets, cylindrical pipes, etc. The vertical bright line present in the

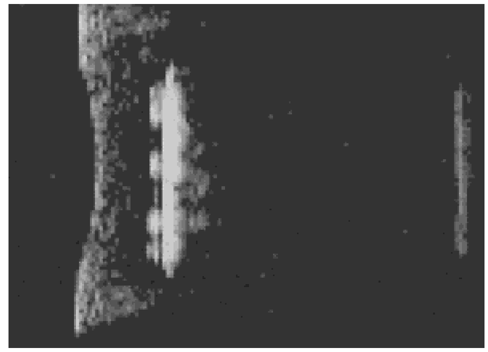

Figure 14. Fan-beam image of the cylinder with the recessed well facing the lens. The object is suspended in the water tank.

bottom of all the three images is due to strong returns from the concave wedge formed by the rib and the plate. Also note the reverberations of the hollow bar attached to the top of the plate, which appear as a series of high-intensity lines parallel to the tube and spaced at regular down-range intervals.

The low-intensity contrast in these images is enhanced by a nonlinear scaling of the average intensity to the middle of the gray-level range, i.e., to 127. A logarithmic scaling of the intensities produces a similar effect. These images can be further enhanced through a number of standard 2D image-processing techniques (e.g., [14,15]). The clarity of these images is clearly not as good as those obtained by the point-focus lens, partly because the returns are averaged over a wider angle. Furthermore, interpretation of these images is made more complicated by the appearance of phantom surfaces caused by interreflections among object features that produce concave surfaces. Interreflections have a longer time delay than direct returns, and thus appear at longer ranges blurring features that are farther from the lens.

Figure 14 shows the image of the cylinder (Fig. 8) with the recessed well facing the lens, while suspended in a water tank. The bottom of the recessed well and the bolt heads as well as the mouth of the well are clearly visible.

We also did experiments by placing objects on a sandy bottom. In these experiments, the images taken with the line-focus system show only that something is there, but it is difficult to identify the object. This is because the returns from the sandy bottom are comparable to those from the object; and in particular, strong interreflections between the bottom and various object features produce phantom features that mix with true features, making the details obscure and the object unrecognizable. For suspended objects, however, this imaging system provides sufficient detail for certain tasks. For example, a four-element line-focus system has been developed and is currently being evaluated for ship hull inspection, for eventual integration in an automated system for barnacle cleaning.

V. CONCLUDING REMARKS

Lens-based systems with 3-MHz transducers provide images that have sufficient contrast and resolution to identify centimeter-size features. The line-focus lens provides 2D images with adequate

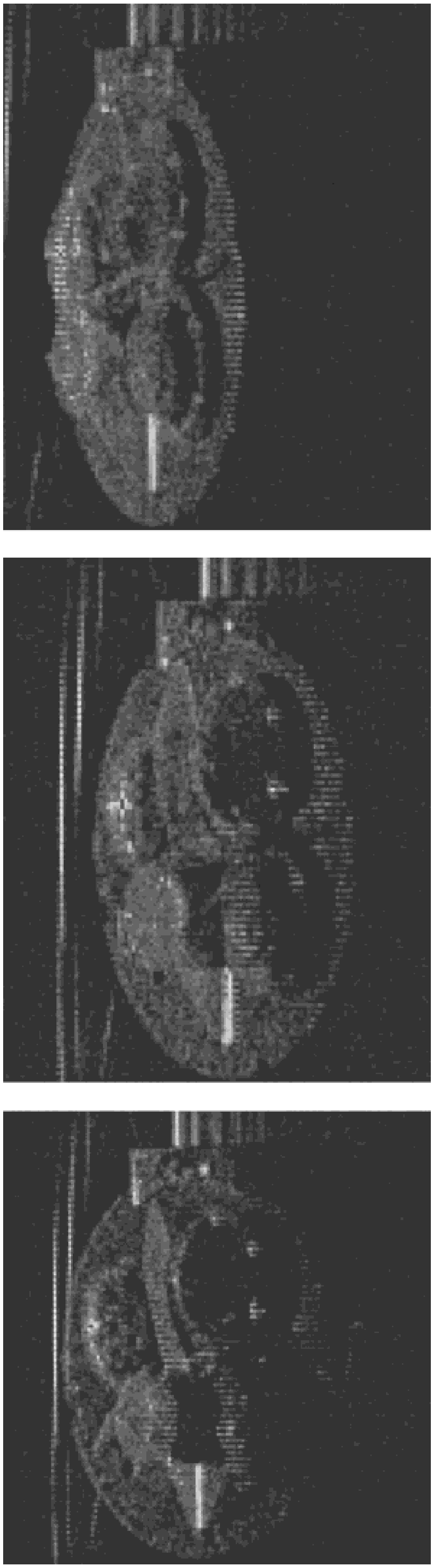

Figure 13. Fan-beam images of the plate at three different angles of incidence: 30°, 45°, and 60°. The object is suspended in the water tank.

contrast for objects suspended in water—that is, when interreflections between objects, or object and bottom, do not significantly blur images. Indeed, a line-focus system has been developed for ship hull inspection and is currently being tested. The point-focus lens typically provides high-quality 3D data which can yield vivid, detailed 3D images of suspended objects as well as those on the bottom. Therefore, it appears that an imaging system with a point-focus lens will be most useful under general field conditions.

The 3-MHz point-focus solid lens imaging system discussed here has a single ceramic transducer at the focal point of the lens. The images are acquired by mechanically scanning the scene, which mimics a 100×100 transducer array. Thus, the resulting images have no acoustical or electrical cross-talk noise among the transducers, and no distortion due to off-axis lens aberrations. They represent the limit of what can be achieved with such an imaging system. Image acquisition, however, is far from real time. To speed up image acquisition, complementary approaches are currently under investigation. One approach is to populate the focal plane with many transducers (in conjunction with compound lens design to prevent off-axis aberrations), and the other is to use innovative scanning strategies. For example, a 96-line-focus beam system is already developed and being tested to investigate cross-talk effects on images. This, together with Mills cross projection, approaches a point-focus result with a technology slightly more complex than line-focus technology [3].

One of our goals is to develop a compact 2D array-based underwater lens imager, or an acoustic video converter (AVC), that would be capable of taking several images per second. To this end, we are attempting to develop 2D array manufacturing, signal processing, and interface electronics technologies. Several approaches are under investigation drawing on technologies from the medical ultrasound, infrared, and international defense communities. For example, a particular design for the AVC, which is expected to take images at a rate of up to 10 frames/s, has a focal plane densely packed with an array of PZT 3-MHz transducers [3]. To cope with the massive amount of data collected by such a system, the associated electronics is designed to detect and transmit only the location and amplitude of the first strong return (first maximum above a threshold) along each beam. This reduces the data by two orders of magnitude, which makes processing and displaying images manageable in real time. In our work, we have seen that taking the first maximum along each beam produces images similar to those obtained by extracting surfaces from the full 3D sampled data. There may even be an advantage to taking the first maximum along each beam, since interreflections between, say, object and bottom, which lead to the appearance of phantom surfaces, are eliminated.

Depending on the AVC design, the image data will be range only, intensity only, or range and intensity. The range image will typically show the large-scale structure of the object, while the intensity image will show strong reflectors even though they may be small, such as corner reflectors (Fig. 12). In either case, displaying the image is straightforward. The electronics of the AVC that can output range/intensity images is more complex. The question of how best to make use of these two complementary images is open. We are conducting experiments to determine this. The possibilities include displaying both images side by side, or overlaying the two images (e.g., a 2D image shaded by range and colored by intensity), or texture mapping the intensity image over a 3D scene reconstructed from the range data. There are many techniques in the literature that may be used for reconstructing surfaces from range image data, e.g., [4,6]. These techniques work well for low-noise images. In noisy cases, they typically result in smooth surfaces at the cost of losing small features. To suppress range image noise while preserving small-scale features, we are developing image-processing algorithms that exploit the intensity information through shape-from-shading techniques [7].

REFERENCES

1. R. Avila, T. He, L. Hong, A. Kaufman, H. Pfister, C. Silva, L. Sobierajski, and S. Wang, ''VolVis: A diversified volume visualization system,'' in *Proceedings of the IEEE Visualization '94 Conference,* Washington, DC, October 1994. R. D. Bergeron and A. E. Kaufman, Ed., IEEE CS Press, Los Alamitos, CA, 1994, pp. 31–38.
2. E. O. Belcher, D. Steiger, and L. J. Rosenblum, ''A forward looking active acoustic lens,'' Tech. Rep. APL-UW TR 9113, Applied Physics Laboratory, University of Washington, Seattle, WA, 1991.
3. E. O. Belcher, D. L. Folds, B. Kamgar-Parsi, and B. Johnson, ''High resolution acoustic imaging for diver-held applications: FY90-FY95 overview report,'' APL-UW TR 9601, Applied Physics Laboratory, University of Washington, Seattle, WA, 1996.
4. R. M. Bolle and B. C. Vemuri. ''On three-dimensional surface reconstruction methods,'' *IEEE Trans. Pattern Anal. Machine Intell.* **13,** 1–13 (1991).
5. H. Fuchs, M. Levoy, and S. M. Pizer. ''Interactive visualization of 3D medical data,'' *Computer* **August,** 46–51 (1989).
6. H. Hoppe, T. DeRose, T. Duchamp, M. Halstead, H. Jin, J. McDonald, J. Schweitzer, and W. Stuetzle, ''Piecewise smooth surface reconstruction,'' in *Proceedings of the SIGGRAPH '94 Conference,* Orlando, FL, July 1994, Addison-Wesley, Reading, MA, 1994, pp. 295–302.
7. B. K. P. Horn, *Robot Vision* (MIT Press, Cambridge, MA), 1986.
8. B. Johnson, D. Scroggins, D. Folds, B. Kamgar-Parsi, and E. Belcher. ''3-D acoustic imaging with a thin lens,'' in *Proceedings of the IEEE Oceans '93 Conference,* Victoria, BC, October 1993, Vol. 3, pp. 444–449.
9. B. Kamgar-Parsi, ''Underwater imaging with acoustic lens: image processing and visualization,'' Tech. Rept. NRL/FR/5580-96-9826, Naval Research Laboratory, Washington, DC, August, 1996.
10. B. Kamgar-Parsi, L. J. Rosenblum, and E. O. Belcher. ''Underwater imaging with a moving acoustic lens,'' *IEEE Trans. Image Process.* (in press).
11. A. Kaufman, *Volume Visualization* (IEEE CS Press, Los Alamitos, CA), 1991.
12. D. Lee. ''Coping with discontinuities in computer vision: their detection, classification, and measurement,'' *IEEE Trans. Pattern Anal. Machine Intell.* **12,** 321–344 (1990).
13. S. R. Marschner and R. J. Lobb, ''An evaluation of reconstruction filters for volume rendering,'' in *Proceedings of the IEEE Visualization Conference,* Washington, DC, October 1994. R. D. Bergeron and A. E. Kaufman, Ed., IEEE CS Press, Los Alamitos, CA, 1994, pp. 100–107.
14. A. Rosenfeld and A. C. Kak, *Digital Picture Processing* (Academic Press, New York), 1982.
15. J. C. Russ, *The Image Processing Handbook,* 2nd ed. (CRC Press, Boca Raton, FL), 1994.
16. D. M. Scroggins and B. Johnson, ''Test report for high resolution acoustic imagery experiments,'' Technical Rep. TR-50-94-1, Naval Explosive Ordnance Disposal Technology Division, Indian Head, MD, 1993.
17. S. W. Smith, G. E. Trahey, and O. T. van Ramm. ''Two-dimensional arrays for medical ultrasound,'' *Ultrasound Imag.* **14,** 213–233 (1992).
18. R. Szeliski, D. Tonnesen, and D. Terzopoulos, ''Modeling surfaces of arbitrary topology with dynamic particles,'' in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,* New York, NY, June 1993, IEEE CS Press, Los Alamitos, CA, 1993, pp. 82–87.