AD NUMBER

ADB275123

NEW LIMITATION CHANGE

TO

Approved for public release, distribution unlimited

FROM

Distribution authorized to U.S. Gov't. agencies only; Proprietary Info.; Jan 2002. Other requests shall be referred to U.S. Army Medical Research and Materiel Command, 504 Scott St., Ft. Detrick, MD 21702-5012.

AUTHORITY

USAMRMC ltr, dtd 28 July 2003

AD_____

Award Number: DAMD17-99-1-9034

TITLE: Ultrasound Imaging Initiative

PRINCIPAL INVESTIGATOR: Raj Shekhar, Ph.D.

CONTRACTING ORGANIZATION: The Cleveland Clinic Foundation
Cleveland, Ohio 44195

REPORT DATE: January 2002

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

NOTICE

USING GOVERNMENT DRAWINGS, SPECIFICATIONS, OR OTHER DATA INCLUDED IN THIS DOCUMENT FOR ANY PURPOSE OTHER THAN GOVERNMENT PROCUREMENT DOES NOT IN ANY WAY OBLIGATE THE U.S. GOVERNMENT. THE FACT THAT THE GOVERNMENT FORMULATED OR SUPPLIED THE DRAWINGS, SPECIFICATIONS, OR OTHER DATA DOES NOT LICENSE THE HOLDER OR ANY OTHER PERSON OR CORPORATION; OR CONVEY ANY RIGHTS OR PERMISSION TO MANUFACTURE, USE, OR SELL ANY PATENTED INVENTION THAT MAY RELATE TO THEM.

LIMITED RIGHTS LEGEND

Award Number: DAMD17-99-1-9034
Organization: The Cleveland Clinic Foundation

Those portions of the technical data contained in this report marked as limited rights data shall not, without the written permission of the above contractor, be (a) released or disclosed outside the government, (b) used by the Government for manufacture or, in the case of computer software documentation, for preparing the same or similar computer software, or (c) used by a party other than the Government, except that the Government may release or disclose technical data to persons outside the Government, or permit the use of technical data by such persons, if (i) such release, disclosure, or use is necessary for emergency repair or overhaul or (ii) is a release or disclosure of technical data (other than detailed manufacturing or process data) to, or use of such data by, a foreign government that is in the interest of the Government and is required for evaluational or informational purposes, provided in either case that such release, disclosure or use is made subject to a prohibition that the person to whom the data is released or disclosed may not further use, release or disclose such data, and the contractor or subcontractor or subcontractor asserting the restriction is notified of such release, disclosure or use. This legend, together with the indications of the portions of this data which are subject to such limitations, shall be included on any reproduction hereof which includes any part of the portions subject to such limitations.

THIS TECHNICAL REPORT HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION.

_____ _____

_____ _____

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	January 2002	Annual (21 Dec 00 - 20 Dec 01)

4. TITLE AND SUBTITLE
Ultrasound Imaging Initiative

5. FUNDING NUMBERS
DAMD17-99-1-9034

6. AUTHOR(S)
Raj Shekhar, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
The Cleveland Clinic Foundation
Cleveland, Ohio 44195

E-mail: shekhar@bme.ri.ccf.org

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
Report contains color

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Distribution authorized to U.S. Government agencies only (proprietary information, Jan 02). Other requests for this document shall be referred to U.S. Army Medical Research and Materiel Command, 504 Scott Street, Fort Detrick, Maryland 21702-5012.

12b. DISTRIBUTION CODE

13. ABSTRACT ***(Maximum 200 Words)***

This objective of this project is to build a real-time 3D ultrasound imaging system for combat casualty care. The high frame rate necessary for real-time 3D imaging is obtained using a synthetic aperture beamforming technique. The technique uses a fraction of the transmit pulses required by a conventional imaging system and permits very rapid image acquisition with no degradation of image quality. The hardware and software development for our beamformer, that uses a network of high-speed digital signal processors for generating data at the desired speed, has been completed. The fabrication of a new multi-layer transducer array has been completed as well. This array has been incorporated into a compact probe head designed to rock the array for real-time 3D imaging. The assembly of the final system is currently under way. Besides the scanner development, an extensive set of software tools have been developed to provide easy and accurate analysis of the resulting 3D data. These tools allow real-time viewing of arbitrary two-dimensional planes through the data set and quantitative assessment and display of the 3D anatomy.

14. SUBJECT TERMS
Ultrasound, 3-D Imaging, Image Processing, Synthetic Aperture

15. NUMBER OF PAGES
96

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

TABLE OF CONTENTS

I. FRONT COVER .....

II. SF-298 2

III. TABLE OF CONTENTS 3

IV. INTRODUCTION 4

V. BODY 5

VI. KEY RESEARCH ACCOMPLISHMENT 6

VII. REPORTABLE OUTCOMES 6

VIII. CONCLUSIONS 8

IX. APPENDICES 9

INTRODUCTION

Providing accurate medical imaging at an aid station or remote field hospital is difficult. Portable ultrasound instrumentation can be designed for these applications, but the expertise required for an accurate diagnosis can only be obtained through years of training. Much of the information in an ultrasound examination is obtained by exploiting the real-time nature of the imaging modality. A successful diagnosis relies on the skill of the diagnostician to transform mentally dynamic two-dimensional (2D) images into the complex three-dimensional (3D) anatomy. Locating anatomical landmarks and moving the scan plane throughout the volume of interest are both critical components of this process. Without extensive training, it would be difficult for a medical corpsman to perform this procedure.

Researchers have started exploring tele-medicine in combination with 3D ultrasound data acquisition as a solution to this problem. This combination could potentially transfer the skill required to scan the patient and make the diagnosis from a medical corpsman to an imaging expert. Unfortunately, acquiring a 3D ultrasound data set is difficult. Modern 3D ultrasound systems are essentially conventional scanners modified to collect a series of 2D images. The images are later 'stacked' to represent the 3D anatomy. Although modern scanners are designed to collect 2D images in real-time (20 2D images/s), 3D image acquisition is slow. Slow image acquisition introduces the problem of how to align adjacent 2D images collected at different times. The patient and imaging probe can be immobilized to reduce movement of the anatomy between adjacent images. Cardiac and respiratory gating can also be applied. Even in a carefully controlled clinical setting, the resulting 3D data set is often badly distorted. If the patient, the anatomy, or the transducer moves during 3D image acquisition, the data must be discarded. Providing the care needed to obtain 'good' 3D data in the clinic is troublesome, on a battlefield it would be very difficult.

We are building an ultrasound imaging system that would avoid these problems. The two key components of our system are the following: 1) a high speed scanner that can collect 3D data 40 to 80 times faster than current 3D imaging approaches, and 2) a set of software tools for rapid image manipulation and analysis at a remote site. Real-time 3D image acquisition eliminates the need for patient immobilization, and cardiac and respiratory gating. A medical corpsman would simply place a small probe on the patient and position the sample volume by viewing a real-time two-dimensional image of the anatomy. Once the probe is correctly placed, a three-dimensional data set would be recorded in real-time (0.05 s for each 3D data set). Following data acquisition, the images would be transmitted to a central hospital for post-processing and analysis. An expert clinician could 're-scan' the patient by looking at multiple 2D planes through the data sets, or examine a computer reconstruction of the three-dimensional anatomy. The 3D data set could also be analyzed quantitatively to calculate dynamic changes in the anatomy. The entire imaging procedure, including subject preparation, would take only a few minutes.

BODY

A description of progress in each of the areas outlined in the statement of work is given below.

Phase III: Month 24 – 36

Scanner/Array Development:

- System Integration – Assemble and debug prototype scanner.

IN PROGRESS – The beamformer, data acquisition system, high voltage electronics, scan converter, transducer array, array module, and motion system have been assembled and tested individually. A test of the system from the level of the data acquisition system to image display has been performed successfully. However, the sensitive pre-amplifiers and protection circuits that form the interface from the array to the data acquisition system are still underdevelopment.

- Evaluation Experiment – Evaluate the static and dynamic performance of the scanner using test objects and tissue equivalent phantoms.

NOT COMPLETED – Work on the evaluation experiments will begin once the entire system has been assembled and debugged.

- Report on scanner performance.

NOT COMPLETED – A report on the scanner performance will be prepared following the evaluation experiments.

User Interface & Image Processing Software:

- Apply software to phantom, animal and clinical data.

PARTIALLY COMPLETED – We have applied our software to phantom and clinical data. Clinical cardiac data has been used for real-time scan conversion, interactive 3D visualization (see Appendix 4) and 3D image registration (see Appendix 5). Phantom, animal and clinical data are also being used for 3D image segmentation, currently in the refinement stage. The software will be applied to real-time 3D images produced by our scanner as soon as the system integration is completed.

- Refine clinical toolboxes and post-processing algorithms.

PARTIALLY COMPLETED – The breast imaging toolbox has been completed and refined. The automated diagnostic tool developed allows differentiating between cysts and non-cysts. The technique is described in detail in a manuscript submitted for publication (see Appendix 6).

Visualization, registration and segmentation make up the cardiac toolbox. 3D Visualization and registration algorithms for the cardiac toolbox have been refined and integrated into the main

application. These algorithms have been published (Appendices 4 and 5). The work on 3D image segmentation is ongoing.

The real-time scan conversion program has been completed, refined and tested with the 3D beamformer and a simulated probe. The final testing will be performed when the entire system is assembled.

KEY RESEARCH ACCOMPLISHMENTS

- Development and testing of 64-channel synthetic aperture beamformer (see Appendix 2)
- Completion of 64-element array fabrication and motion system development (see Appendix 3)
- Completion of online, real-time 3D scanner converting display
- Completion and refinement of 3D data visualization (see Appendix 4)
- Completion of 3D image registration software development (see Appendix 5)
- Completion of breast imaging toolbox (see Appendix 6)

REPORTABLE OUTCOMES

Manuscripts, abstracts, presentations:

- Michael Inerfield, Geoffrey R. Lockwood, and Steven L. Garverick, "A sigma-delta-based synthetic aperture beamformer for real-time 3-D ultrasound," *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control* (in press).

- Christopher R. Hazard and Geoffrey R. Lockwood, "Real-time synthetic aperture beamforming: practical issues for hardware implementation", *Ultrasonics Symposium Proceedings*, 2001 (in press).

- Vladimir Zagrodsky, Raj Shekhar and J. Fredrick Cornhill, "Multifunction extension of simplex optimization method for mutual information based registration of ultrasound volumes," presented at SPIE Medical Imaging Symposium, 2001.

- Raj Shekhar and Vladimir Zagrodsky "Interactive visualization of four-dimensional ultrasound data," presented at IEEE Visualization conference, 2001.

- Raj Shekhar and Vladimir Zagrodsky, "Mutual information-based rigid and nonrigid registration of ultrasound volumes," IEEE Transactions on Medical Imaging (in press).

- Radhika Sivaramakrishna, Kimerly A. Powell, Michael L. Lieber, William A. Chilcote and Raj Shekhar, "Texture analysis of lesions in breast ultrasound images," submitted to Journal of Digital Imaging.

Degrees obtained

- Christopher R. Hazard, Doctor of Philosophy, The Ohio State University, 2001
  Dissertation Title: Real-time three-dimensional ultrasound imaging using synthetic aperture imaging,

  Michael Inerfield, Doctor of Philosophy, Case Western Reserve University, 2001
  Dissertation Title: Sigma-delta modulation in real-time three-dimensional sparse synthetic aperture ultrasound imaging systems

Funding obtained

- Development of quantitative real-time 3D stress echocardiography (PI - Raj Shekhar)
  Funding Agency: The Whitaker Foundation
  Duration and Total Costs: 3 years (Sept 01-Aug 04) and $239,939

CONCLUSIONS

We have developed a method for high-speed ultrasound imaging. The method permits the acquisition of 40 to 80 images in the time normally required to collect a single image. The increased acquisition speed will be used to collect a 3D data set in real-time. The objective of this proposal is to build a prototype scanner and develop the software tools required to analyze and interpret the 3D data.

The beamformer is integral to allowing high-speed ultrasound imaging and collection of 3D data sets in real-time. We reported previously assembly of the hardware and development of optimized assembly language routines required for high-speed beamforming. In the past year, all hardware and software developments were completed. The beamformer was tested with a prototype probe along with its motion system. Furthermore, this beamformer was successfully interfaced with the real-time scan converting display, whose development is also finished.

A unique rocking probe permits the collection of 3D data. A prototype probe was delivered by Tetrad Corporation earlier in the year for testing the beamformer and other components of the system. The final probe as per the original specifications was delivered to us in December 2001. The development of all scanner comments is complete and we are currently assembling the full system.

The user interface and image analysis software development focused on refining, testing and integrating the developed tools. As described above, scan converting display, crucial for previewing images online and positioning the probe, was completed. 3D visualization for "virtual re-scanning" a patient was fully developed and fully integrated into our application's final user interface. Image registration and segmentation tools, the building blocks of the cardiac imaging toolbox have been developed and are being ported to the main application. The development is also over for the breast imaging toolbox.

Real-time 3D imaging combined with tele-medicine and a set of image analysis tools will enable ultrasound imaging for forward echelon combat casualty care. This technology will be equally effective in civilian emergency care.

APPENDICES

Appendix 1: Preprint of the publication - Michael Inerfield, Geoffrey R. Lockwood, and Steven L. Garverick, "A sigma-delta-based synthetic aperture beamformer for real-time 3-D ultrasound," *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control.*

Appendix 2: Preprint of the publication - Christopher R. Hazard and Geoffrey R. Lockwood, "Real-time synthetic aperture beamforming: practical issues for hardware implementation", *Ultrasonics Symposium Proceedings*, 2001.

Appendix 3: Progress report from Tetrad Corporation.

Appendix 4: Reprint of the publication - Raj Shekhar and Vladimir Zagrodsky "Interactive visualization of four-dimensional ultrasound data," presented at *IEEE Visualization* conference, 2001.

Appendix 5: Preprint of the publication - Raj Shekhar and Vladimir Zagrodsky "Mutual information-based rigid and nonrigid registration of ultrasound volumes," *IEEE Transactions on Medical Imaging.*

Appendix 6: Submitted manuscript - Radhika Sivaramakrishna, Kimerly A. Powell, Michael L. Lieber, William A. Chilcote and Raj Shekhar, "Texture analysis of lesions in breast ultrasound images," submitted to Journal of Digital Imaging.

Appendix 1: Preprint of the publication - Michael Inerfield, Geoffrey R. Lockwood, and Steven L. Garverick, "A sigma-delta-based synthetic aperture beamformer for real-time 3-D ultrasound," *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control.*

A Sigma-Delta-Based Sparse Synthetic Aperture Beamformer for Real-Time 3-D Ultrasound

Michael Inerfield, *Student Member, IEEE*, Geoffrey R. Lockwood, *Member, IEEE*,
Steven L. Garverick, *Senior Member, IEEE*,

***Abstract*—Sigma-delta modulation allows delay resolution in ultrasound beamformers to be achieved by simple clock cycle delays applied to the undecimated bit-stream, greatly reducing the complexity of the signal processing and the number of bits in the datapath. The simplifications offered by this technique have the potential for low power and portable operation in advanced systems such as three-dimensional and color Doppler imagers. In this paper, an architecture for a portable, real-time, three-dimensional sparse synthetic aperture ultrasound beamformer based on sigma-delta modulation is presented, and its simulated performance is analyzed. Specifically, with a 65-element linear phased array and 3 transmit events, this architecture is shown to achieve a 1.1^0 beam width, a -54-dB secondary lobe level, and a theoretical frame rate of 1,700 frames/sec, at $\lambda/64$ delay resolution using a second-order lowpass sigma-delta modulator. Finally, a technique for modifying the proposed multi-beam architecture to allow improved A/D resolution by premodulating the input signal for bandpass sigma-delta modulation is also presented.**

This work was supported by the United States Army Medical Research and Material Command, Grant No. DAMD17-99-1-9034.

Michael Inerfield and Steven L. Garverick are with the Department of Electrical Engineering and Computer Science, Case Western Reserve University, Cleveland, OH, USA 44106.

Geoffrey R. Lockwood was with the Cleveland Clinic Foundation, Cleveland, OH USA 44106. He is now with the Department of Physics, Queen's University, Kingston, Ontario, Canada K7L3N6.

I. INTRODUCTION

Since the 1950's, medical ultrasound imaging has progressed from simple, analog A-mode imaging to far more sophisticated digital B-mode and color Doppler systems. These newer digital systems have greatly benefited from improvements in semiconductor electronics, which have improved the speed and accuracy of the necessary data conversion. While these advancements have resulted in high-quality, two-dimensional real-time imagers that are used regularly in hospitals, an extension of this technology to produce three-dimensional real-time images of comparable quality has not yet been realized. Recent advances in semiconductor electronics have made it possible to develop ultrasound beamformers that can acquire and process data at a sufficiently high rate to produce real-time 3-D images. Due to limitations imposed by the speed of sound, a real-time 3-D imager based on a linear phased array must form 2-D images from a relatively small number of transmit events to avoid distortion due to movement. This requirement is well matched with the sparse synthetic aperture technique proposed in [1]. Since a significant fraction of a 2-D data set has to be collected and processed over the course of a single transmit event, beam formation must either occur serially with a great amount of memory and speed, or in parallel using duplication of hardware.

A sparse synthetic aperture system has recently been proposed that forms beams serially, based on commercially available Nyquist-rate analog-to-digital converters (A/Ds) and digital signal processors (DSPs) [2]. In this system, data collected from each of the receive channels is stored in memory. A DSP must serially form all of the beams from the collected data before acquisition of the next set of data occurs. Serial beam formation in this system places a heavy burden on the DSPs, thus the filter applied to form the fine delays must be relatively simple. Even the simple interpolation algorithm proposed in [2] requires DSPs running at 4 or more times faster than the A/D converter. The need for simplicity in the delay calculation forces the need for an A/D that samples well in excess of the Nyquist rate to reduce error in calculation of the fine delays. While such a system is feasible, there is great potential for further hardware reduction.

Duplication of conventional circuitry for parallel beamforming would require an overwhelming amount of hardware. For example, a system with a 65-element array forming 90 beams would require 5,850 beamformer channels, each consisting of dynamically changing delays and DSP circuitry. The duplication of dynamically changing delays can be alleviated if multiple beams share the same focusing delays, with some sacrifice to image quality, as proposed in [3], [4]. However, a system using a conventional Nyquist-rate A/D converter would still be of excessive size due to the required number of channels.

An interesting method of achieving such a high data conversion rate, while making major simplifications in the signal processing hardware, is the use of a sigma-delta ($\Sigma\Delta$) modulator as the A/D converter in the beamformer, as proposed by [5], and developed in greater detail in [6], [7], [8]. When a sigma-delta modulator is operated with a sampling frequency (f_s) of 32x the center frequency (f_c) of the ultrasound signal or higher in a conventional system [5], the number of bits involved in the signal processing is greatly reduced, and the need to interpolate between samples is eliminated. It then becomes feasible to achieve parallel beam formation via duplication of hardware, which has been drastically simplified over the conventional case. The price of this simplicity is the need for a high-speed sigma-delta modulator.

Although such high-speed operation of the sigma-delta modulator might appear to be wasteful of power, sigma-delta modulators use significantly less power than Nyquist-rate A/Ds operating at the same speed, because they supply only a single bit of resolution at that speed. That single bit of resolution, however, is enough to provide the required delay accuracy, saving the conversion to high resolution until after the complex beamformer hardware, where it is needed at a much lower rate. Most importantly, the speed and size of the digital hardware is reduced due to the simplified signal processing and reduced number of bits being processed.

For the proposed beamformer, and typical values of f_c ranging from 3-5 MHz, this translates into a modulator operating at a minimum speed of 96 MHz, capable of providing at least 7 bits of resolution after decimation (see Section II). If the beamformer uses a linear phased array of length 32λ, a center frequency of 3 MHz and images to a maximum depth of 15 cm, the round trip time-of-flight for a single transmit event would be 0.2 msec. For 3 transmit events, this implies a theoretical maximum frame rate of ~1,700 frames/sec.

In this paper we propose a beamformer architecture for real-time 3-D imaging. The system combines sigma-delta A/D conversion with a sparse synthetic aperture beamforming technique. This combination provides high-speed parallel beam formation using relatively simple hardware. The remaining sections of this paper are organized as follows: Section II discusses the proposed beamformer architecture in broad detail, with emphasis placed on describing the hardware and its requirements. Section III discusses specific issues in calculating and implementing separate focusing and steering delays. A unique feature of this design is that the same set of dynamic focusing delays are used for all lines in the image, the consequences of which are quantified in this section. Section IV describes simulation results obtained with a discrete-time model of the proposed beamformer. Finally, a unique way of increasing A/D resolution by combining input signal translation (premodulation) [7] and bandpass sigma-delta modulation with this multi-beam system is proposed in Section V.

II. THE ARCHITECTURE

The requirements for a synthetic aperture beamformer for a real-time 3-D imaging system are quite different from a conventional beamformer. A synthetic aperture image is formed by exciting each array element individually. Since the resulting radiation fills the imaging plane, all the receive beams necessary for imaging can be formed simultaneously. The next element is then excited and the process is repeated. Once all the selected transmit elements have been excited, the final image is formed from the sum of the beamformed energy from each transmit event. The time required to collect a synthetic aperture image is proportional to the number of transmit elements used. Consequently it is possible to achieve frame rates of over 1,000 2-D images/s if only a few transmit elements are used per image. However, obtaining this frame rate requires a beamformer that can form all the receive beams in the time of a few transmit events. If a conventional beamformer where simply duplicated to generate simultaneous beams the resulting system would be massive.

Fig. 1 shows a beamformer architecture that greatly simplifies the hardware required for a real-time 3-D imaging system. The architecture is similar to a conventional beamformer with two important modifications. The first modification is that the conventional A/D converter is replaced with an oversampling sigma-delta modulator. This modification simplifies the design of the digital delays by reducing the number of bits. More importantly, the high clock speed of the converter reduces the delay quantization errors to a level where a fine delay correction is not necessary. The second modification is that the focusing and steering delays are separated so that the expensive focusing delays can be shared for each line in the image. The focusing delays are introduced assuming a beam steered broadside to the array. Separate steering delays are then introduced using a pipelined architecture consisting of a parallel set of steering columns with one column for each line in the image. Within a single steering column, the relative steering delay between adjacent elements is introduced, the signals are added, and then passed down the pipeline where the contribution from the next element is added. At the base of each column, the signal is downconverted and decimated to provide a high-resolution sample of the beamformed signal at the required sampling rate for scan conversion and display.

It is apparent from Fig. 1 that delay can only increase as the signal travels down the pipeline, implying that the beam can only be steered in one direction by the steering delays. To achieve steering in the opposite direction, the direction of the downward arrows in the steering

columns are reversed, and the downconversion and decimation blocks are placed at the new beam outputs. A possible reduction in hardware can be achieved by sharing the hardware for positive and negative angles, switching only the direction of data flow. This would eliminate somewhat less than half of the digital hardware (added switches to properly route the delay and adder I/O would be necessary) at the expense of requiring double the number of transmit events, thus reducing the frame rate.

For small steering angles, this architecture introduces small delays errors due to the shared focusing delays. However, for large steering angles, the delay errors can be significant fractions of a wavelength, and the radiation pattern will be degraded unless these errors are corrected. A simple way to reduce these errors is to adjust the fixed steering delays for each line to produce a perfectly focused beam at a depth somewhere in the middle of the image. On either side of this depth the radiation pattern will still be distorted. However, we will show that a surprisingly good radiation pattern can be achieved using this correction.

In addition to delay errors introduced by the shared focusing delays, errors will also be introduced by the delay and amplitude quantization, rounding errors and by imperfections associated with the sigma-delta modulator. The relative importance of each of these errors was assessed using a computer simulation of the beamformer. In the remainder of this section, the performance of the beamformer is assessed assuming an ideal analog to digital converter. A discussion of issues related to the calculation of the steering delays and a more realistic simulation including a model of a second-order sigma-delta modulator are described later.

The most important decision to be made in implementing this architecture is in the choice of delay resolution. The effects of various levels of delay quantization on the sparse synthetic aperture beamformer are shown in Fig. 2 [8]. These simulations were generated with delay quantization applied to a computer model of an ideal 65-element beamformer using 3 transmit events. No amplitude quantization is explicitly applied, thus calculations are carried out at the floating-point accuracy of the computer. The pulse used in this simulation has a 30% fractional bandwidth and is given by $\cos(2\pi\lambda)e^{-0.5\lambda^2}$ (Fig. 3). The results indicate the best performance that can be expected from a beamformer for a given delay resolution. Note that to achieve secondary lobe levels approaching –60 dB—the minimal acceptable level for most modern beamformers—the delay resolution must be $\lambda/64$ or better. With a fixed delay resolution, further improvement can be had at the expense of size and power by increasing the number of receive elements, or at the expense of frame rate by increasing the number of transmit events.

Fig. 4 shows the affects of amplitude quantization on the radiation pattern. This figure is generated at $\lambda/64$ delay resolution for a 65-element array using 3 transmit events. It shows that delay resolution dominates the levels of the secondary lobes as long as the A/D provides at least 7 bits of resolution. This requirement also approximately holds for delay resolutions of $\lambda/32$ and $\lambda/128$ which are also used in this paper.

The consequences of fixing the steering delays on the radiation pattern are illuminated in Fig. 5a and b. In this simulation an ideal beamformer model was modified to include the effects of finite delay resolution ($\lambda/64$) and amplitude quantization (7 bits). Furthermore, focusing and steering delays were separated, with the steering delays fixed at values ideal for a distance f/4 from the center of the array. These plots are generated by analyzing radiation patterns at 1^0

increments from 0^0 to 45^0 for focal distances of f/3, f/4, f/6, f/8, and f/10. Average secondary lobe levels are determined by averaging (in dB) all points in the radiation pattern at 9^0 or greater from the target location. Beam widths are the angular difference (in degrees) between –3-dB points in the radiation pattern.

The results of this experiment show that for $\lambda/64$ delay resolution, fixing the steering delays raises the average value of the secondary lobes in the radiation pattern by less than 9 dB at the most extreme steering angle (45^0) at f/10, and by a maximum of 6 dB from 0^0 to 35^0 from f/3 to f/10. The effect on beam width due to the fixed steering delays is minor until a target angle of 35^0, increasing by only 0.5^0 from the nominal value of 1.1^0 at a target angle of 0^0 from f/3 to f/10. This simulation mimics the effects of the major non-idealities in the system and is a good indication of what can be expected from a physical implementation.

III. CALCULATING THE DELAYS

The delays used to focus ultrasound energy in the proposed beamformer are separated into dynamically changing focus delays and static steering delays. The following sections outline the process of calculating those delays.

A. Focusing delays

Signal energy returns to the transducer array after an amount of time proportional to its radius from the location of the transmitter. As signals from increasing distances from the array become available, the dynamic focusing delays must change so that the receive focus is aimed at the point from which the signal is emanating. The sharing of focus delays for each beam significantly reduces the size of the hardware, as dynamic delays are expensive to implement.

Fig. 6 illustrates how focus delays are calculated for a 9-element array. Focus delays are calculated only for the 0^0 beam. Steering is achieved via an array of fixed delays whose calculation is discussed in Section IIIB. An important point in this calculation is that only the "relative delay" differences between a target point and the various elements need to be compensated during focusing. The current radius $\boldsymbol{R_i}$ (from the transmitter) of returning signal energy is tracked by a master clock, with each clock cycle i since the pulsing of the transmitter equal to an increment of $\lambda/2k$ in space ($k = f_s/f_c$). At any given point in time, the appropriate relative delay differences among the receive elements, corresponding to the current radius $\boldsymbol{R_i}$, are corrected for each element to form a focus.

Clearly a reference element is needed about which the delays for the other elements should be applied. A natural choice for this reference is the center receive element (which is the transmitter for the center transmit event), because it is located along the center axis of the image to be formed. Because the system is to generate evenly spaced points emanating from the center of the array, the center receive element has the exclusive property that during the center transmit event, the data point it collects at every clock cycle is the desired data point. This implies that the delay applied to the center receive element during the center transmit event is a constant with no rounding error. Furthermore, the delay τ_c is always the shortest, thus the delay applied to the center element, τ_z-τ_c, is always the longest. If the center receive element is used as a reference, the maximum relative delay error applied to other receive elements is half a clock cycle. Were

the center receive element not used as a reference, a maximum relative error of a full clock cycle could result, and the beamformer performance would suffer significantly.

For all other elements, the desired datum lies between two collected points, and the data point nearest to the desired one ("nearest neighbor") is used. This process is synonymous with rounding the delays to the nearest clock cycle. For the center transmit event, the focusing delays for the other elements are calculated relative to the constant delay applied to the center transducer element. Focusing delays for off-center transmit events are all calculated relative to the delay applied to the center element during the center transmit event.

It is convenient when describing the beamformer to express round trip time-of-flights and phase shifts in units of wavelengths (λ). For example a relative path distance of 1 wavelength between adjacent elements would correspond to a "phase shift" or "time difference" of 1 wavelength. The actual phase shift, in degrees, or time, in seconds, can be calculated based on the speed of sound and frequency of the transducer.

Referring again to Fig. 6, the delays applied at each clock cycle must undo the relative time-of-flight delays experienced in the medium. At clock cycle i since the transmit pulse, the beamformer focuses at the radius $\boldsymbol{R}_i$ along the axis. The flight times (τ_i) taken for the signal to travel from the center transmitter to $\boldsymbol{R}_i$ and back to the individual receive elements are contained in the vector $\mathbf{T}_{i,0^\circ}$. Flight times $\tau_{c,128k}$ and $\tau_{z,128k}$ are τ_i (in units of λ) for the center and end elements at the closest radius to be imaged, which occurs at i = 128k for an initial focal point at f/2. $\mathbf{D}_i^{\text{focus}}$ is equal to the column vector of digital delays (in clock cycles) applied to the receive array at clock cycle i, and is given by (1).

$$\mathbf{D}_i^{\text{focus}} = round\left(k[\tau_{z,128k} - \tau_{c,128k}]\right) - round\left(k[\mathbf{T}_{i,0^\circ} - \tau_{c,i}]\right) \qquad (1)$$

The first term in (1) is the constant delay applied to the center element (not including any offset required for transmitter alignment). It acts as a constant offset for the other elements, keeping the center of the "lens" in Fig. 6 fixed for transmitter 2. The center of the "lens" changes for the off-center transmit events, however. This point is discussed below. The second term contains the quantized relative delay differences between the off-center elements and the center element. $\mathbf{D}_i^{\text{focus}}$ is always a positive number.

The situation for off-center transmit events is somewhat more complicated. The relative delay applied to each element remains the same for all transmit events. This is apparent since the path length differences from the transmitter to $\boldsymbol{R}_i$ and back to the receive elements do not depend on the location of the transmitter. However, the data returning to the center receive element are no longer the desired evenly spaced points, as can be reasoned from Fig. 7. For off-center transmitters at distances +/- 16λ vertically from the center of the array, $\Delta\tau_i$ in units of λ is given by

$$\Delta\tau_i = \sqrt{(16\lambda)^2 + (m\lambda)^2} - m\lambda \qquad (2)$$

To obtain a point a distance mλ along the center axis takes $k\Delta\tau_i/f_s$ longer for an off-center transmit event than it does for the center transmit event. Thus delay offsets that decrease with increasing radius from the center must be added to each column (i) of $\mathbf{D}^{\text{focus}}$ to align the center

transmit data with the off-center transmit data. Equation (2) essentially provides an adjustment to the first term in (1) to account for the fact that the off-center transmitters are no longer in the line-of-sight of the transmitted energy. Alternatively, a constant offset can be added to all of the center delays, and the correction factors $\Delta\tau_i$ can be subtracted from the columns of the off-center focus delays. The second method keeps the center element's focus delay constant for the second transmit event, which is desirable, since the data it receives is without rounding error. The resulting focus delays for three transmit events are shown in Fig. 8.

Note that the largest focusing delay in the system, and thus the required size of the focusing delay registers, is determined by the closest imaging depth and the maximum vertical position of the various transmitters. For the system described here, with the closest imaging depth at f/2, and off-center transmitters located at +/-16λ from the center element, the focusing delay register length must be:

$$round\left(k[\tau_{z,128k} - \tau_{c,128k}]\right) + round\left(k\Delta\tau_{128k}\right). \qquad (3)$$

A final issue concerning focus delays is in the frequency with which they are allowed to change. In principle, the focus delay should be updated each clock cycle. Since a clock cycle corresponds to a distance of $\lambda/2k$, the memory required to image from f/2 to f/10 would be 32 x 2 x 8 x k, or ~33 kb per channel for $\lambda/64$ delay resolution.

However, since the data will ultimately be decimated and compressed into pixels, it makes sense to reduce the rate at which the focusing delays are updated, thus reducing the size of the memory required to store them. In 2k clock cycles the ultrasound energy travels a distance of 1λ each way in the tissue, corresponding to the approximate round-trip length traveled in forming a pixel. Simulations show that restricting the maximum delay change frequency to be $2k/f_s$ has little effect on the radiation pattern. For this reason, all simulations of the sigma-delta beamformer have delays calculated for λ increments in space and which are repeated 2k times during beam formation.

B. Steering delays

The steering delays determine the beam angle from array broadside at which the focus is to take place. Since arcs in the image must be built up simultaneously, a separate pipeline for each beam is necessary. Parallel beam formation has been previously presented in the literature for systems based on conventional beamformers in [3], [4], and [9]. In these references, parallel beams span only a few degrees, as is imposed by the finite transmit beam width. Corrections which are a linear function of receive element position are added to the focusing delays to produce adjacent beams in these systems.

The synthetic aperture system proposed here does not focus transmit energy, thus parallel processing of all of the beams in the image is possible. The small angle approximation does not apply to this system, since the beamformer is to produce beams spanning a wide range of angles in the near field of the array. The resulting corrections are not a linear function of receive element position. The nonlinearity accounts for the effective reduction in the width of the receive aperture for nonzero angles.

Steering delays do not change dynamically, and thus can only be calculated for a fixed radius from the center of the array. At this radius, the steering delays are implemented to an

accuracy within half a clock cycle. Clearly at radii offset from this reference radius, there is error between the required and applied delay. The reference radius should be chosen to minimize this error. Fig. 9 shows the error in the steering delays at 35^0 produced for a delay resolution of $\lambda/64$ for delays optimized for f/3, f/4, f/5, and f/6 with target locations varying from f/3 to f/10.

Fig. 10a and b quantify the effects of the steering delay reference radius on the radiation pattern. The figures show nominal (target at f/4, 0^0) and peak (typically occurring for targets at f/10, 45^0) average secondary lobe levels and beam widths for reference radii varying in 8λ increments from f/3 to f/5 at $\lambda/64$ delay resolution. The data for these plots was collected through a set of simulations, each identical to those that produced Fig. 4 with the exception that steering delay reference radius was varied.

The appropriate steering delays are calculated by computing the focus delays for a given beam angle at a fixed radius from the center element, and subtracting from them the corresponding focus delays (for the 0^0 beam) given by equation (1). From this the differential steering delays must be calculated for use in the architecture shown in Fig. 1. Since the data output from a particular channel is delayed by each of the steering delays that follow its entrance point to the column, the sum of those delays must equal the desired steering delay. It is important to calculate the steering delays incrementally, rounding after each successive differential delay has been calculated, and using the rounded value in the calculation of subsequent delays. If the rounded delay is not used in the calculation of successively computed delays, rounding error will propagate through the rest of the calculations and the resulting steering delays will be far from optimal.

The calculation of the steering delays is formalized in equation (4) below. $\mathbf{D}^{\mathbf{steer}}$ is a matrix with elements $d_{r,\theta}^{steer}$, each of which contain the net steering delay (in clock cycles) that should be experienced from data entering the θ–steering column from receive channel r. $\Delta\mathbf{D}^{\mathbf{steer}}$ is a matrix containing the differential steering delays required by Fig. 1. The notation for element r,θ of the matrix $\Delta\mathbf{D}^{\mathbf{steer}}$ is $\Delta d_{r,\theta}^{steer}$. Steering delays in (4) are optimal at f/4, which occurs 256k clock cycles after the center transmit event.

$$\mathbf{D}_{\theta}^{\mathbf{steer}} = k\left[\max\left(\mathbf{T}_{256k,\theta}\right) - \mathbf{T}_{256k,\theta}\right] - \mathbf{D}_{256k}^{focus}$$
$$\Delta d_{r,\theta}^{\mathbf{steer}} = round\left(d_{r,\theta}^{steer} - \sum_{j=1}^{r-1} \Delta d_{j,\theta}^{steer} \right) \qquad (4)$$

In general, the steering delays calculated according to (4) increase or decrease monotonically, depending on the steering angle. The direction of data flow is always chosen so that the steering delays progressively increase, with r=1 being the last row to enter each steering column (see Fig. 1). The delay for element r=1 is thus 0 (not including alignment corrections). Because the differential delays must be positive, the effective steering delay each channel sees is positive. Thus, there is an added delay bias to each beam. This additional delay does not disrupt the focusing of the data; it simply delays the time at which pixels are output. Delay can be added at the end of the steering column to realign the beams in time, or more efficiently, after the downconversion and decimation stage. A similar effect occurs between the equivalent beams

from multiple transmit events. These beams must be realigned in time before they are summed for quadrature downconversion.

The set of fixed differential steering delays, ΔD^{steer}, (not including the realignment delays) for beams 0^0 to 45^0 in the architecture in Fig. 1 are shown in Fig. 11. It should be noted that the rounding algorithm occasionally produces a negative steering delay. This can be implemented simply by adding 1 to every steering delay and 65-r to every focus delay.

IV. SYSTEM SIMULATIONS

The prototype system to be simulated uses 3 transmit events and 65 receive channels, a nominal carrier frequency (f_c) of 3 MHz, a 30% fractional bandwidth, and a 2^{nd}-order sigma-delta modulator. The system is tested at delay resolutions of $\lambda/32$ (f_s=96 MHz), $\lambda/64$ (f_s=192 MHz), and $\lambda/128$ (f_s = 384 MHz).

Fig. 12a. shows a model of the 2^{nd}-order sigma-delta modulator to be used in this beamformer. The modulator can be tested by inputting a full-scale sinusoid (scaled by 0.78 at the input to avoid overload) and calculating the SNR produced at the modulator output. Fig. 12b shows the noise spectrum produced by the 2^{nd}-order modulator operating at a sampling frequency of 192 MHz converting a 3-MHz input. The SNR produced over a 3.45-MHz bandwidth in this simulation is 53.1 dB.

Fig. 13 shows the downconversion and decimation algorithm implemented in the beamformer model. In this figure, TX #1-3 are the steered data for a particular beam for transmit events 1-3. A high-order lowpass filter with a 3.45-MHz bandwidth removes the high-frequency quantization noise produced by the sigma-delta modulator. The output of this filter can then be downsampled to 12 MHz ($4f_c$), where quadrature downconversion can take place with a simple delay implementing the 90^0 phase shift. This algorithm allows idealized downconversion of the signal to take place for the purposes predicting the best possible performance that can be produced with this system. It is unrealistic for a practical implementation, however. In a real hardware implementation, the high-order lowpass filter and decimator would typically be replaced with one of the $sinc^K$-type filters described in [10] that are usually used in decimating sigma-delta bit streams.

Combining the focusing and steering delays calculated in Section III with the sigma-delta modulator and downconversion/decimation blocks, it is possible to do transient simulations of the architecture as it is shown in Fig. 1. It should be noted that the model also includes a "delay control" block as indicated in Fig. 1. This block senses delay changes and zeros out any repeated samples so they do not corrupt the encoded sigma-delta bit stream, a technique presented in [6]. Because +1 and –1 are the only valid modulator outputs, allowing a zero output from the focus delay register requires the register to have 2-bit outputs.

The resulting beamformer model is constructed exclusively of components that have direct circuit implementations (comparators, delays, adders, gain blocks, and shift registers), and thus should accurately predict the performance of a real circuit implementation. The sigma-delta architecture model must generate and process every sample the real system would—from the firing of each transmitter to the last point collected for the image—to accurately model the

operation of the sigma-delta modulator. This model is significantly more computationally intensive than the ideal beamformer model (with approximations of nonidealities) that produced Fig. 2, Fig. 4, Fig. 5, and Fig. 10, which computes data only at the desired pixel locations. The idealized simulations and those of the sigma-delta architecture are compared for delay resolutions of $\lambda/32$, $\lambda/64$, and $\lambda/128$ in Fig. 14a, b, and c. Excellent agreement has been obtained.

V. MODIFICATION FOR HETERODYNING

Heterodyning or premodulating the input signal is a useful way of increasing the number of effective bits the sigma-delta modulator can generate for systems which use color Doppler processing and require greater resolution from the A/D [7]. Premodulating the input signal poses a problem for a multi-beam architecture like this one because the phase of the premodulator must be adjusted to allow for the dynamic phase shifts introduced by the focus delays and the static phase shifts introduced by the steering delays. The need for this phase adjustment is briefly explained below. A thorough explanation is given in [7].

A changing of focus delays in a digital beamformer repeats samples. It is desired that the *output of the focus delay register* be modulated by the premodulation signal. However, since the output of the focus delay register includes the repeated samples, this cannot be the case unless the premodulation phase accounts for the inclusion of those extra samples. Thus the premodulation phase must be adjusted dynamically to account for changing focus delays. Since the focus delays are shared by all beams, their phase correction can be implemented by the delay control, as indicated in Fig. 15.

The concurrent formation of multiple beams, however, does pose a problem. Steering delays introduce constant phase shifts between the various receive channel rows which need to be summed coherently. Since the outputs of each of the focus delay registers are shared by all of the beams, they must be individually corrected for each combination of row and column before summation. Were the modulation sequence not composed of +1's and –1's, this might pose a daunting problem.

If this architecture were combined with an A/D converter that did not use discriminatory noise shaping, the problem could be fixed for the case of premodulation by $f_s/4$, as described in Fig. 16. In this figure, the input signal is assumed to be a constant 1, modulated by the sequence 1, 1, -1, -1. A constant input sequence is used for purposes of illustration, so that the effect of modulation and clock cycle phase shifts can be easily discriminated. This technique works for more realistic inputs as well.

The basic idea is to shift the phase of one of the two signals being added by multiplying it by a sequence of +1's and –1's. This can be implemented fairly simply in the architecture shown in Fig. 15 by using the extra bit already added for sample nulling as a sign bit to be toggled. Other methods are possible as well. For the case of the modulation signal 1, 1, -1, -1, only four possible phase-shifting sequences are required to allow coherent summations between receive channels. A steering delay of 1 clock cycle is corrected by multiplication of the input to the summation block by the sequence 1, -1, 1, -1. For a steering delay of 3, the sequence is -1, 1, -1, 1. For a steering delay of 2, the whole input sequence is negated. No correction is needed for a steering delay of 4. Cases with steering delays greater than 4 are equivalent to one of the above

cases for reasons of periodicity. Since the phase correction sequences are either constants or simple, clock-divided sequences, this solution requires little memory or other extra hardware. The proper sequence need only be routed to the appropriate input of the adder/subtractor.

The situation is more complicated when used with a noise-shaping A/D. For the special case of premodulation by $f_s/4$ in combination with a bandpass modulator [8], the two phase correction sequences for odd steering delays clearly re-modulate the signal by $f_s/2$. Modulation by $f_s/2$ reflects the signal in frequency over the axis $f=f_s/4$. In a non-noise shaping system, this has no consequences, as the signal after premodulation is symmetric about $f_s/4$. Unfortunately, this is not the case for the sigma-delta modulator, whose noise shaping destroys that symmetry.

However, the problem for the bandpass sigma-delta modulator with $f_s/4$ premodulation is correctable if odd steering delays are not allowed. This eliminates the possibility of odd phase shifts, meaning phase correction is achieved simply by selecting whether an addition or subtraction should take place at the summer locations in Fig. 15. This effectively halves the resolution of the steering delays, and may require operation at double the speed necessary to satisfy quantization resolution requirements alone. However, once this price is paid, the increase in SNR achievable with premodulation can be much greater than is achievable without premodulation, depending on the system specifications. If a lowpass modulator without premodulation were used in this system, its speed would typically have to be at least doubled to provide adequate SNR. For the case of the bandpass modulator with premodulation, its speed must be doubled to compensate for degraded delay resolution. However, given that the speed penalty must be paid in both situations, the bandpass modulator solution will usually provide significantly more A/D resolution.

Although this premodulation scheme would also work for a lowpass sigma-delta-based multi-beam system, the premodulation frequencies used to translate the input signal close to dc are much lower than $f_s/4$, because of the delay resolution requirement. This means the steering delay resolution degradation required to avoid remixing the signal would be much worse, and is thus not a practical solution for premodulation with lowpass converters.

VI. CONCLUSION

An architecture for a portable, real-time, three-dimensional sparse synthetic aperture ultrasound beamformer based on sigma-delta modulation was presented and analyzed. The beamformer is capable of achieving frame rates of up to 1,700 frames/sec. Practical issues concerning its implementation were developed in detail. Of particular interest is the computation of the steering delays to minimize the error normally incurred by sharing the focus delays for all beams. A nonlinear model of the beamformer constructed out of signal processing primitives was developed and its performance compared to results predicted from an ideal beamformer model at various delay resolutions. The sigma-delta beamformer proposed is capable of producing a radiation pattern with secondary lobes at –54 dB and a beam width of 1.1^0 at $\lambda/64$ delay resolution. Finally, a technique for modifying the architecture to allow for premodulation specifically suited for bandpass sigma-delta modulation of the input signal was discussed.

REFERENCES

[1] G. R. Lockwood, J. R. Talman, and S. S. Brunke, "Real-time 3-D ultrasound imaging using sparse synthetic aperture beamforming," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, vol. 45, no. 4, pp. 980-988, Jul, 1998.

[2] C. R. Hazard and G. R. Lockwood, "Theoretical assessment of a synthetic aperture beamformer for real-time 3-D imaging," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, vol. 46, no. 4, pp. 972-980, Jul, 1999.

[3] D. P. Shattuck, M. D. Weinshenker, S. W. Smith, and O. T. von Ramm, "Explososcan: A parallel processing technique for high speed ultrasound imaging with linear phased arrays," *J. Acoust. Soc. Amer.*, vol. 75, no. 4, pp. 1273-1282, 1984.

[4] O. T. von Ramm and S. W. Smith, "High speed ultrasound volumetric imaging system—Part II: Parallel processing and image display," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, vol. 38, no. 2, pp. 109-115, Mar. 1991.

[5] S. E. Noujaim, S. L. Garverick, and M. O'Donnell., "Phased array ultrasonic beam forming using oversampled A/D converters," Apr. 1993, U.S. patent no. 5,203,335.

[6] S. R. Freeman, M. K. Quick, M. A. Morin, R. C. Anderson, C. S. Desilets, T. E. Linnenbrink, and M. O'Donnell, "Delta-sigma oversampled ultrasound beamformer with dynamic delays," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, vol. 46, no. 2, pp. 320-331, March 1999.

[7] S. R. Freeman, M. K. Quick, M. A. Morin, R. C. Anderson, C. S. Desilets, T. E. Linnenbrink, and M. O'Donnell, "Heterodyning technique to improve performance of delta-sigma-based beamformers," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, vol. 46, no. 4, pp. 771-789, July 1999.

[8] M. Inerfield, S. L. Garverick, and G. R. Lockwood, "Bandpass sigma-delta modulation in ultrasound beamforming," submitted to *IEEE Trans. Circuits Syst.—II.*

[9] M. O'Donnell, "Efficient parallel receive beam forming for phased array imagine using phase rotation," Proc. 1990 IEEE Ultrasonics Symposium, vol. 3, p 1495-1498, 1990.

[10] S. R. Norsworthy and R. E. Crochiere, "Decimation and interpolation for $\Delta\Sigma$ conversion," in *Delta-Sigma Data Converters, Theory, Design, and Simulation,* ed. S. R. Norsworthy, R. Schreier, and G. C. Temes, New York: IEEE Press, 1997.

FIGURE CAPTIONS

Fig. 1. Three-dimensional beamformer architecture

Fig. 2. Effect of delay quantization. These radiation patterns for $\lambda/16$, $\lambda/32$, $\lambda/64$, $\lambda/128$, and no delay quantization are formed using a point target at f/4, 0^0 with 3 transmit events. The receive element array has 65 elements spaced $\lambda/2$ apart. Reprinted from [8].

Fig. 3. Ultrasound pulse with 30% fractional bandwidth, as used in all beamformer simulations.

Fig. 4. Effect of amplitude quantization. Radiation patterns are for $\lambda/64$ delay resolution and are formed using a point target at f/4, 0^0 with 3 transmit events. The receive transducer array has 65 elements.

Fig. 5a. and b. (a) Average secondary lobe level vs. target angle at various target radii for radiation patterns produced with a fixed-steering delay beamformer. (b) Beam width vs. target angle at various target locations for radiation pattern produced with a fixed-steering delay beamformer. The steering delays were calculated at a reference distance f/4 from the center of the array. The delay resolution is $\lambda/64$, amplitude quantization is 7 bits, and three transmit events were used.

Fig. 6. Calculation of focusing delays for 9-element array. The transducer array is labeled with travel times τ from the transmitter to each receive element.

Fig. 7. Time-of-flight difference between center and off-center transmit events.

Fig. 8. Focus delays for three transmit events, $\lambda/64$ delay resolution.

Fig. 9. Steering delay error (λ) at 35^0 vs. receive element #. Fixed steering delays are calculated at $\lambda/64$ accuracy at focal distances of (a) f/3, (b) f/4, (c) f/5, and (d) f/6 and subtracted from the ideal delays at focal distances ranging from f/3 to f/10.

Fig. 10. (a) Nominal (target at f/4, 0^0) and peak average secondary lobe levels (dB) for steering delay references from f/3 to f/5. (b) Nominal (target at f/4, 0^0) and peak beam widths (deg) for steering delay references from f/3 to f/5. The delay resolution is $\lambda/64$, and the amplitude quantization is 7 bits.

Fig. 11. Steering delays, $\Delta \mathbf{D}^{\mathbf{steer}}$, for selected beams from 0^0 to 45^0 for $\lambda/64$ delay resolution.

Fig. 12a. and b. (a) Second-order sigma-delta modulator model. (b) Spectrum for a 3-MHz input and 192-MHz sampling. Simulated SNR is 53.1 dB.

Fig. 13. Downconversion and decimation algorithm.

Fig. 14. Comparison of radiation patterns generated by ('o') a detailed model of the proposed architecture with a second-order sigma-delta modulator as the A/D and by ('-') an ideal beamformer model with delay quantization. Delay resolutions are a) $\lambda/32$, b) $\lambda/64$, and c) $\lambda/128$. The receive array has 65 elements spaced $\lambda/2$ apart, and a 3-MHz, 30% fractional bandwidth pulse was used to generate 3 transmit events.

Fig. 15. Receive channel structure modified to allow premodulation.

Fig. 16. Premodulation phase correction for a multi-beam architecture. A premodulation sequence (1, 1, -1, -1) is used to modulate a constant (1) input.

Fig. 1., Inerfield, et. al.

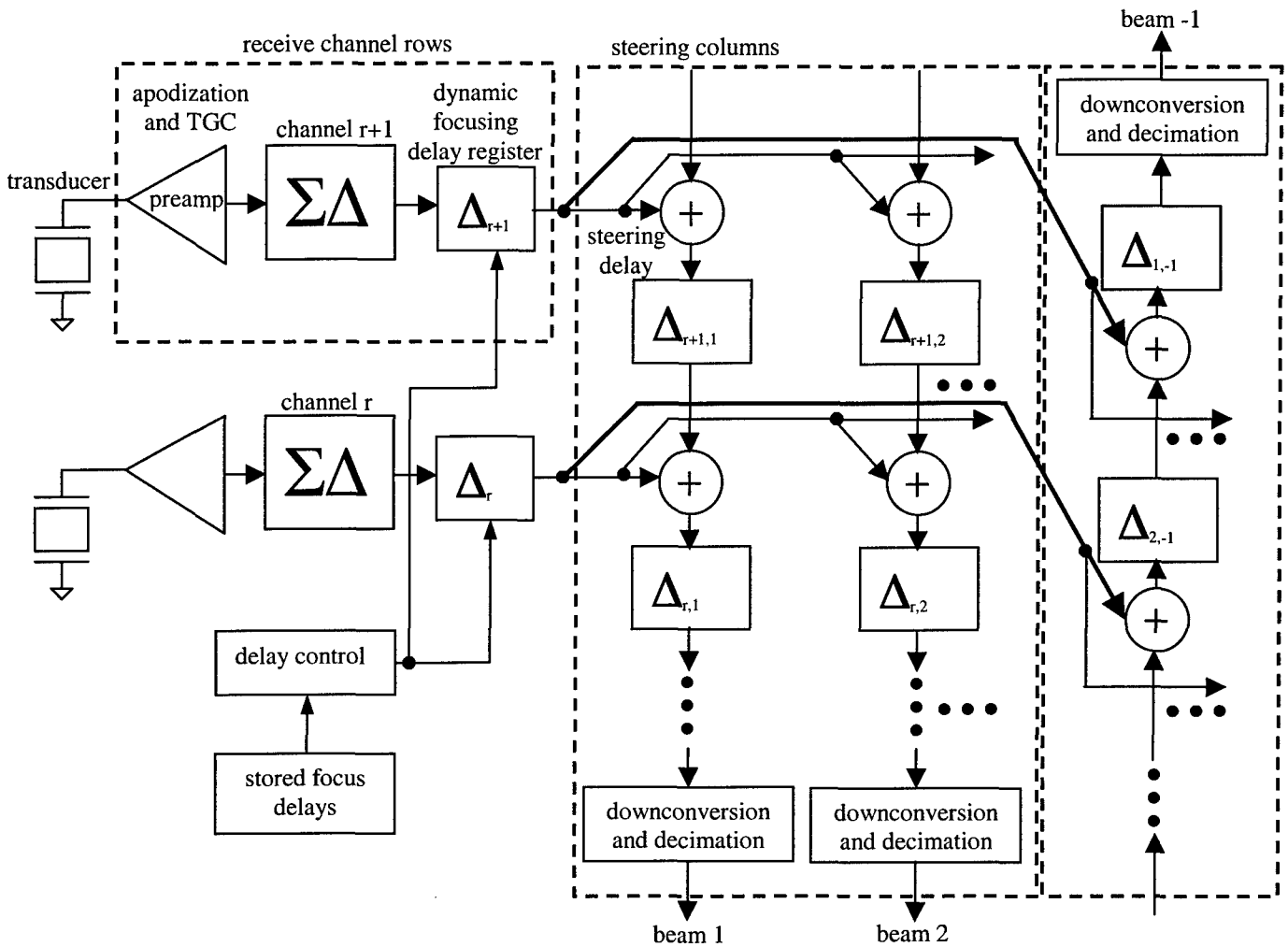


Fig. 2. Inerfield, et. al.

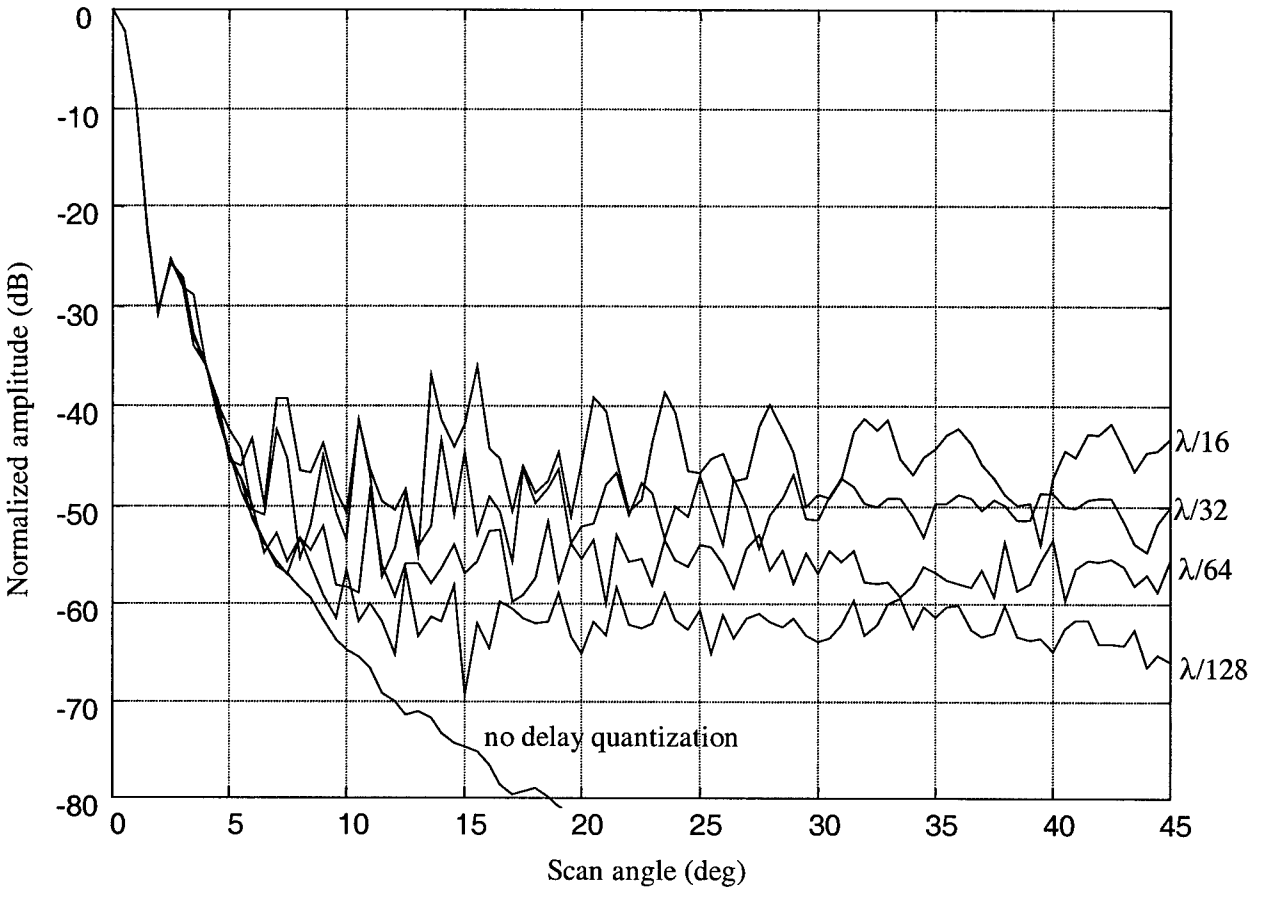


Fig. 3. Inerfield, et al.

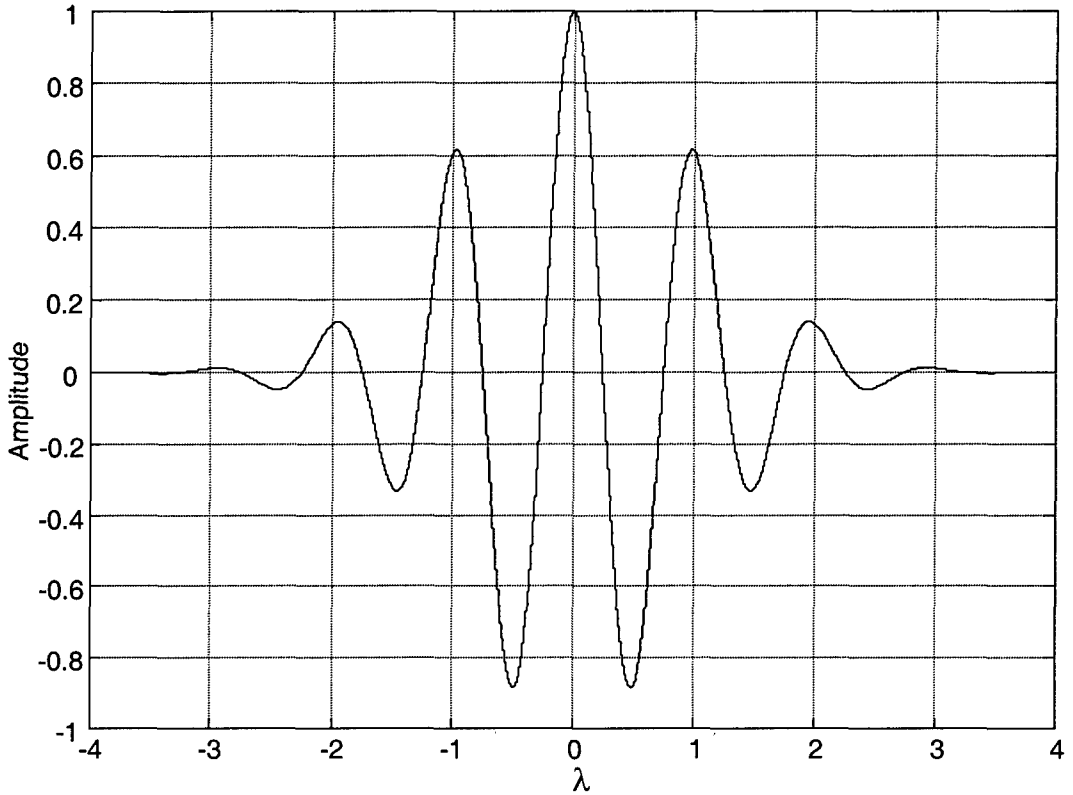


Fig. 4. Inerfield, et. al.

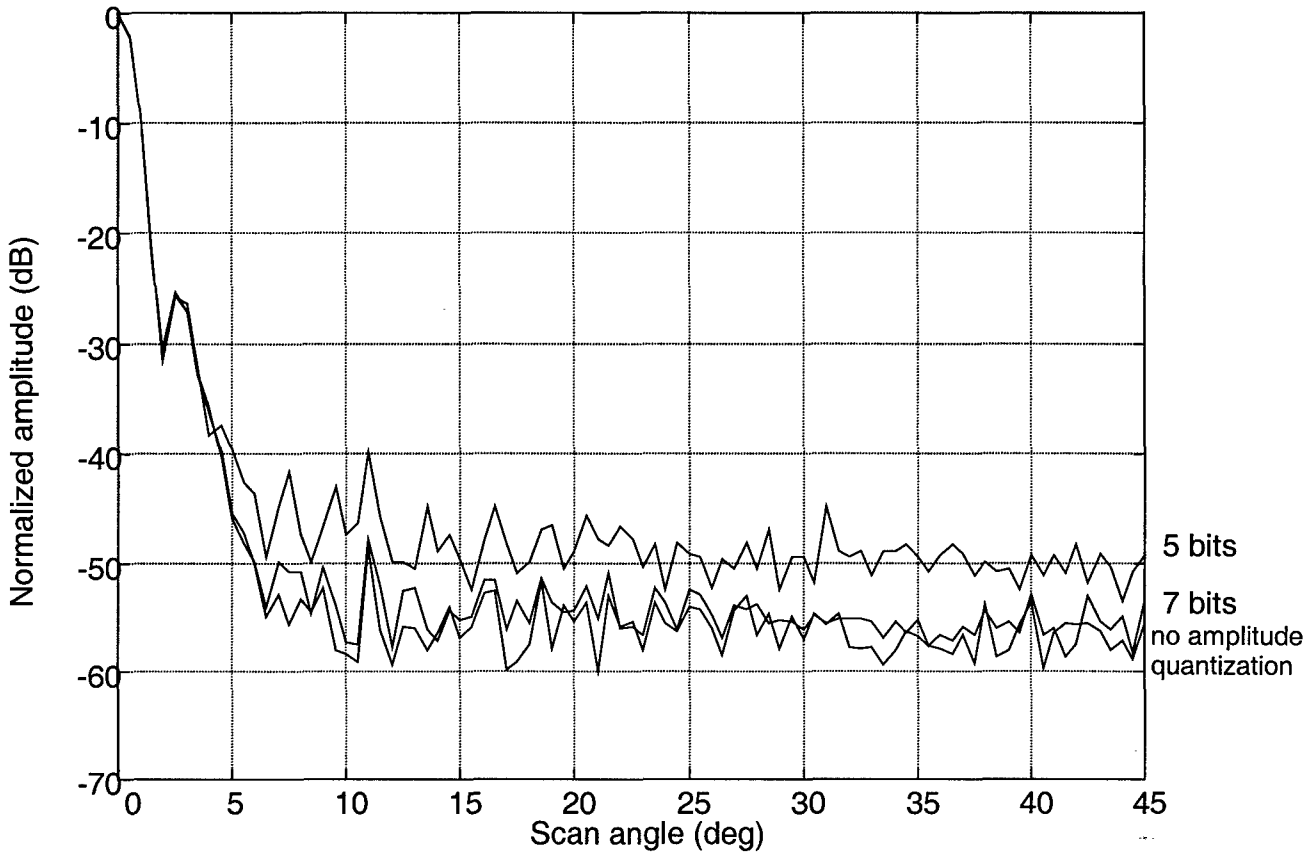
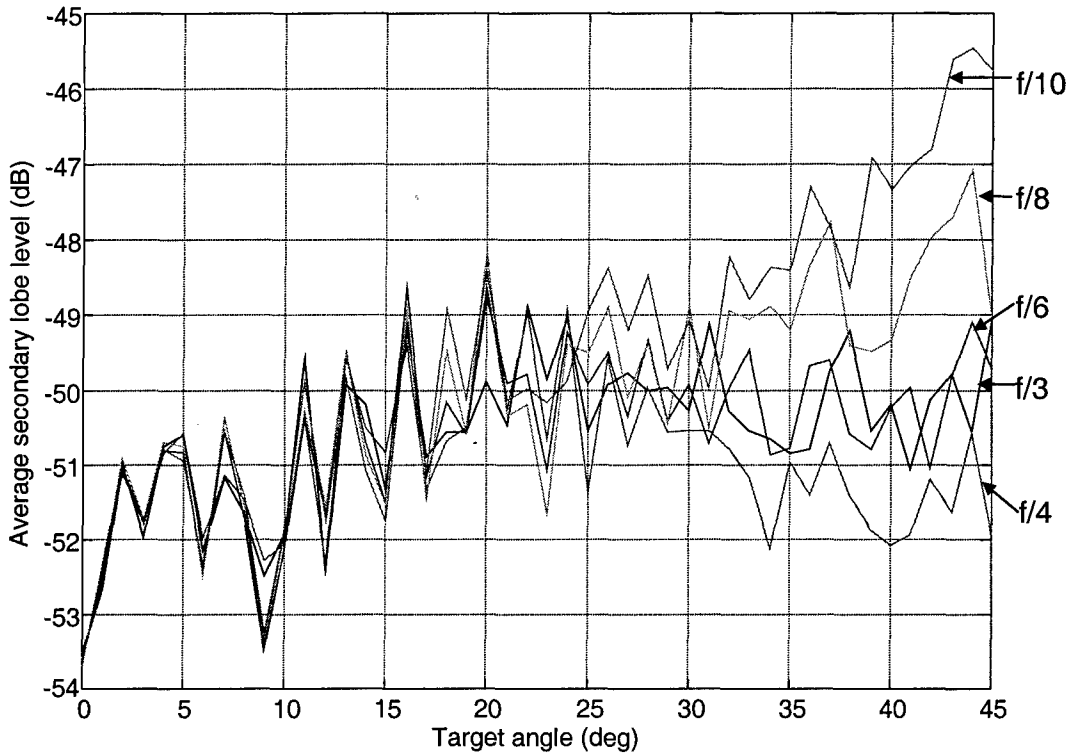


Fig. 5a and Fig. 5b., Inerfield, et. al.

(a)



(b)

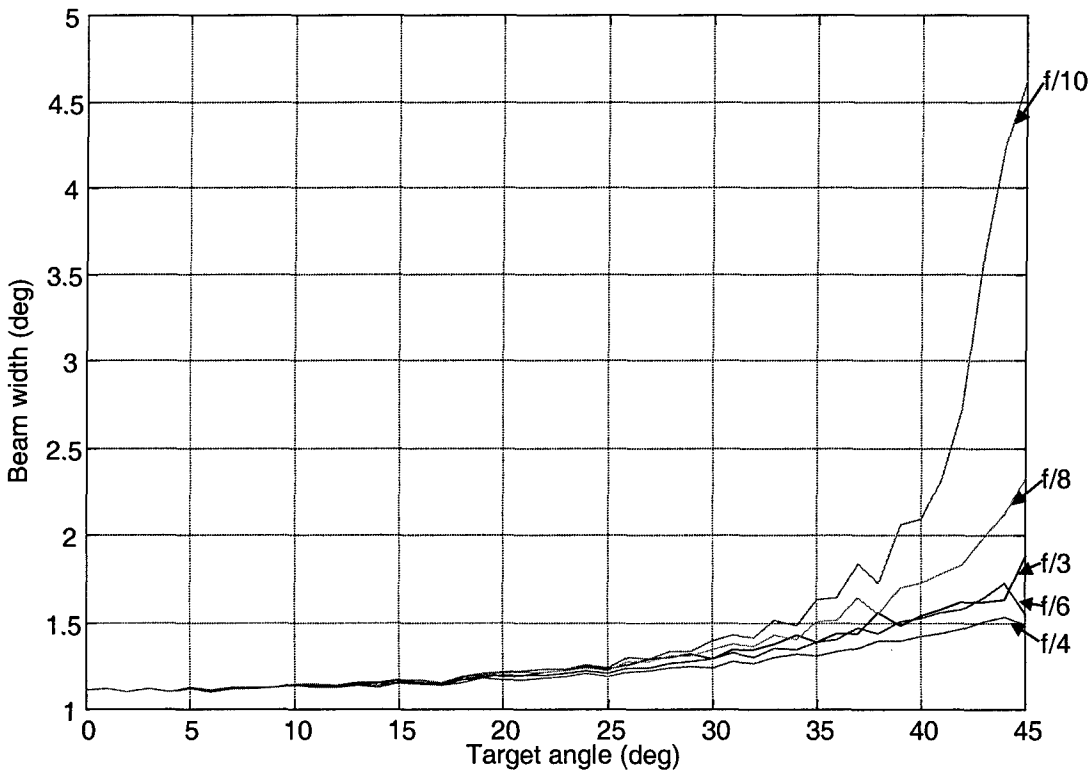


Fig. 6., Inerfield, et. al.

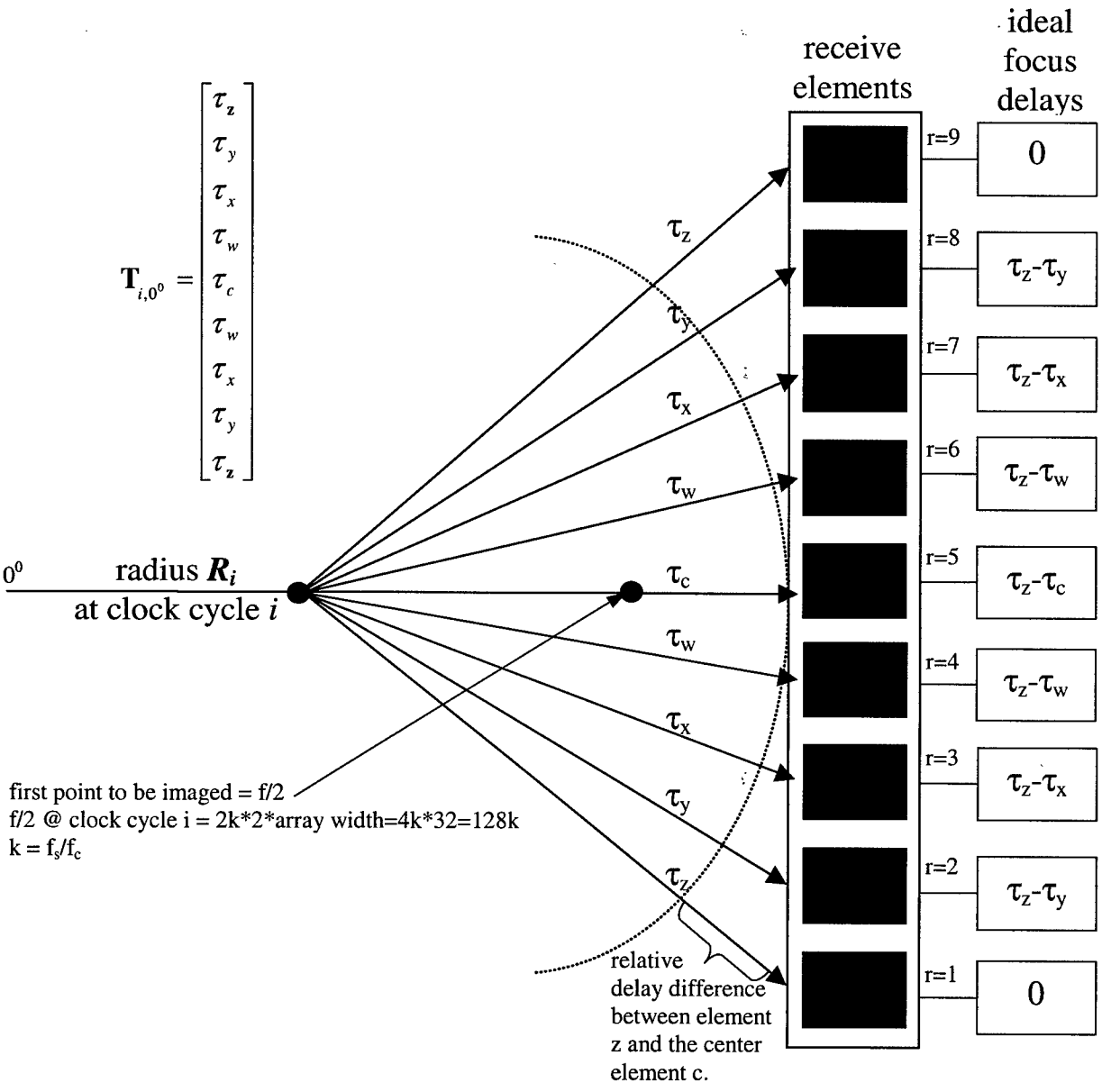


Fig. 7. Inerfield, et al.

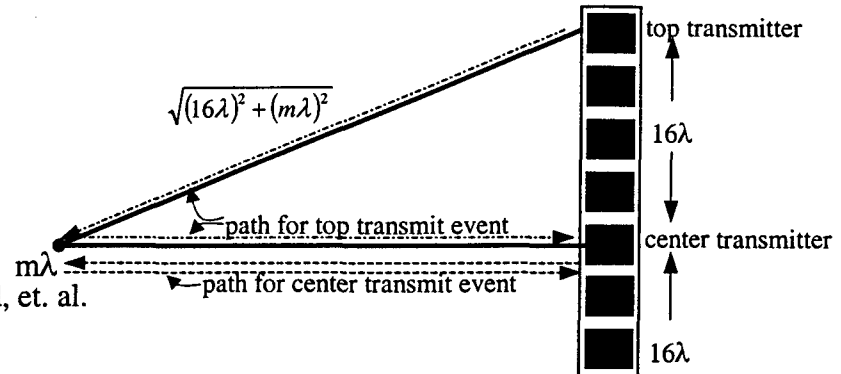


Fig. 8., Inerfield, et. al.

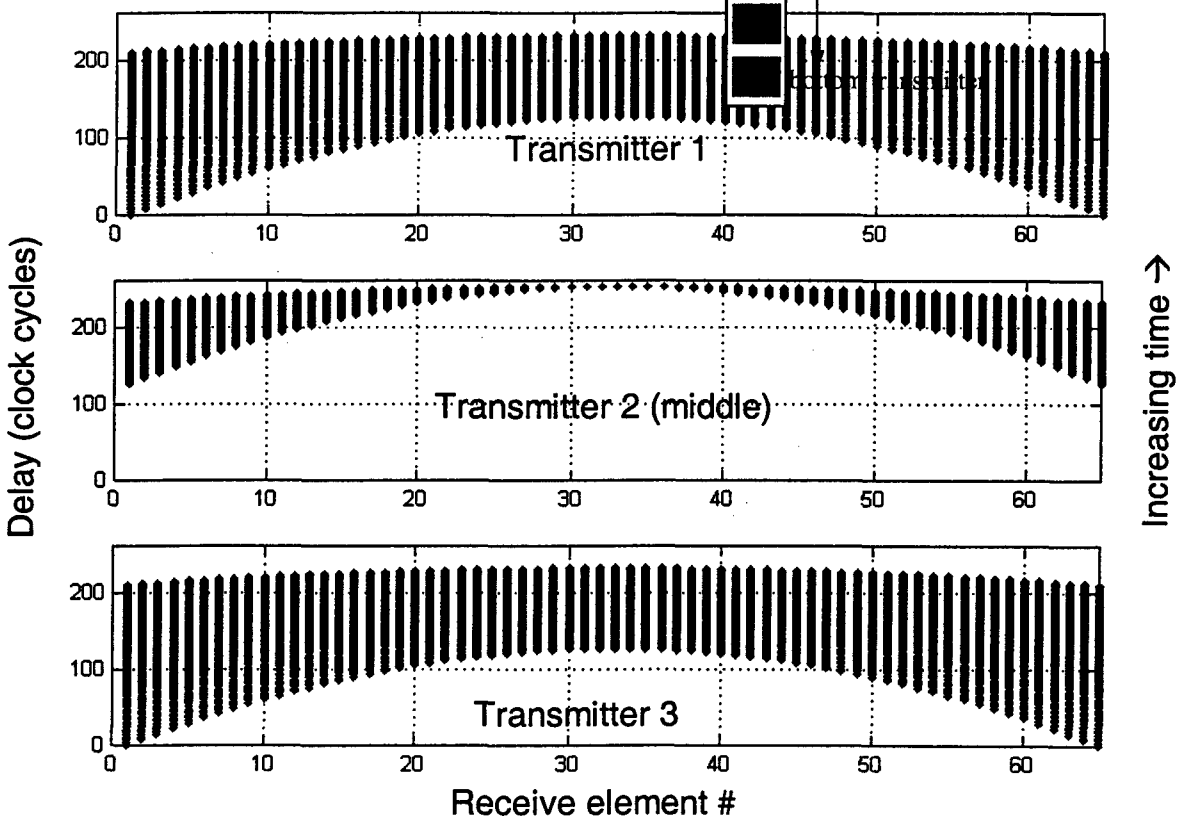


Fig. 9. Inerfield, et al.

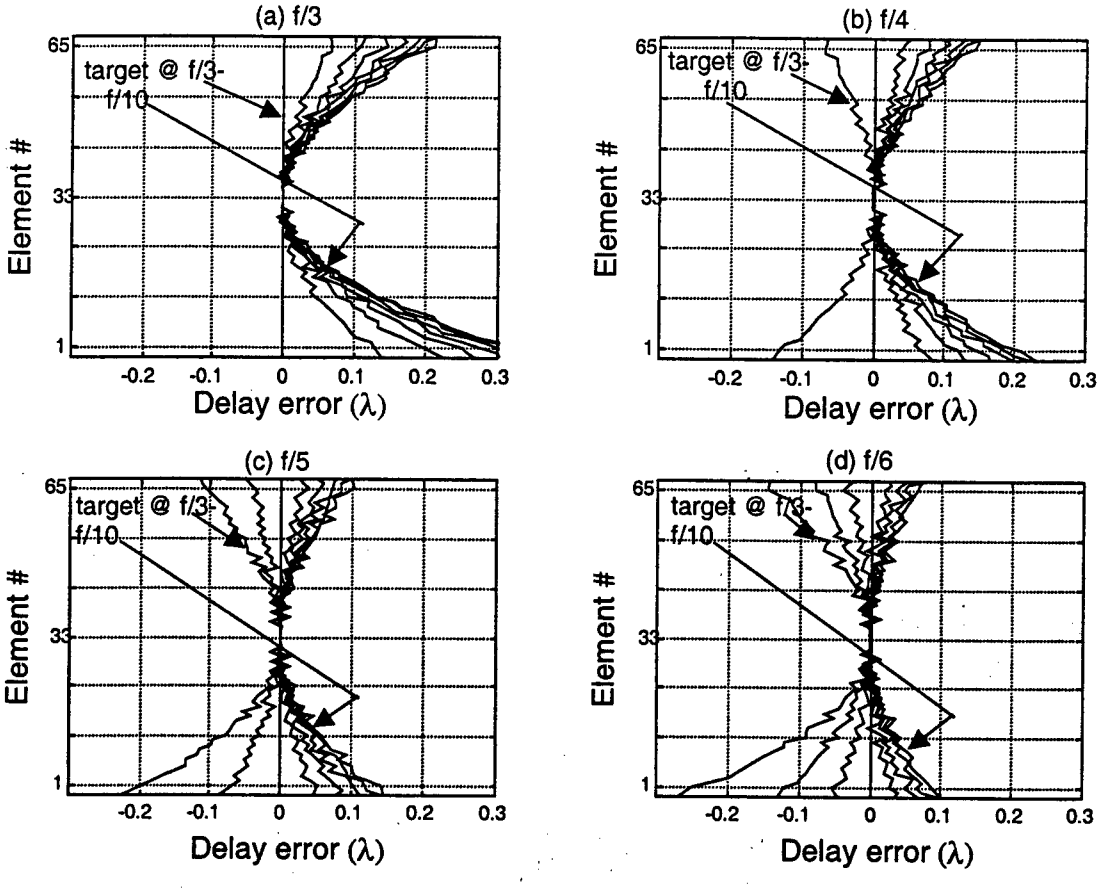


Fig. 10. Inerfield, et al.

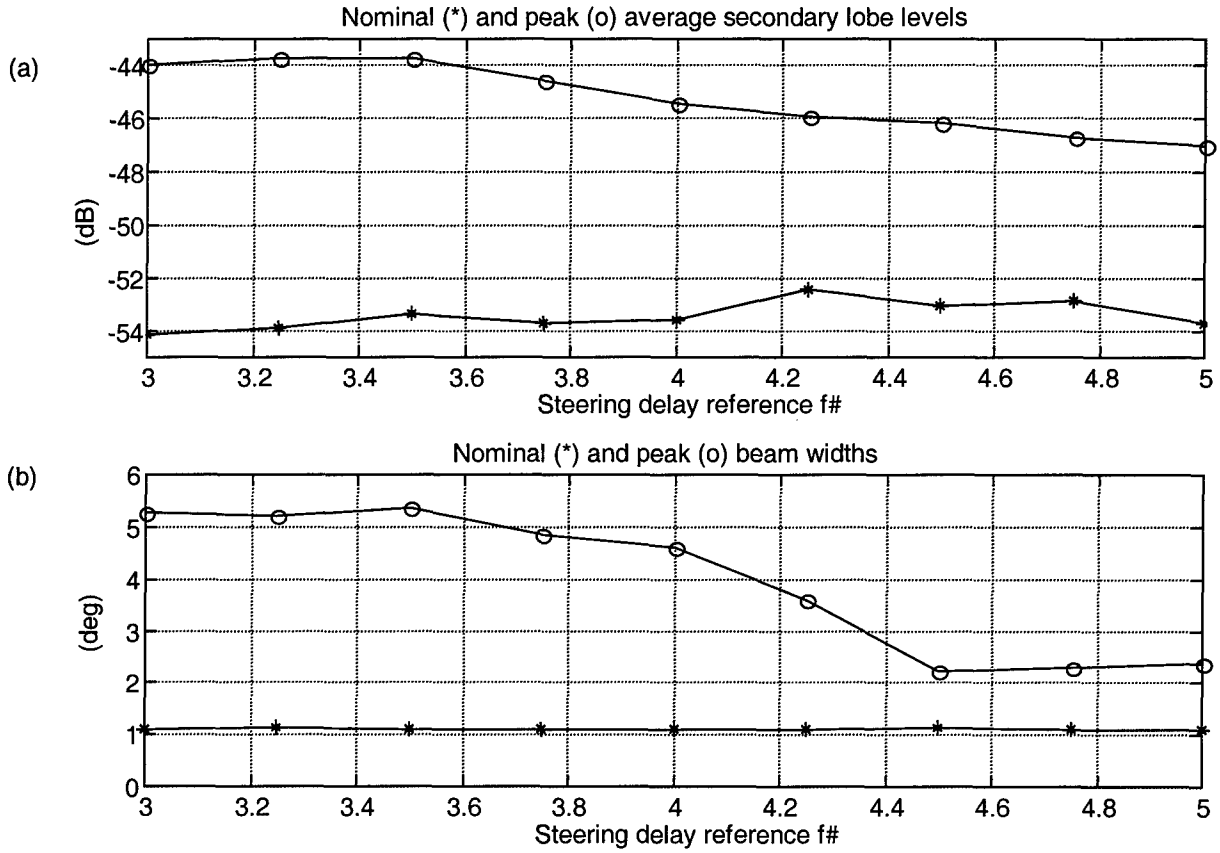


Fig. 11., Inerfield, et al.

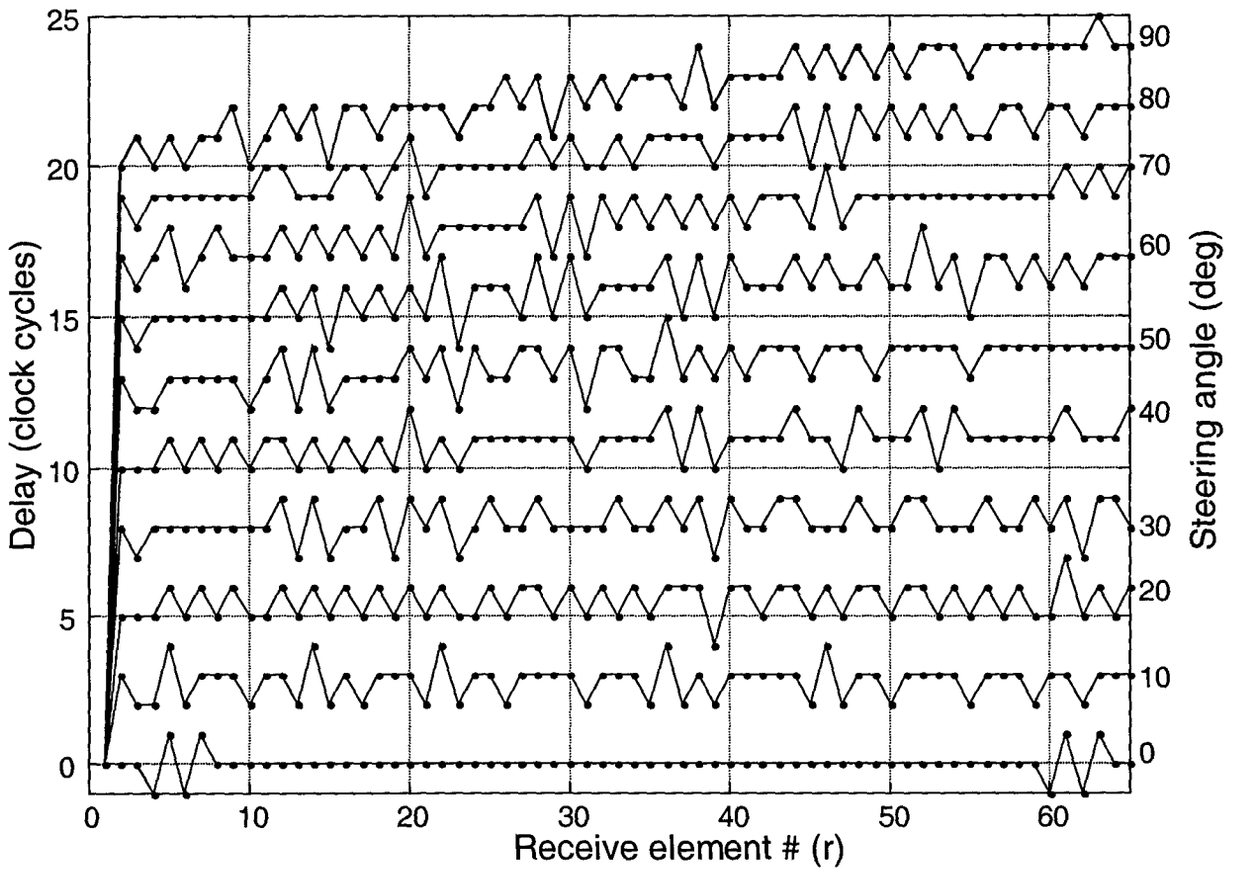


Fig. 12a., Inerfield, et al.

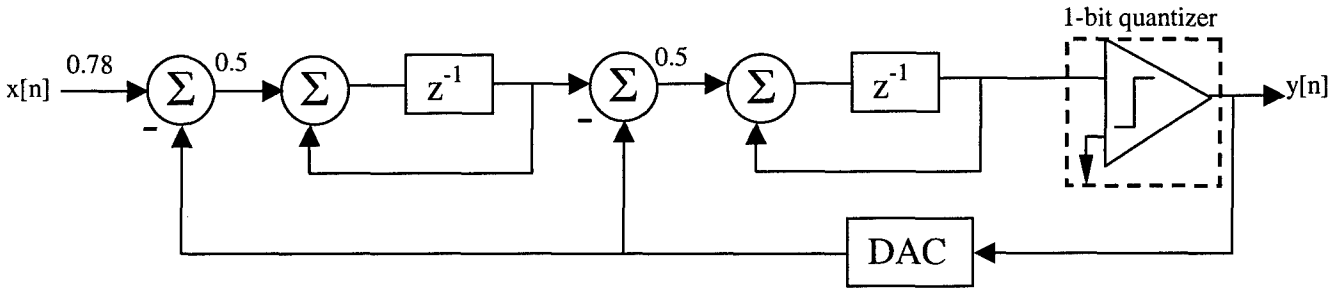


Fig. 12b., Inerfield, et al.

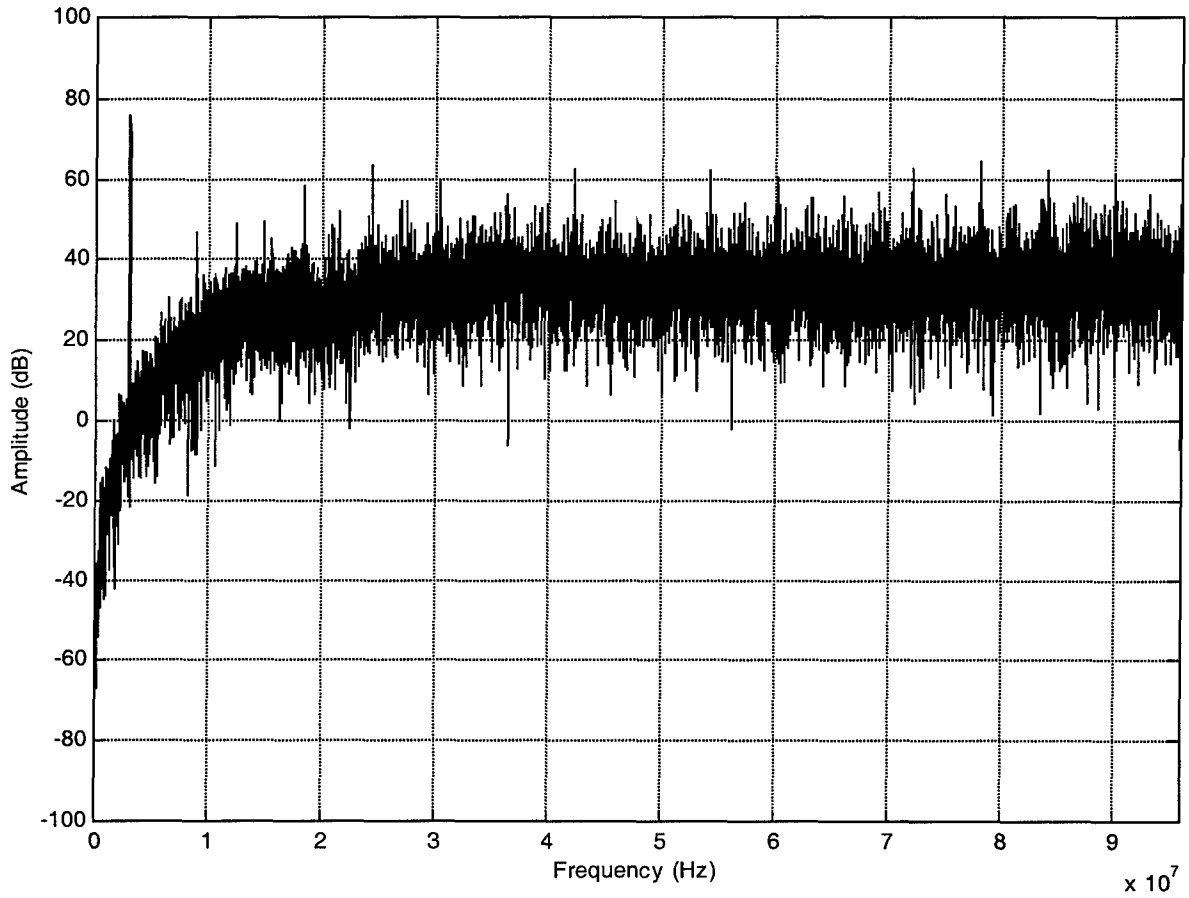


Fig. 13. Inerfield, et al.

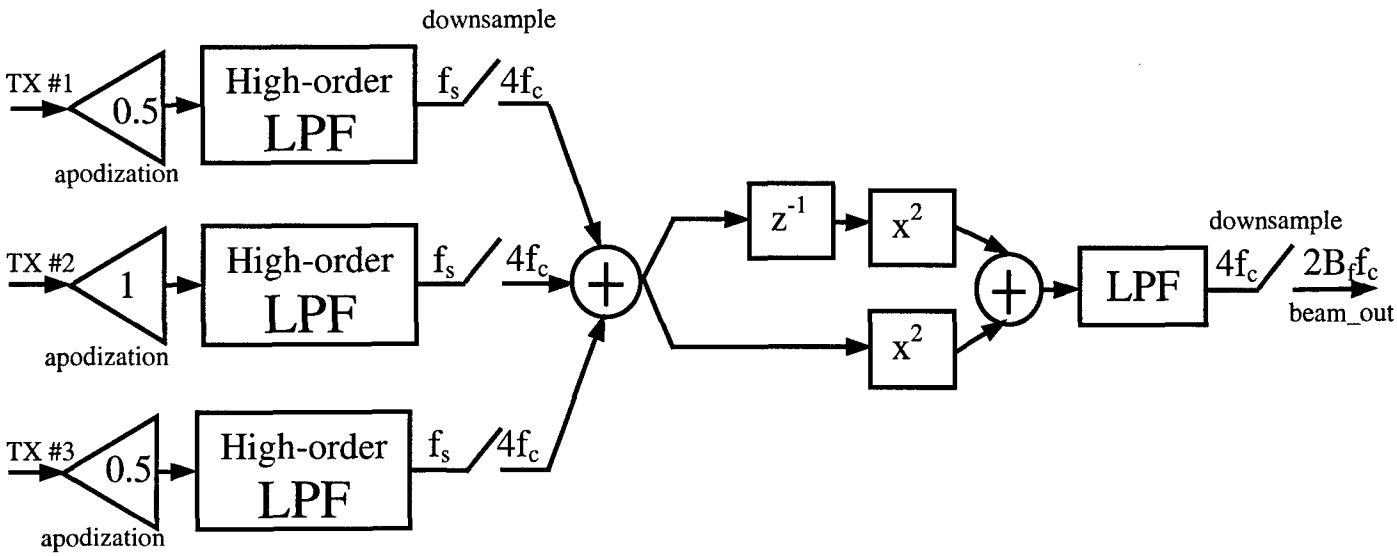


Fig. 14a., Inerfield, et al.

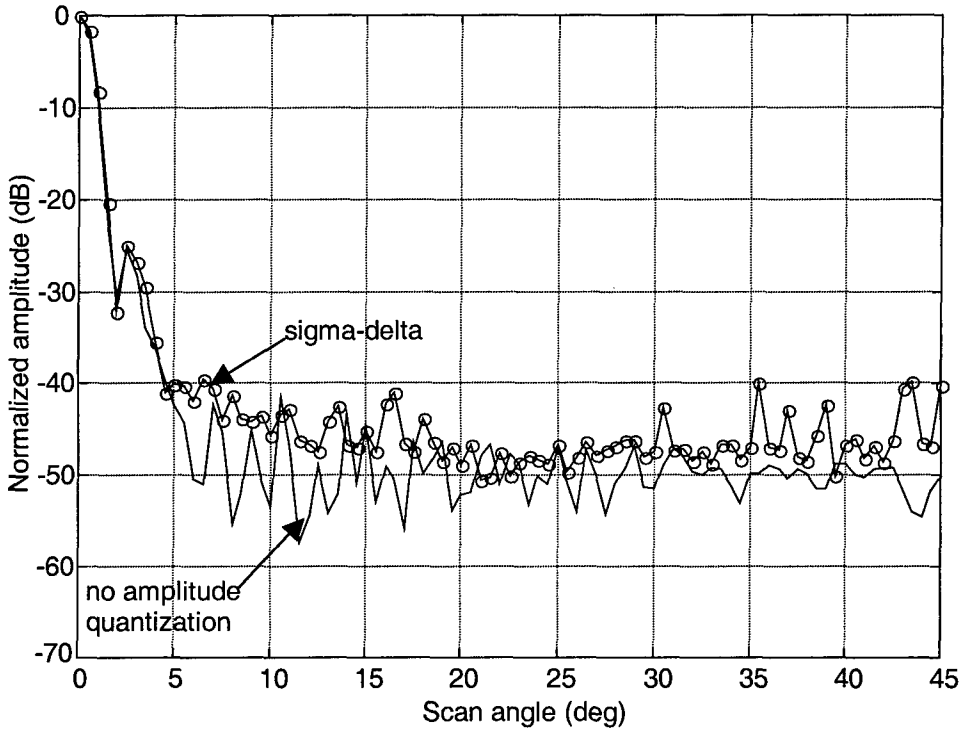


Fig. 14b., Inerfield, et al.

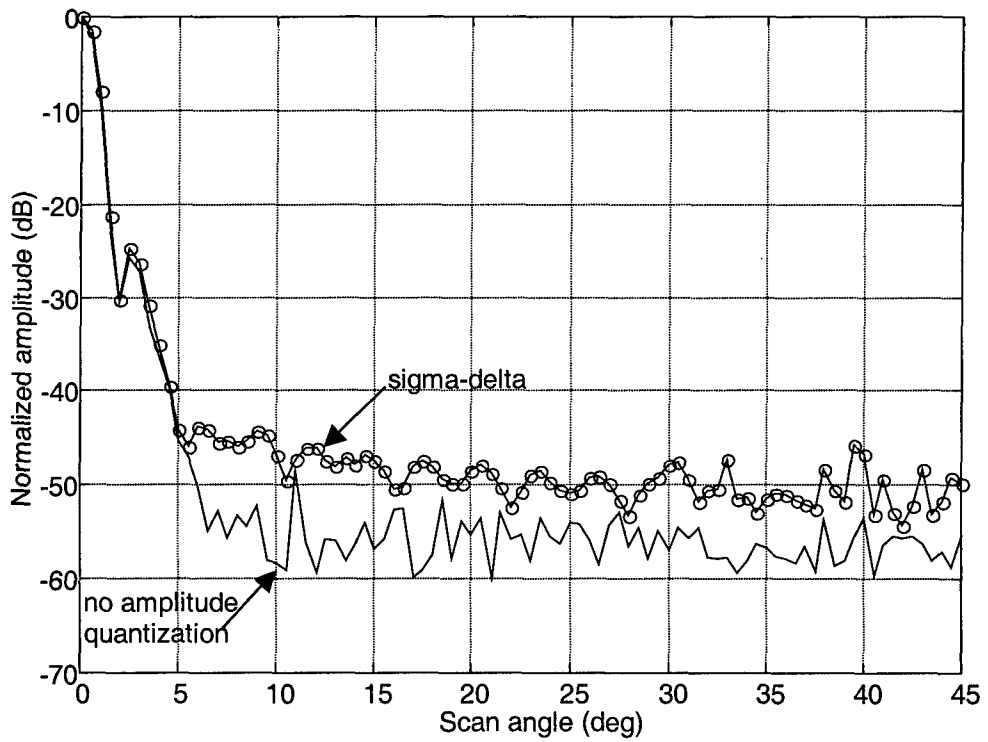


Fig. 14c., Inerfield, et al.

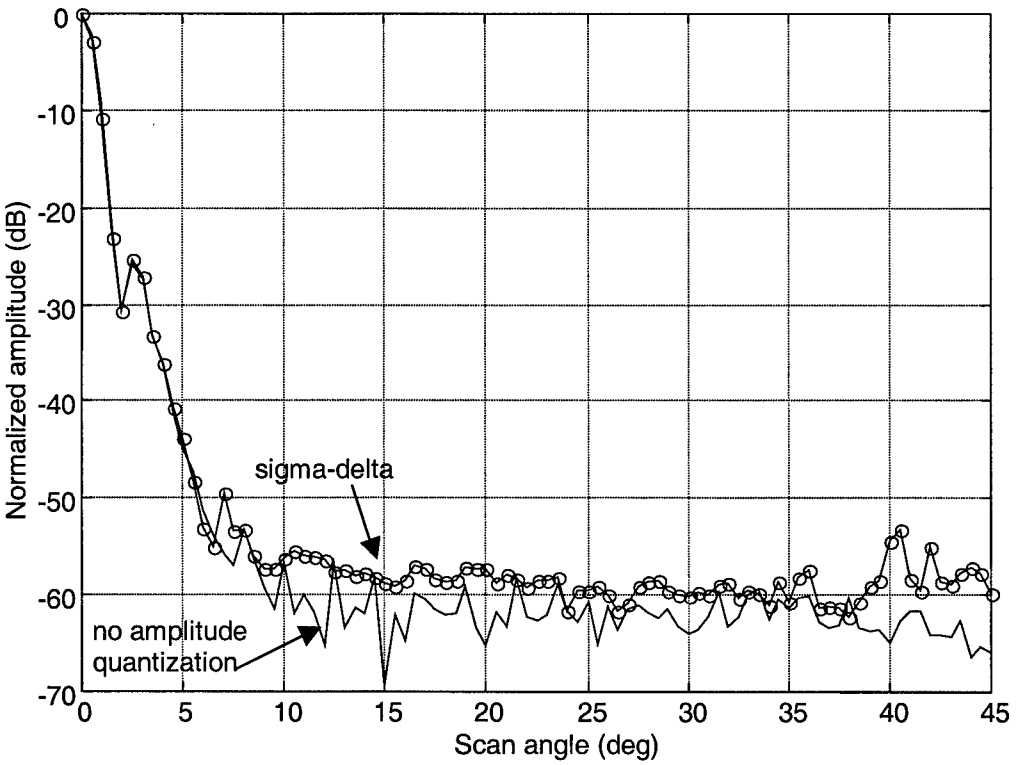

0
-10
-20
-30
-40
-50
-60
-70
Normalized amplitude (dB)
sigma-delta
no amplitude
quantization
0
5
10
15
20
25
30
35
40
45
Scan angle (deg)

Fig. 15., Inerfield, et. al.

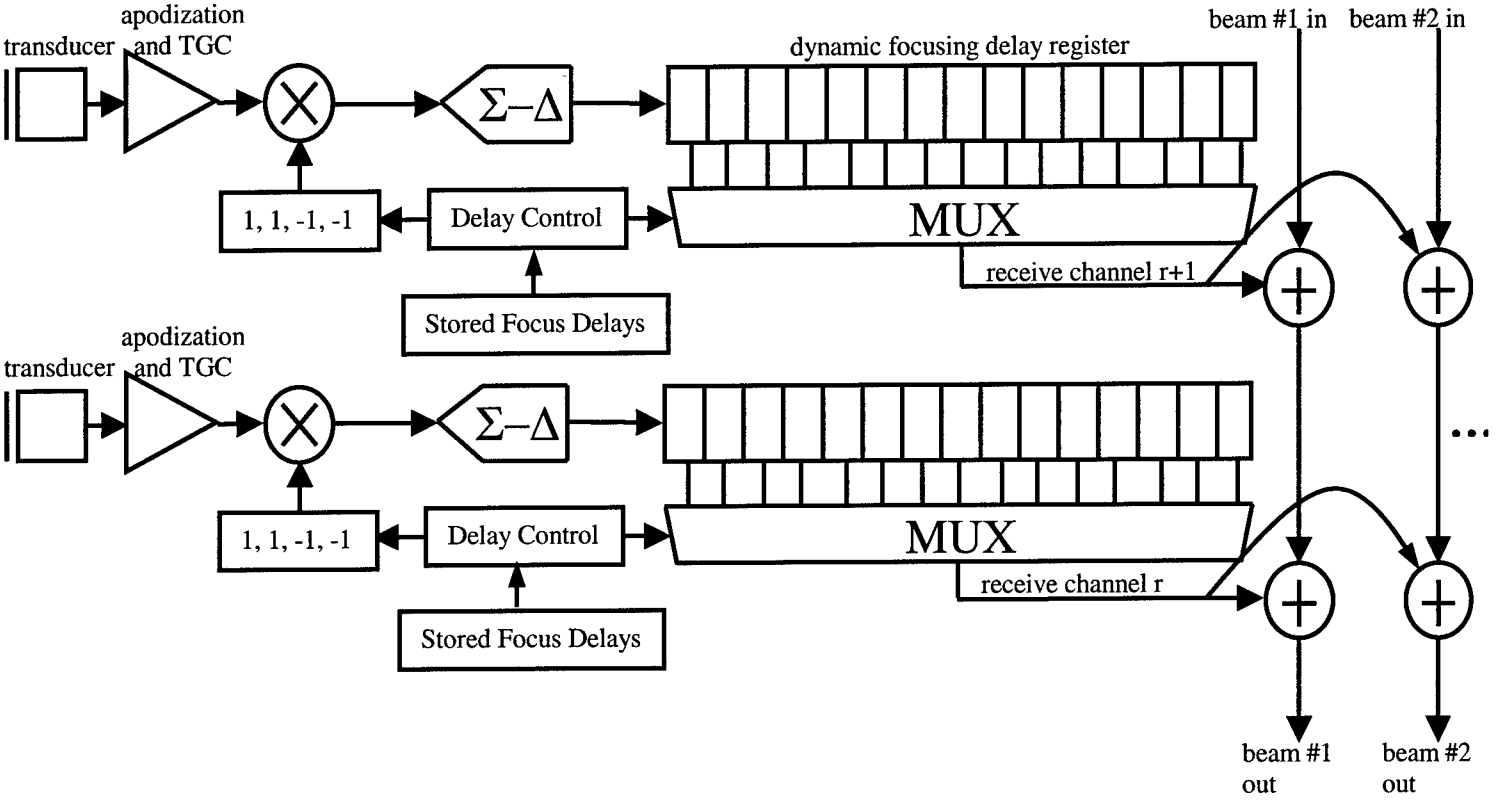
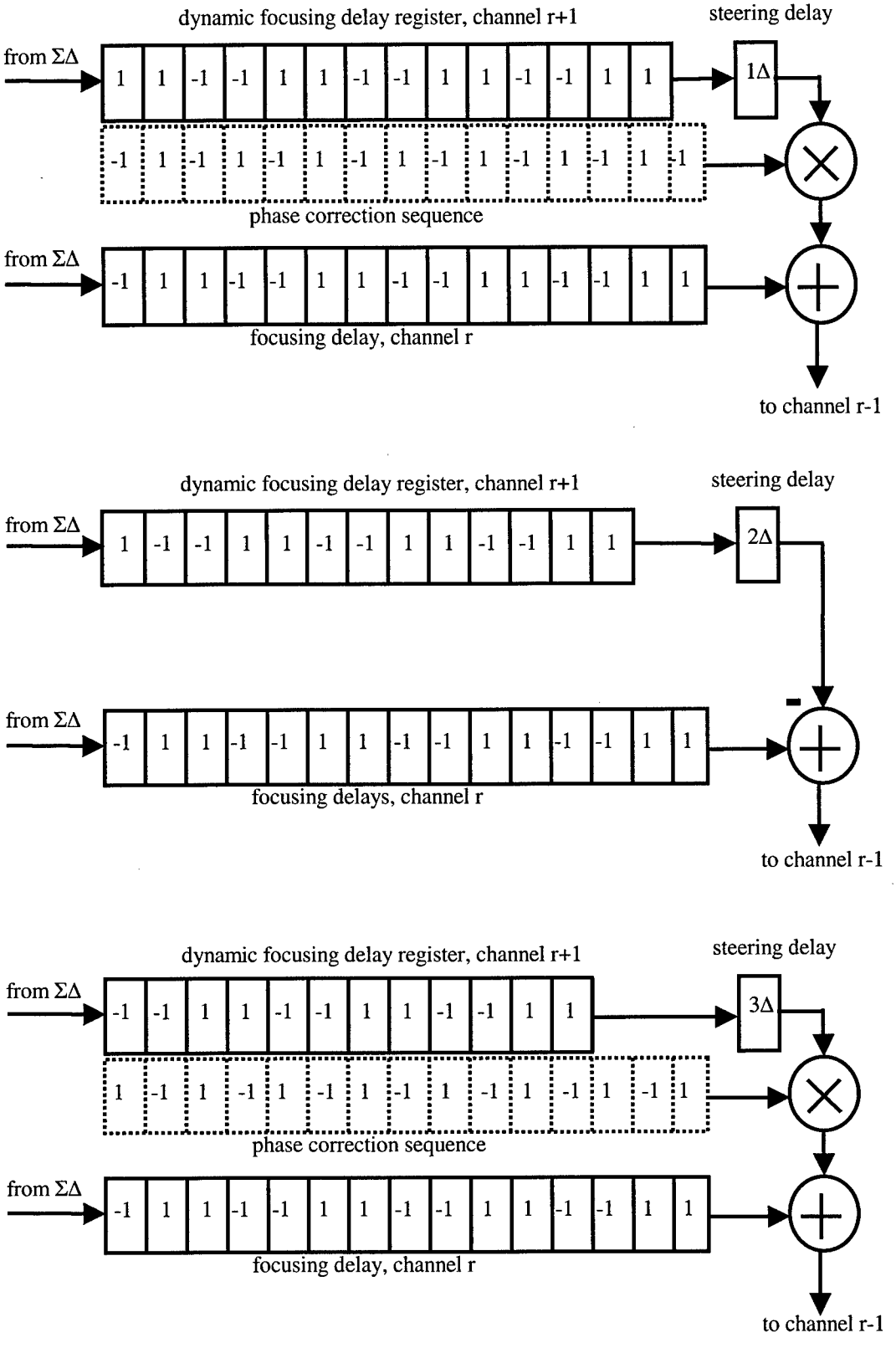


Fig. 16., Inerfield, et. al.



Appendix 2: Preprint of the publication - Christopher R. Hazard and Geoffrey R. Lockwood, "Real-time synthetic aperture beamforming: practical issues for hardware implementation", *Ultrasonics Symposium Proceedings*, 2001.

REAL-TIME SYNTHETIC APERTURE BEAMFORMING: PRACTICAL ISSUES FOR HARDWARE IMPLEMENTATION

C. R. Hazard[1,2], G. R. Lockwood[1,3]

[1] Department of Biomedical Engineering, The Cleveland Clinic Foundation, Cleveland, Ohio.
[2] Biomedical Engineering Program, The Ohio State University, Columbus, Ohio.
[3] Department of Physics, Queens University, Kingston, Ontario, Canada.

Abstract - We have implemented a synthetic aperture beamformer for real-time 3D imaging using a network of digital signal processors. This system is capable of beamforming 6.6 million points per second. Using simulated acoustical inputs, the point response of the beamformer has been evaluated using the actual hardware. The output of the beamformer using pre-calculated inputs is in agreement with simulations. The dynamic range of the system as a function of noise at the inputs has been deterimined. The experimentally measured root-mean-square noise level of the A/Ds is less than 1 mV, which provides over 65 dB of noise-free dynamic range in the image. Calibrations to remove DC offsets, to correct for gain differences, and to correct for relative delays between channels have been developed. The relative delays can be calibrated to within less than 1/1000th of a cycle. A real-time DC correction has been implemented, which corrects for the DC offsets introduced by the truncation error associated with the 16-bit fixed-point summing. The preliminary performance of the beamformer has been characterized using a limited number of narrowband inputs for beamforming. Images made at 650 frames per second using sine wave inputs are in good agreement with simulations.

I. INTRODUCTION

We have developed a synthetic aperture beamformer capable of generating 6.6 million beamformed points per second using linear interpolation to increase delay accuracy. The beamformer is part of a prototype real-time 3D ultrasound imaging system. Previous modeling addressed the effects of sampling rate and amplitude quantization on the linear interpolation algorithm. It was shown that 10-bit A/Ds sampling at 40 MHz suppress the secondary lobes by more than 50 dB [1]. The effects of motion on such a beamformer were also modeled and experimentally verified to be negligible at physiologically relevant velocities [2]. The initial hardware and software implementation of the system using the TMS320C6201 DSP (Texas Instruments, Dallas, TX) was previously presented [3]. Here the practical matters associated with the actual implementation of the system are presented. These include calibrating the system, removing DC artifacts introduced by the fixed point processing, and assessing the effects of noise at the front end of the system.

II. INPUT NOISE AND DYNAMIC RANGE

To provide performance criteria for the front-end electronics, the noise response of the beamformer was simulated using Gaussian white noise as the input to each channel of the beamformer. The root-mean-square (rms) value of the noise was varied from 0.001 V to 0.1 V and the resulting dynamic range in the I and Q sums was calculated for each noise level. The input range of the A/Ds is +/- 1 V. Figure 1 shows the simulation results. The calculations were repeated three times for each rms noise level and the error bars on the graph show the range for the three trials. An rms noise input of 1 mV results in a dynamic range greater than 68 dB. The rms noise introduced by the A/Ds and input buffers for the prototype system were determined by digitizing zero inputs on each channel of the system. The maximum rms noise level across all the channels was less than 0.75 mV. This provides design specifications for the front-end amplifiers and the noise performance of the transducer.

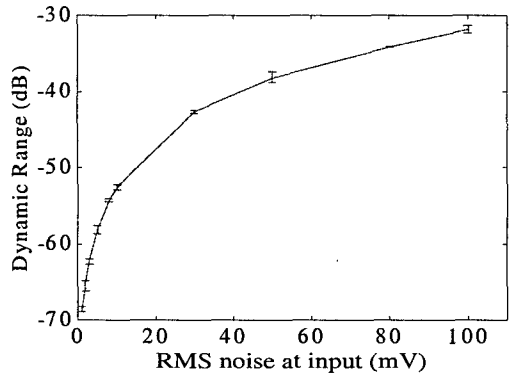


Figure 1: Dynamic range vs. rms noise

III. CALIBRATIONS

Three calibrations are used in the prototype beamforming system: DC offset calibration, relative delay calibration, and gain calibration. The beamforming software is capable of subtracting a DC offset from the digitized inputs before beamforming. The DC offset, for each channel, is determined by digitizing zero inputs and calculating the mean. The DC offsets are calculated offline and loaded into the beamformer at initialization. Subtracting this value reduces the DC errors introduced by the front-end electronics or A/D.

A synchronization amplifier is used to allow the multiple channels of the system to collect data simultaneously. The clock and synchronization signals are distributed to each of the A/Ds using separate ribbon cables. Small relative delay differences occur from channel to channel due to different cable lengths and delays introduced by buffers in the synchronization distribution amplifier. These small relative delays are fixed and can therefore be measured. To determine the relative delays, a sine wave generated by a function generator, is split between two matched cables. The pairs of signals are then digitized by a reference channel and each of the other channels in the system pairwise. Using cross correlation and interpolation the relative delays are calculated to within $1/1000^{th}$ of a cycle. Once the delays are determined, the addresses and coefficients used for beamforming are adjusted to compensate for the delays.

The gain for each of the 64 channels is slightly different. Using the same set of signals collected to determine the relative delay errors, the relative gain for each channel can be calculated. Correcting the gain differences is simply a matter of adjusting the linear interpolation coefficients to correct for the different gain values. This is a simple modification of the apodization functions applied.

IV. PRELIMINARY PERFORMANCE

The performance of the beamformer can be evaluated independently of the transducer and other system components. The current system uses three transmit events and 64 receive channels to generate images. The theoretical maximum frame rate is determined by the imaging depth and by the number of transmits. For an imaging depth of 15 cm, the round trip time of flight is 200 microseconds. The minimum time for three transmit events would be 600 microseconds, making the theoretical maximum frame rate greater than 1666 frames per second. In practice, the time between transmits must be increased to allow long-range echoes to subside. If the time between transmits is 300 microseconds, the maximum frame rate is closer to 1000 frames per second.

The current beamformer is capable of generating 6.6 million points per second. This limitation is imposed by the bandwidth available for interprocessor communication. This corresponds to a frame rate of approximately 650 2D frames per second for a 10,000 pixel image or 1000 frames per second for a 6000 pixel image. Since the current hardware limits how many points can be calculated, it is important to optimize the sampling of the image plane. The beamformed points can be positioned anywhere in the image, and the pixels in a rectangular image could be directly calculated. However, for a square image the maximum size would be approximately 100 x 100 pixels. Because the lateral beamwidth increases as the distance from transducer increases, polar sampling is more efficient for large image sizes. Polar sampling exploits the loss of lateral resolution with distance, by using fewer samples for points further away from the transducer.

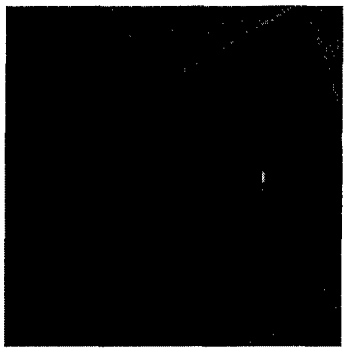


Figure 2: Image of point target using hardware and simulated acoustical inputs.

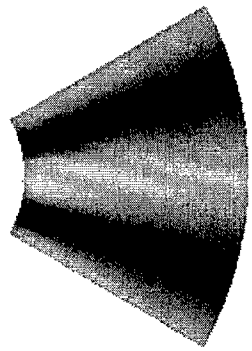


Figure 3: Image using two sine wave inputs without a DC offset correction.

To verify that the hardware implementation agreed with the simulation of the system, the hardware was used to beamform simulated acoustical inputs. The image produced by the beamforming hardware was compared to a simulated image using the same simulated acoustical inputs. Only a slight modification to the hardware system was required for this test. The DMA channel, which transfers data from the A/D to the internal SRAM buffer, was re-programmed to transfer data from an external SB-SRAM buffer, which contained the simulated acoustical inputs, instead of the A/D. No other modifications were made. Figure 2 shows the resulting image of a point target at F/8 using the simulated inputs and the actual hardware. The images from the hardware and simulator are identical. This demonstrates that the hardware is performing as expected.

As a simple test of the beamformer, an image was made using two sine-wave inputs. To simplify the resulting images, only the central transmit was used to beamform the image. The sine waves were generated using a function generator (Hewlett Packard 8116A). The signal from the function generator was split between two channels using a 50-ohm splitter (Mini-Circuits 15542). For this test, only 60 channels of the beamformer were used and the signals from the function generator were connected to the 31st and 34th channel of the beamformer.

Figure 3 is the image formed using the two sine wave inputs. The frequency of the sine wave was 3.5 MHz, the designed center frequency of the system, and the amplitude was approximately 0.5 V. The dynamic range of the resulting image has been adjusted to normalize the maximum value in the image to the maximum display value. Two dark bands, which represent the complete destructive interference between the sine waves on the two channels, are seen in the image. These nulls occur for angles that can be calculated based on the geometry of the system. The beamformer is designed for an array with elements spaced at 0.57 wavelengths. The two channels used for the test image are the 31st and 34th channels. This gives a separation between the channels of 3*0.57 wavelengths. The angle, which corresponds to destructive interference at this spacing, is approximately 17 degrees. This agrees well with the angle in the resulting image. Simulations of the same input show similar results.

V. Real-time DC Correction

The narrow-band test image shown in Figure 3 has a periodic variation in the amplitude of the image. This is perhaps most clearly seen as a jagged edge along the dark bands in the image. The ripple is the result of a DC offset in the I and Q sum values. It can be shown theoretically that a DC offset creates a periodic variation in the image by examining the algorithm used to calculated the envelope of the signal in the prototype system. The envelope of the signal is estimated using second order sampling [4]. The estimate of the envelope, A, is given by Equation (1).

$$A \approx \sqrt{I^2 + Q^2} \qquad (1)$$

where I is the delayed signal value and Q is a value calculated with an additional quarter period delay. In the narrow-band approximation used in this system, the I and Q values are given by equation (2).

$$I = A(t)\cos(\omega\, t+\phi) \qquad (2)$$
$$Q \approx A(t)\sin(\omega\, t+\phi)$$

where ϕ is a fixed arbitrary phase factor.

Equation (3) shows the I and Q values with a DC offset, d.

$$I = A(t)\cos(\omega t+\phi)+d \qquad (3)$$
$$Q \approx A(t)\sin(\omega t+\phi)+d$$

The resulting envelope estimated using the DC contaminated I and Q values is given in Equation (4).

$$\sqrt{I^2+Q^2} = \qquad (4)$$
$$A\sqrt{1+2\sqrt{2}(d/A)\sin(\omega_o t+\phi_2)+2(d/A)^2}$$

Where $\phi_2 = \phi + \pi/4$. The DC offset introduces two error terms in Equation (4), including the term which is seen as a ripple in the image.

The DC offset is introduced by the truncation of the linearly interpolated delayed values to allow summing over all the channels using a 16-bit buffer. In a twos complement binary numbering system, the truncation of a number results in either the same value or a smaller or more negative number. As a result, the sum of many truncated linearly interpolated values will have a net negative truncation error. For any particular interpolated value, the truncation error cannot be determined a priori, but the net truncation error for the sum over all the channels will be an average error that can be estimated. On average, the truncation error on each channel will be negative with a magnitude of one half of the least significant un-truncated bit. The truncation error for the sum of N channels will be N/2. Simulations of the system have shown this to be the case.

A real-time DC correction was added to correct this truncation induced offset. The mean value of the I and Q sums is calculated and subtracted before the squaring and square root operations. Figure 4 shows the resulting image using this correction. No ripple is present.

VI. Conclusions

A prototype beamformer has been built using a high speed network of off-the-shelf digital signal processing hardware. The beamformer uses a synthetic aperture algorithm and linear interpolation to increase delay accuracy resulting in improved contrast resolution. The system is capable of producing 6.6 million beamformed points per second. The hardware algorithms have been verified using simulated acoustical inputs. The system must be calibrated to correct for DC offsets, gain differences, and relative delays between channels. Images made using sine wave inputs match simulations. Real-time correction of DC offsets has been implemented to overcome truncation induced offsets. Future work will involve evaluating this prototype beamformer using real acoustical signals from a rocking transducer array.

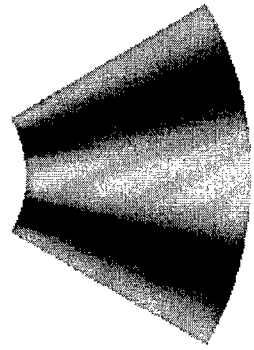


Figure 4: Image using two sine wave inputs with DC correction.

VII. References

[1] C.R. Hazard, G.R. Lockwood, "Theoretical assessment of a synthetic aperture beamformer for real-time 3D imaging," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, vol. 46(4), pp. 972-980, 1999.

[2] C.R. Hazard, G.R. Lockwood, "Effects of motion on a synthetic aperture beamformer for real-time 3D ultrasound," *Proc. 1999 IEEE Ultrason. Symp.*, vol. 2, pp. 1221-1224, 1999.

[3] C.R. Hazard, G.R. Lockwood, "Developing a high speed beamformer using the TMS320C6201 digital signal processor," *Proc. 2000 IEEE Ultrason. Symp.*, vol. 2, pp. 1755-1758, 2000.

[4] T.K. Song, S.B. Park, "A new digital phased array system for dynamic focusing and steering with reduced sampling rate," *Ultrasonic Imaging*, vol. 12, pp. 1-16, 1990.

Appendix 3: Progress report from Tetrad Corporation.

To:	Geoff Lockwood
Project:	Cleveland Clinic 3-D Mechanically Steered Phased Array Probe
Customer Grant #:	DAMD17-99-1-9034
Reports:	Delivery of probe to Cleveland Clinic; Final report
Milestones:	#s 19 and 20
Amount:	$42k and $20k
Date:	November 27, 2001
Written By:	Mike Zipparo
Organization:	Tetrad Corporation
Address:	357 Inverness Dr. South, Suite A
	Englewood, CO 80112
Phone:	303-754-2309
FAX:	303-754-2329
E-mail:	mzipparo@tetradcorp.com

Milestone #19 – Delivery of probe to Cleveland Clinic.

An array has been incorporated into a completely new wobbler mechanism and housing. The array used is a different one than what was reported on in Milestone 17. Concerns about the flex circuit arose, particularly its ability to withstand prolonged flexing under wobbling. Rather than risk losing elements after the probe was completed, a complete new array was constructed from scratch. The complete probe, along with a completely new driver, has been shipped to the Cleveland Clinic. The wobbler was tested extensively both before and after the probe was placed into it. All of the elements were functional after the probe housing was closed, as was the rotation mechanism. A complete set of test data was sent with the probe. A summary of that will be presented here.

The completed probe was measured at GTS. Figure 1 shows measurements of received voltage (Vpp), pulse width (PW), center frequency (fo), and fractional bandwidth (FBW) for the final array mounted in the wobbler mechanism. The average center frequency was about 3.4 MHz. The bandwidth measured through the acoustic window averaged about 60%. At Tetrad, the array averaged almost 75% bandwidth. It is unclear whether the discrepancy is due to the acoustic window or to some difference in the test conditions such as the pulser and receiver used.

Figure 2 shows the element uniformity measured for the array, which is very good for all of the measurements. All but four of the elements are within ± 0.5 dB of the mean sensitivity, and those four elements are within ± 1.0 dB. The mean bandwidth varied between 45 and 68%.

Milestone #20 – Final Report.

Reports for all milestones have been submitted and the final probe, which incorporated a 3-layer ceramic stack into a rotation mechanism, has been delivered to The Cleveland Clinic, care of Geoff Lockwood at Queens University. This fulfills our requirements under this contract.

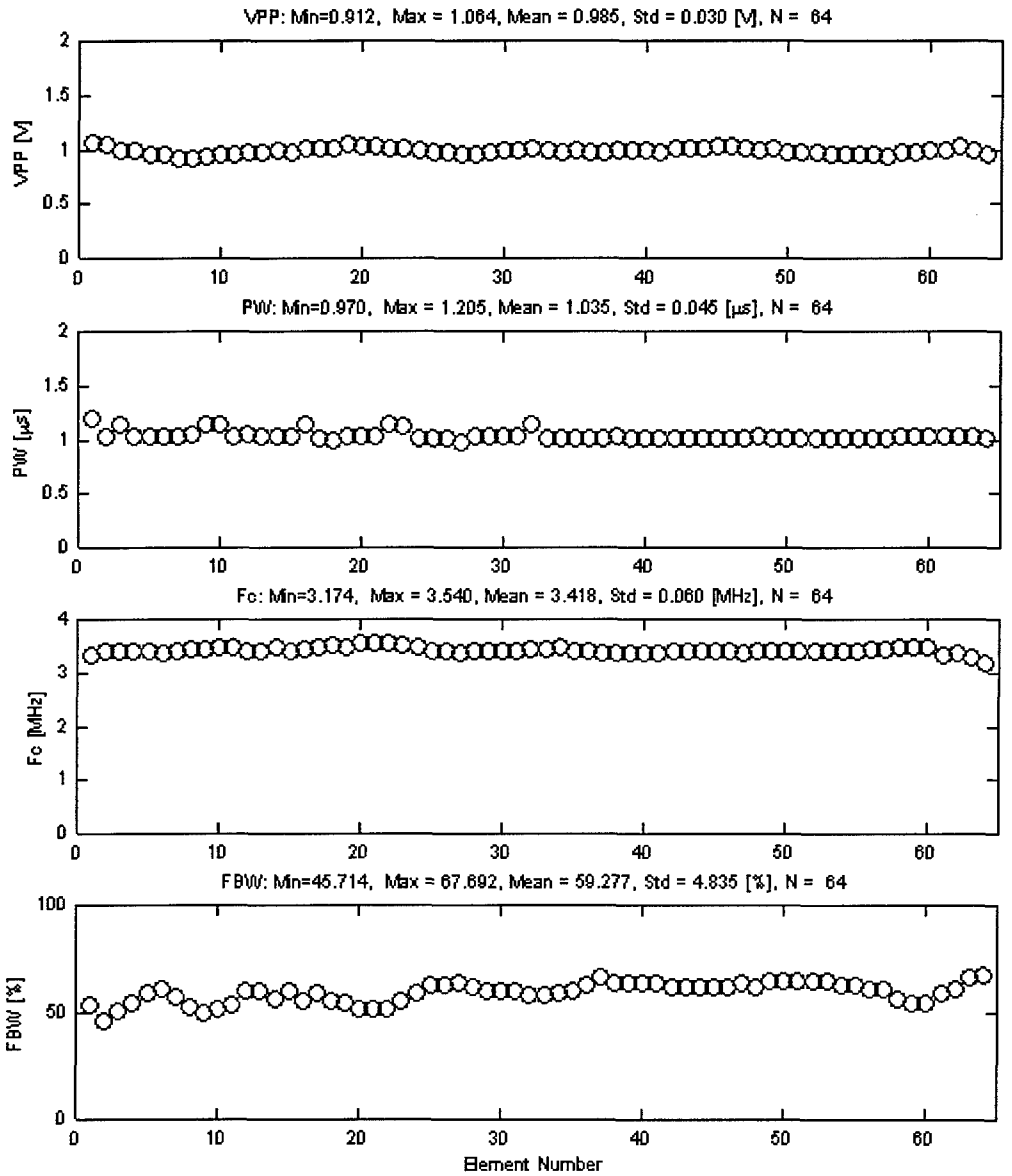


Figure 1) Measured transmit-receive response for the final array mounted in the wobbler.

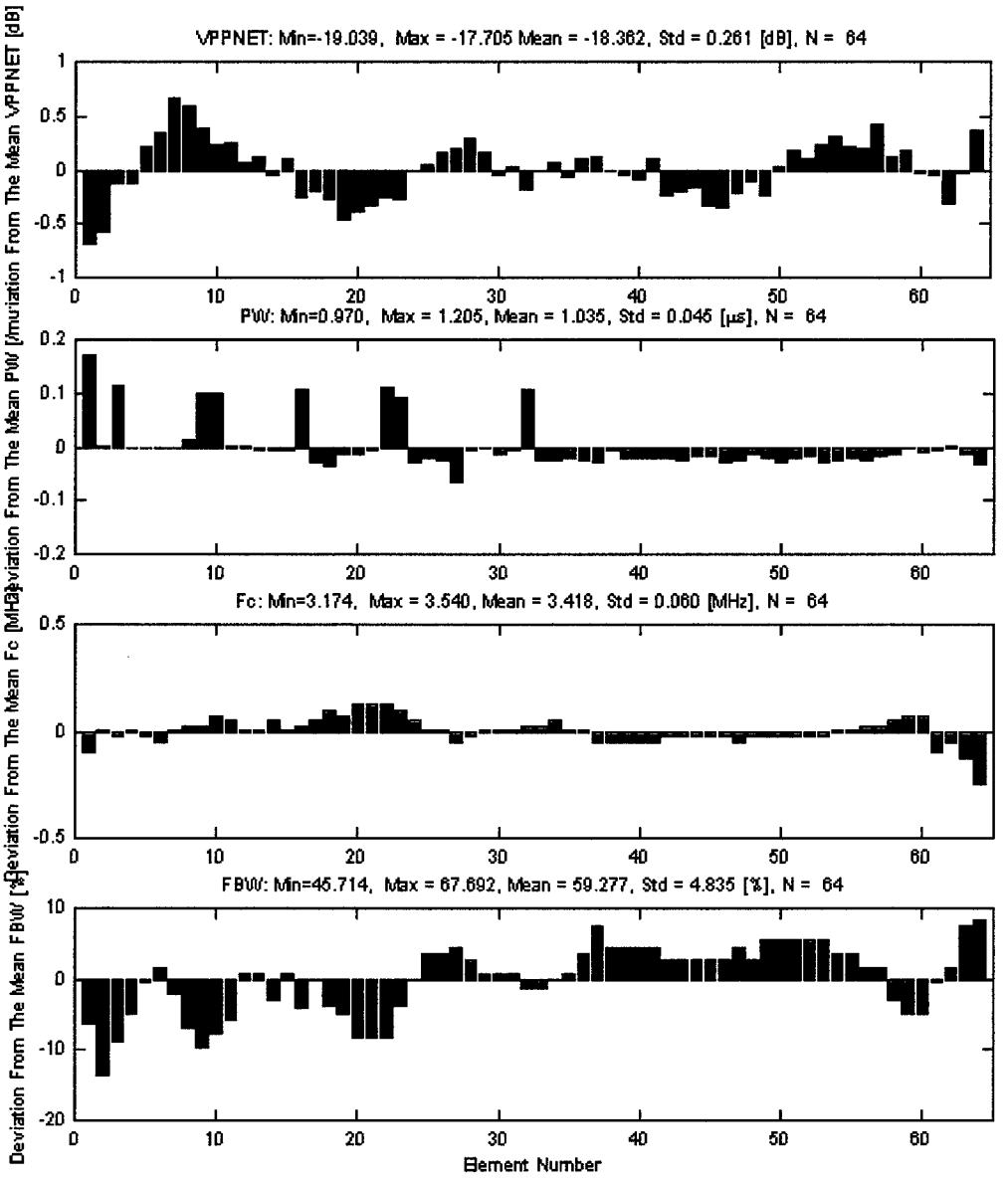


Figure 2) Element to element variation in the measured response.

Appendix 4: Reprint of the publication - Raj Shekhar and Vladimir Zagrodsky "Interactive visualization of four-dimensional ultrasound data," presented at *IEEE Visualization* conference, 2001.

Interactive Visualization Of Four-Dimensional Ultrasound Data

Raj Shekhar and Vladimir Zagrodsky

The Cleveland Clinic Foundation, Cleveland, Ohio

1 INTRODUCTION

Acquisition of four-dimensional (4D) ultrasound, especially of the heart, is gaining popularity. 4D acquisition is powerful in that it reveals the complex three-dimensional (3D) geometry and motion of the heart. Historically, 4D images have been assembled with planar images taken at multiple locations over multiple heart cycles. 3D localizers have been used to record the orientation of the planes and electrocardiograms have been employed to tag each planar image with its correct phase in the heart cycle. This 4D acquisition protocol is slow and, despite great care, one cannot avoid distortion in the final 4D image due to patient motion and inherent inaccuracies in 3D localization mechanisms. Real-time 3D acquisition capability, a faster, more accurate and more convenient alternative, was recently introduced and is an area of active research and development. Real-time 3D acquisition is arguably the future of 4D ultrasound.

Irrespective of the acquisition mechanism, challenges for 4D visualization remain the same. Methods to visualize 4D data must handle large data size (100-300 MB) and maintain a frame rate (20-30 Hz) so as not to alter the underlying heart motion. Important diagnostic cues are derived from the heart motion; maintaining the original heart motion is, therefore, critical to making an accurate diagnosis. Additionally, many applications require visualizing two 4D images, either side-by-side or overlaid, simultaneously, thus doubling the data size requirement.

The use of 3D texture mapping hardware for accelerating volume rendering is well reported in the literature [1-3]. Many volume rendering libraries, such as Volumizer by Silicon Graphics, have been built upon this technology. The work reported so far has focused on mostly on 3D data. We report here the use of 3D texture mapping hardware as a means to achieve the desired frame rate in 4D visualization. The rendering speed of 3D texture mapping hardware is near instantaneous as long as the image data fit in the accelerated texture memory. The above condition is, however, seldom met when using 4D data. We describe here our novel use of data subdivision and caching schemes to meet challenges unique to 4D visualization. As for terminology, the term "texture memory" will imply "accelerated 3D texture memory" in the remainder of this article.

2 DATA SUBDIVISION

A 4D image (>100 MB) is typically larger than the size of the texture memory (32-64 MB) available even on most high-end graphics boards. Therefore, to still be able to use the limited texture memory, we "brick" the 4D data by dividing up each 3D image (or frame) of the sequence into 3D subblocks of identical size. An array of 3D texture objects equal in size to a 3D subblock are created in the texture memory. The required data subblocks are copied to the existing texture objects before and/or during rendering.

Address: 9500 Euclid Ave, ND20, Cleveland, OH 44195
Email: {shekhar,zagrodv}@bme.ri.ccf.org

Not all data contribute to a visualization task. Data subdivision, hence, provides the "granularity" to pick the subblocks that contribute to visualization and discard the ones that do not, thus reducing data requirement. The discarded subblocks are the subblocks either away from the cutting plane in multi-planar reformatting (MPR) (Fig.1) or containing transparent voxels in volume rendering. The size of the subblocks plays a key role in determining the amount of reduced data. In general, smaller-sized subblocks lead to more specific data selection, but they also cause a greater percentage of data duplication because neighboring subblocks must overlap by a single layer of voxels. Computation costs also rise with decreasing subblock size. We are studying the effect of subblock size on data reduction and overall performance to determine the optimal subblock size.

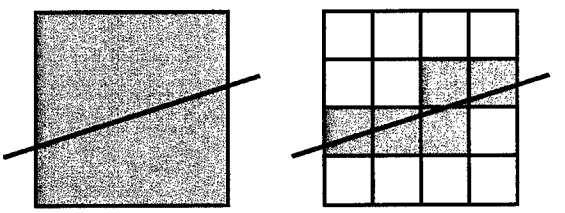


Fig.1 A schematic to show the concept of data subdivision. The oblique line shows the plane of MPR. Without data subdivision, the entire data must be copied to the texture memory; only the shaded subblocks need to reside in the texture memory following data subdivision.

3 CACHING

The data required for visualization, even after data reduction, may still exceed the available texture memory. In such situations, some data subblocks must be brought into the texture memory, overwriting preexisting ones during rendering. In this sense, the texture memory also functions like a cache if data overflow occurs. Initialization of this cache and the associated cache replacement rules influence the frame rate that can be achieved. There are two features that make the current caching task unique. First of all, unlike most caching applications, the cache items (data subblocks) here are identical in size. Furthermore, due to the periodic (looping) nature of 4D visualization, cache items have identical lag time between their successive usages, i.e., subblocks to render the first frame will be required as frequently as those to render the second frame. The familiar Least-Recently Used and Least-Frequently Used cache replacement rules, which assume disproportionate usage, therefore offer the worst performance in our case. Our caching solutions are explained below.

We first explain caching for the steady-state or no-interaction case. We treat all but one texture object as residing in the "long-term" cache, whereas the remaining texture object is used for short-term caching. Given a visualization task and its viewing

parameters, the subblocks needed per frame for all the frames are determined. The long-term cache is then populated with as many selected subblocks as possible. During rendering, a subblock is first searched in the long-term cache; if unavailable, it is brought into the short-term cache to complete the rendering. We use the term "cache initialization" to refer to the initialization of the long-term cache and "leftover" subblocks to refer to subblocks that are left out of the long-term cache. Cache initialization ensures that the number of leftover subblocks per frame is roughly equal and hence the time spent in short-term caching is equally distributed between frames. If a certain subblock is used more than once per frame, it is given priority in the long-term cache.

As long as the viewing parameters are unchanged, the contents of the long-term cache stays intact. However, the subblock requirement changes once interaction begins. Although many existing subblocks continue to be needed, a subset of new subblocks emerges and also a subset of existing subblocks is no longer needed. Since we know how to initialize the long-term cache optimally for any given set of viewing parameters, it is possible to update the long-term cache for each intermediate orientation during interaction. However, such an attempt may be unnecessary because intermediate orientations will likely be temporary during interaction and may not last more than a few frames. An attempt to update the long-term cache for all frames during an interactive session is also damaging to achieving interactive frame rate. We, therefore, constantly compare the contents of the long-term cache with the continuously changing subblock requirement and update the long-term cache entries such that they are optimal for rendering only a few upcoming frames in the current orientation but not the entire sequence. These incremental changes are made in such a way that when the interaction stops, the contents of the long-term cache are restored optimally upon one loop through the sequence.

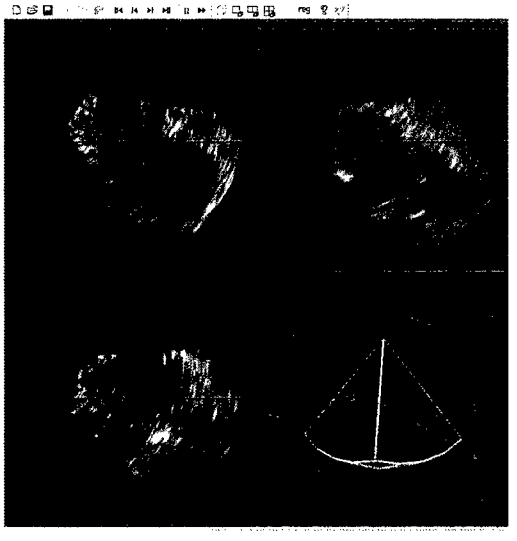


Fig. 2 A snapshot of an interactive cine MPR session.

4 RESULTS

A snapshot of an interactive cine MPR session is shown in Fig. 2 ("cine" refers to the looping feature). The bottom-right viewport is a reference schematic showing the orientation of the three reformatted planes inside the ultrasound acquisition pyramid and its bounding box. A movie clip is available at the URL http://www.lerner.ccf.org/bme/staff/shekhar/research/cinempr.

The 4D data set used for the cine MPR demonstration was a sequence of twenty 128 x 128 x 512 3D images. The data size was thus 160 MB. The maximum frame rate achieved in conjunction with 3DLabs Wildcat 4210 graphics accelerator board with effectively 64 MB of texture memory for byte data was 28 Hz when all three planes were animated simultaneously. The maximum frame rate achieved was slightly higher than the desired frame rate of 25 Hz (equal to the acquisition frame rate). The subblock size was 16 x 16 x 32 and each viewport was sized 360 x 360 pixels.

In addition to interactive cine MPR, we are investigating interactive volume rendering of 4D cardiac data and visualization of two instantaneously fused 4D data sets following spatio-temporal image registration.

5 DISCUSSION AND CONCLUSIONS

4D visualization algorithms need to be developed as 4D data acquisition techniques emerge and gain clinical acceptance. 4D visualization is unique in that even in the steady-state, when no user interaction takes place, the rendered views must be animated at the original acquisition frame rate to prevent any motion distortion. The data subdivision and caching schemes presented here provide a general framework to visualize 4D data using the 3D texture mapping hardware. We have used this framework to perform cine MPR of three perpendicular cross-sections through a 4D cardiac data set, as well as other tasks. We have further shown that we can achieve the necessary frame rate in cine MPR visualization. The visualization framework we present offers many advantages: (1) hardware acceleration, (2) handling of one or more 4D data sets at one time, (3) usability across many visualization tasks, (4) capability to use any-sized 4D data on any-sized texture memory, and (5) achievement of higher frame rates. The maximum frame rate that can be achieved in a given visualization task depends on a number of factors, namely, the amount of data, the amount of texture memory, the data transfer rate between the system and the texture memories, and the size and number of viewports on the screen. Although the hardware acceleration of texture mapping together with our solutions maximizes the frame rate, limited resources, especially limited texture memory, may not permit achievement of the desired frame rates. Fortunately, texture memory is becoming inexpensive, which is an encouraging trend for continued development of interactive 4D visualization algorithms and techniques.

References

[1] B. Cabral, N. Cam, and J. Foran, "Accelerated volume rendering of tomographic reconstruction using texture mapping hardware," in *Workshop on Volume Visualization*, Washington, DC, pp. 91-98, 1994.

[2] A. Van Gelder and K. Kim, "Direct volume rendering with shading via three-dimensional textures," in *Proc. Symp. Vol. Visual.*, pp. 23-30, 1996.

[3] M. Bailey, "Interacting with direct volume rendering," *IEEE Computer Graphics and Applications*, vol. 21, no. 1, pp. 10-12, 2001.

Appendix 5: Preprint of the publication - Raj Shekhar and Vladimir Zagrodsky "Mutual information-based rigid and nonrigid registration of ultrasound volumes," *IEEE Transactions on Medical Imaging*.

IEEE TMI Manuscript 291-00

Mutual Information-Based Rigid and Nonrigid Registration of Ultrasound Volumes

Raj Shekhar and Vladimir Zagrodsky

Department of Biomedical Engineering, Lerner Research Institute
The Cleveland Clinic Foundation, Cleveland, Ohio 44195, USA

Corresponding Author: Raj Shekhar, Ph.D.
Department of Biomedical Engineering (ND20)
Lerner Research Institute
The Cleveland Clinic Foundation
9500 Euclid Avenue
Cleveland, Ohio 44195

Phone: 216/445-3246
Fax: 216/444-9198
Email: shekhar@bme.ri.ccf.org

This project was supported by Department of Defense research grant DAMD17-99-1-9034.

ABSTRACT

We investigated the registration of ultrasound volumes based on the mutual information measure, a technique originally applied to multimodality registration of brain images. A prerequisite for successful registration is a smooth, quasi-convex mutual information surface with an unambiguous maximum. We discuss the necessary preprocessing to create such a surface for ultrasound volumes. Abdominal and thoracic organs imaged with ultrasound typically move relative to the exterior of the body and are deformable. Consequently, four specific instances of image registration involving progressively generalized transformations were studied: rigid-body, rigid-body + uniform scaling, rigid-body + nonuniform scaling, and affine. Registration was applied to clinically acquired volumetric images. The accuracy was comparable to the voxel dimension for all transformation modes, although it degraded as the transformation grew more complex. Likewise, the capture range became narrower with the complexity of transformation. As the use of real-time three-dimensional ultrasound becomes more prevalent, the method we present should work well for a variety of applications examining serial anatomic and physiologic changes. Developers of these clinical applications would match the deformation model of their problem to one of the four transformation models presented here.

Key Words: image registration, mutual information, nonrigid image registration, three-dimensional ultrasound

I. INTRODUCTION

Registration of monomodality medical images is an important first step in successful visualization and quantification of temporal changes in anatomy and physiology. Since the bodily organs are fundamentally three-dimensional (3D), a comprehensive picture of serial change is expected to emerge from the registration of 3D images or volumes. A 3D approach is not only the most natural approach to medical image registration, it is also a worthwhile image processing endeavor.

Image registration has been an area of active research [1], and the state-of-the-art brain image registration solves many difficult clinical tasks [2-5]. However, there is a relative shortage of image registration work outside the brain anatomy, and consequently, a dearth of literature on registration techniques involving ultrasound imaging, a modality not suitable for the brain. Ultrasound, however, is ideal for imaging abdominal and thoracic organs, especially the heart. Nonetheless, only a few researchers have published investigations into ultrasound image registration. This may be due to the relatively poor image quality of ultrasound and the nonrigid nature of organs typically imaged with it. Registration techniques developed for the brain do not extend easily to the characteristically low-quality ultrasound images of nonrigid organs.

The lack of investigation in ultrasound image registration may also be due to the primarily two-dimensional (2D) nature of ultrasound. Whereas magnetic resonance imaging (MRI), computed tomography (CT) and various nuclear medicine image modalities have historically produced 3D images, ultrasound has not. Solutions such as reconstructing 3D object by carefully registering 2D images acquired from a conventional ultrasound scanner have been suggested. A 3D volume is reconstructed by translating, rotating or rocking the transducer head uniformly with the aid of a purpose-built mechanical device such that the spatial relationship between the acquired 2D images is known [6]. Alternatively, a 3D volume could be created by freehand manipulation of the transducer head whose orientation is recorded continuously with a wireless 3D localizer [7]. Regardless of transducer localization mechanism, such solutions are too slow to image a dynamic organ such as the heart. Electrocardiogram and respiratory gatings have been employed to reconstruct 3D images of the heart over multiple cardiac cycles, but such data may still have distortions due to cardiac arrythmias and irregular breathing. Even when imaging a static organ, it is difficult to avoid distortion due to patient motion and inherent inaccuracies in 3D localization mechanisms.

Real-time 3D ultrasound acquisition, the most recent advance in ultrasound imaging, addresses both speed and distortion problems inherent in 3D reconstruction solutions. A commercial real-time 3D ultrasound system (Volumetrics, Inc., Durham, North Carolina) is available now, but at very few hospitals only. Where it is available, its use is limited to research investigations. Images are acquired through a 2D phased array of crystals, capable of directing an ultrasound beam anywhere within a 60 x 60 degree pyramid of space. Through the use of 16:1 parallel processing, it is possible to acquire 64 scan lines in 64 scan planes in less than 30 milliseconds, leading to effective volumetric imaging rates of greater than 25 Hz [8]. Although this scanner has the necessary speed for 3D image acquisition, it provides a lateral image resolution poorer than that of the current clinical images (64 vs. greater than 200 scan lines). Development is under way at our institution to produce a real-time 3D ultrasound scanner that works on the principle of synthetic aperture beamforming [9, 10]. This scanner will not compromise lateral image resolution for speed and produces volumetric images comparable to those based on the current clinical lateral resolution. It is not difficult to envision real-time 3D

acquisition as the future of ultrasound imaging, and most hospitals, especially their cardiology departments, adopting real-time 3D ultrasound in the near future.

There are several approaches to image registration, not all of which are applicable to our problem domain – registration of volumetric ultrasound images of the heart and other characteristically nonrigid organs. The anatomy and unique motion of the heart place special constraints on the registration approach. As commonly applied to brain image registration, frame-based techniques [11] or techniques that rely on the placement of external markers on the patient's body [12] assume a rigid underlying anatomy and a fixed spatial relationship of this anatomy with respect to the outside markers. The heart is not only nonrigid, it can move significantly within the chest cavity, rendering such approaches inappropriate. These prospective techniques, in general, have little clinical acceptability because they involve time-consuming acquisition protocols. Retrospective image registration, on the other hand, is nonobtrusive to the existing clinical practice and perhaps the only alternative in the case of abdominal and thoracic organs. Retrospective registration approaches utilize internal anatomic point, contour and surface landmarks, or voxel similarity [1]. Techniques based on internal landmarks, although generalizable to nonrigid transformation, have limitations because they involve some form of image segmentation. Not many point landmarks in ultrasound images of abdominal and thoracic organs can be reliably identified and used for registration. On the contrary, the requirement for the number of point landmarks is even higher to solve for nonrigid transformation. Contour- and/or surface-based approaches rely on accurate segmentation of one or many anatomical structures in the images to be registered. Segmentation of ultrasound images is a difficult problem that usually requires manual intervention for optimal robustness and accuracy. If manual steps are involved, the accuracy of segmentation becomes user-dependent and is always suspect. Segmentation-based registration is limited by the accuracy, reliability and speed of segmentation. A voxel similarity-based approach provides the current best framework for ultrasound volume registration. No segmentation of points, contours or surfaces is required, thereby removing any extrinsic limits on the accuracy and speed. There is also no theoretical limitation on the nature of transformation (rigid or nonrigid) involved. A voxel similarity-based technique has the potential for full automation – another reason for its selection in the present investigation.

We here report results of registration of ultrasound volumes using the mutual information measure [13, 14] of voxel similarity. In particular, we report the accuracy, capture range and execution time for four different modes of possible transformation between two image representations of an anatomy. Although the results are presented for cardiac images, the approach is generalizable to ultrasound images of most other anatomical sites.

II. RELATED WORK

Voxel similarity-based techniques of image registration, especially those involving ultrasound images, form the background for our work. The flexibility of using voxel similarity for image registration has been recognized in the literature [1]. The superiority of a volume-based over a surface-based approach for multimodality registration of brain images has been shown [15]. Many studies [16, 17] have compared various measures of voxel similarity and concluded that mutual information is the most accurate and robust measure for 3D image registration. Although these comparative studies have been performed on non-ultrasound data (brain MRI and single photon emission computed tomography (SPECT) in [16], liver MRI in [17]), we believe, based on these studies and a preliminary study by our group [18], that the same holds true for ultrasound. These studies provide sufficient confidence that mutual information is

a reasonably good measure of voxel similarity. Instead of repeating experiments comparing voxel similarity measures, we have focused on preprocessing, recovery of rigid and nonrigid deformations, and, in general, effectiveness of the voxel similarity approach applied to ultrasound volume registration.

One of the first applications of voxel similarity for registering ultrasound volumes was by Rohling *et al.* [19]. The objective was to register up to six different freehand swept, volumetric ultrasound images of the gall bladder from slightly different viewing directions so that they could be spatially compounded (averaged) to create a 3D image free of acquisition distortions, artifacts and speckles. The specific voxel similarity measure used was the correlation coefficient of gradient images. The authors report the effectiveness of the voxel similarity approach and the adequacy of rigid-body registration to eliminate most artifacts due to organ movement between image acquisitions.

A study with a similar focus to ours is that by Meyer *et al.* [20], who used the mutual information measure successfully to register 3D ultrasound images of the breast. The objective was to register a pair of color flow and/or power Doppler images to create a difference image for serial monitoring of patients in response to chemotherapy or radiation therapy. If imaged from multiple directions, superimposition followed by registration allowed filling in of the flow information missing in a single view. Starting with an approximate registration based on user-selected point landmarks, the investigators refined the registration using voxel similarity. Rigid-body, full affine and elastic transformations were compared. The authors concluded that the affine transformation modeled the deformation of the breast between scans with clinically acceptable accuracy.

III. REGISTRATION METHODS

In this section, we briefly describe the general theory of mutual information-based 3D image registration and explain the ultrasound-specific processing steps we have introduced.

A. *Mutual Information-Based Registration*

The algorithm assumes the existence of two data volumes: one (primary) is kept stationary, and the other (secondary) is transformed iteratively until the optimal alignment between the data volumes, corresponding to the maximum of mutual information function, is achieved. An optimization method searches for the mutual information maximum in the domain of transformation parameters. The mutual information $I(A,B)$ between two data volumes A and B is a function of the individual probability density functions $p(a)$ and $p(b)$ and the joint probability density function $p(a,b)$ of voxel intensities in the overlapping zone of A and B.

$$I(A,B) = \sum_a \sum_b p(a,b) \log\left(\frac{p(a,b)}{p(a)p(b)}\right) \tag{1}$$

Physically, mutual information conveys the amount of information that A contains about B, or vice versa [13, 14].

In our formulation of the problem, the goal of image registration is to obtain a 4 x 4 transformation matrix T_0, in homogeneous coordinates, such that the mutual information measure, $I(A, TB)$, between the primary volume (A) and the transformed secondary volume (TB) is maximized at $T = T_0$. T_0 refers to rigid-body transformation if it incorporates rotation and translation only. The transformation is affine if it includes scaling and shearing as well. In homogeneous coordinates, a 3D vector $\{x, y, z\}$ is represented as $\{x, y, z, 1\}$ and a 3 x 3 linear transformation matrix as a 4 x 4 matrix. Homogeneous coordinates are a handy mathematical

means that allows formulating translation as matrix multiplication just like rotation, scaling and shearing are [21].

B. Modes of Transformation

A generalized affine transformation ($\boldsymbol{T}$) is the cumulative effect of a series of scaling ($\boldsymbol{S}$), shearing ($\boldsymbol{H}$), rotation ($\boldsymbol{R}$) and translation ($\boldsymbol{D}$). Although individual transformations could be combined in many ways, we have restricted the order to the following.

$$\boldsymbol{T} = \boldsymbol{D} \times \boldsymbol{R} \times \boldsymbol{H} \times \boldsymbol{S} \tag{2}$$

The expanded affine transformation matrix appears as below

$$\boldsymbol{T} = \begin{vmatrix} r_{xx} & r_{xy} & r_{xz} & d_x \\ r_{yx} & r_{yy} & r_{yz} & d_y \\ r_{zx} & r_{zy} & r_{zz} & d_z \\ 0 & 0 & 0 & 1 \end{vmatrix} \tag{3}$$

where $\{d_x, d_y, d_z\}$ is the translation vector, and the nine elements of the upper-left 3 x 3 submatrix encompass the combined effect of three rotations $\{\phi_x, \phi_y, \phi_z\}$, three scalings $\{s_x, s_y, s_z\}$ and three shearings $\{\theta_{xy}, \theta_{yz}, \theta_{zx}\}$ (refer to the Appendix for further formulation).

In the present work, we have investigated four different global transformation modes with progressively increasing complexity. The simplest mode corresponds to rigid-body transformation, whereas the most complex is a full affine transformation. The two intermediate modes are limited forms of affine transformation. The first limited affine mode corresponds to rigid-body transformation plus uniform global scaling, whereas the second relaxes the uniformity condition to nonuniform scaling along the three principal axes. Shearing is not allowed in either mode. Totally elastic deformation, in which each data sample of the secondary volume has a unique transformation, is conceivable, but such a registration may warp a pathological region of tissue loss or growth to match perfectly with a region of healthy tissue, defeating the purpose of serial follow-up by subtraction in medical applications.

The transformation mode determines the dimensionality of the search space for optimization. Six parameters, three translations $\{d_x, d_y, d_z\}$ and three rotations $\{\phi_x, \phi_y, \phi_z\}$, are searched in the rigid-body (RB) transformation mode. The uniform scaling (RB+US) mode searches for seven parameters – a global scaling parameter $\{s_x: s_y = s_z = s_x\}$ in addition to six transformation parameters of the RB mode. Three distinct scaling parameters $\{s_x, s_y, s_z\}$, one per principal axis, make the number of parameters searched in the nonuniform scaling (RB+NS) mode equal to nine. The last case, affine transformation (AT) mode, involves 15 geometric parameters that include three translations, three rotations, three scalings and six shearings. The effect of the 15 geometric parameters is, however, expressed by only 12 algebraic parameters in the 4 x 4 homogeneous matrix formulation (see Eq. 3). In the Appendix, we show that only three shearing parameters are unique; the effect of the other three is a combination of the remaining geometric parameters. Therefore, in the AT mode, three shearing parameters $\{\theta_{xy}, \theta_{yz}, \theta_{zx}\}$ were searched in addition to nine parameters of RB+NS mode without any loss of generality.

C. Capture Range

Image registration using the mutual information measure searches for the maximum of mutual information function in the domain of transformation parameters. The desired solution, in general, is a strong local maximum but not necessarily the absolute maximum over the entire search space. This idea can be explained by an extreme example, in which the two volumes overlap, for example, by 5% only. If the intensity distribution in the volume of overlap happened

to be nearly identical in the two data volumes, mutual information would be extremely high. However, such a relative orientation is clearly not the desired solution because the two volumes virtually do not overlap. An implicit assumption in our method, or any other voxel similarity-based method, is that the starting alignment between the two data volumes is in the neighborhood of the desired solution. We use the term *capture range* to convey the notion of a space around the perfect solution such that launching registration anywhere in this space makes convergence on the desired solution extremely likely.

D. Creation of Smooth Mutual Information Function

A requirement for successful registration is that the mutual information function (or any other measure of voxel similarity) at least within the capture range must be quasi-convex with as few local maxima (or ripples) as possible. Although this requirement is often met in mono- or multimodality registration of any combination of typical MRI, CT and nuclear medicine images, the speckle noise in ultrasound images in combination with interpolation artifacts makes this a difficult requirement to meet. In our implementation, interpolation is required to resample the secondary volume on the grid points of the primary volume to build individual and joint histograms. Interpolation artifacts appear in the histograms and get propagated to mutual information computation, subsequently.

Fig. 1 shows surface plots of mutual information function against two forms of elementary misalignments – translations along two principal axes (d_y and d_z) from an arbitrary initial relative orientation – to exemplify the issue of ripples. The mutual information surface for a pair of MRI and SPECT 3D images in panel (a) is quite smooth, whereas the same surface for a monomodality 3D ultrasound image pair in panel (b) has ripples. Based on preliminary experimentation, we hypothesized that: (1) the ripples originate from the combination of speckle noise and the known ill effects of trilinear interpolation [14, 22] in histogram building and mutual information computation, and (2) the ripples confound the search for the maximum corresponding to the desired solution. Removal of undesired local maxima in the mutual information function is key to making optimization robust and reliable. We accomplished this objective in three preprocessing steps. We show the effect of each step individually before showing their combined effect on the shape of the mutual information function.

1) Median Filtering: 3D Ultrasound images were median filtered in a preprocessing step by a 3 x 3 x 3 median filtering kernel. Median filtering suppressed speckles and in turn smoothed the resulting mutual information function, as is apparent from the plot in panel (c) of Fig. 1.

2) Intensity Quantization: Sample points of ultrasound data are typically 1 byte or 8 bits long. Using fewer than 8 upper voxel intensity bits attenuates both signal and noise, but the signal-to-noise ratio (SNR) seems to improve first before decreasing, as the number of bits employed goes from 8 to 1. The higher the SNR, the smoother is the mutual information function. In the preliminary investigation [18], we showed that using either 5 or 6 upper bits allowed the most reliable convergence. Panels (b) and (d) of Fig. 1 have the mutual information surface plots for 8 and 6 bits of intensity quantization, respectively, without median filtering. The surface is smoother and steeper at 6 bits of intensity quantization. All results reported in this study have been compiled for mutual information computed at 5, 6 and 7 bits of intensity quantization. We have used our idea of multifunction simplex [23] that allows optimization by consensus when multiple, similar but slightly differing versions of a function are present.

3) Trilinear Partial Volume Distribution Interpolation: The desirability of trilinear partial volume distribution (PV) interpolation over nearest-neighbor and trilinear interpolations in producing a smoother mutual information function and hence achieving a more robust

optimization behavior has been demonstrated [14]. PV interpolation creates a joint histogram by accumulating fractional weights that trilinear interpolation would use. Unlike trilinear interpolation, however, no new intensity values are created that may be arbitrary. PV interpolation smoothes out the mutual information surface dramatically, as shown in panel (e) of Fig. 1.

The combined effect of median filtering, intensity quantization and PV interpolation on the mutual information surface can be seen in panel (f) of Fig. 1. Although the combined effect may appear only marginally superior to the individual effects (panels (c), (d) and (e)), we expect the overall improvement to be more profound in the multi-dimensional space.

E. Optimization Algorithm

There are several approaches to optimization. Gradient-based approaches, although fast, are sensitive to errors in gradient calculation and the presence of local maxima. Gradient-based search algorithms were ruled out in the present study because of the limitation of the preprocessing to rid mutual information function completely of ripples and local maxima. The simulated annealing approach, despite its robustness, was also discarded because of the excessive time it took to converge. Based on our preliminary experimentation, we selected the simplex method of Nelder and Mead [24] as a compromise between robustness and convergence time. As mentioned earlier, we, in fact, use an enhanced version of the simplex method [23].

Choosing the size of the initial simplex in a multidimensional parameter space is an important step in using simplex optimization. Each axis of the multidimensional parameter space corresponds to a geometric transformation parameter. For RB transformation mode, the space is 6-dimensional with d_x, d_y, d_z, ϕ_x, ϕ_y and ϕ_z as parameters. Prior to determining the size of the initial simplex, a normalization is desired such that a unit step along any parameter axis results in approximately the same physical displacement of the data volume in the spatial domain. The relationship between translation and physical displacement of data volume is direct – a unit translation moves all voxels of a volume by the same fixed amount. The same is, however, not true for rotation, where the displacement of a voxel is dependent on its distance from the axis of rotation. The physical displacement associated with rotation was estimated by the excursion of the farthest vertex from the axis of rotation passing through the center of the data volume. The displacements associated with scaling and shearing were similarly the excursions of the farthest vertex. In the present work, a physical displacement on the order of the voxel dimension was chosen as the normalizing distance. The required translation (in millimeters), rotation (in degrees), scaling (unitless) or shearing (in degrees) to produce this physical displacement was treated as one unit of that parameter in the transformation parameter domain. For the data sets used in the study, a unit parameter distance corresponded to 1.25 mm of translation, 1 degree of rotation, 2% of scale and 2 degrees of shear.

In principle, the size of the initial simplex should be greater than the unit dimension along each axis so that it does not get stuck in a local maximum or in the ripples. The initial size should also not be greater than the capture range, otherwise the convergence may not occur. An initial size roughly 3-5 units along each axis was found satisfactory.

When using optimization, one must also decide on termination conditions. We employed a two-part condition; meeting both parts stopped the optimization. The first part checked for the size of the simplex. If it became smaller than a unit hypercube in the parameter space, the condition was considered met. The second part looked for the range of mutual information values at simplex vertices. This condition was considered met if and when the range became less than 0.001. There was also a failsafe condition, simply the number of iterations, which was

empirically kept at 1000 to prevent optimization from executing indefinitely. The failsafe condition was rarely encountered, as the search would end due to the physically meaningful first condition.

IV. EXPERIMENTAL METHODS

A. *Data Description*

The images used in this study were obtained from a real-time 3D ultrasound scanner manufactured by Volumetrics, Inc. This scanner produced a sequence of volumes, shaped approximately like a truncated pyramid, with 60 degrees azimuth and elevation angular spans, at a frame rate of 25 Hz. The actual number of volumes depended on the heart rate; it varied from 12 to 22 for the data sequences used in the study. The scan depth was 140 mm for the data sets used in the present study.

Volumetrics data sets were acquired natively on a spherical grid with a higher spatial sampling rate along the radial direction (henceforth referred to as the z-axis). The number of samples along the z-axis was 512 or fewer; the actual number depended on the scan depth and the length of the null space coinciding with the near field of the ultrasonic beam. There were 64 samples along azimuth and elevation angles; the lateral sampling rate consequently varied with the depth. To preserve the relatively higher spatial resolution along the z-axis, the data volumes were scan-converted to a 128 x 128 x 512 rectilinear grid, with 512 samples along the z-axis. The rectilinear grid coincided with the rectangular bounding box of the spherical grid.

Five data sequences from three different patients showing the left ventricle in different phases throughout a complete cardiac cycle were used. Images were taken at two different times in the same day in the two patients with two data sequences each. These data were acquired in the Department of Cardiology at our institution.

B. *Data Preprocessing*

Registering the original 128 x 128 x 512 resolution volumes is possible, although, as we demonstrate later, its excessive computational requirement and therefore excessively long execution time posed a practical problem. The results we present required performing registration thousands of times, thus prohibiting use of the data at the original resolution. The volumes were, therefore, spatially subsampled by a factor of two using a 2 x 2 x 2 uniform averaging kernel to create 64 x 64 x 256 resolution data. Median filtering for speckle suppression was performed with a 3 x 3 x 3 kernel in the original resolution prior to subsampling. Given the overall data dimension of 140 mm x 140 mm x 140 mm, the voxel size (dx x dy x dz) in the subsampled data was 2.19 mm x 2.19 mm x 0.55 mm.

C. *Mutual Information Computation*

Creation of 3D binary masks, one corresponding to the primary volume and the other to the secondary volume, preceded mutual information computation. The masks distinguished image voxels from the background, i.e., the null space outside the original ultrasound acquisition pyramid. Each mask was a 3D array whose size matched the size of its corresponding preprocessed data volume. Image voxels were assigned a value of 1 in the masks, whereas the background voxels were given a value of 0. The mask corresponding to the secondary volume, or the secondary mask, underwent the same transformation that the secondary volume underwent during registration. Following each new transformation, the secondary mask was resampled on the grid points of the primary mask, or equivalently, the primary volume, using nearest-neighbor interpolation. The intersection of the primary mask with the transformed secondary mask defined the volume of overlap needed for computing mutual information as described below.

Mutual information was computed using the histogram method [14], which required preparing a joint intensity histogram of the two volumes as well as their individual intensity histograms. Eq. 1 was still usable, except individual histograms were used to represent individual probability density functions and the joint histogram the joint probability density function in the formula. Within the volume of overlap, the secondary volume was resampled on the grid points of the primary volume. The joint histogram was prepared using PV interpolation in all experiments. Trilinear interpolation was used only for a few illustrations in Fig. 1 that compared interpolation methods. When using all 256 gray levels of the original ultrasound images, the resulting joint histogram represented a 256 x 256 discrete function. The individual histogram for the primary volume was prepared by summing the joint histogram entries along the axis corresponding to the secondary volume. Summing along the other axis produced the individual histogram of the secondary volume.

Mutual information at different levels of intensity quantization was computed in the following manner. Subsampling the original histograms by a factor of two, i.e., merging two neighboring bins in the individual histograms and 2 x 2 bins in the joint histogram, produced histograms for 7 bits of intensity quantization. Repeating the process one more time on the subsampled histograms produced histograms for 6 bits of intensity quantization. Yet another subsampling stage produced histograms for 5 bits of intensity quantization The mutual information at 5, 6 and 7 bits of intensity quantization was computed by simply using the corresponding individual and joint histograms.

D. Average Distance Error Computation

As discussed before, the actual physical displacement of each voxel of a 3D data upon a complicated transformation (involving more than translation) is not identical. The physical displacement is usually greatest at the farthest corners of the vertices of the volume. However, it was imperative that we define a metric to quantify the degree of misalignment necessary for computing and comparing registration accuracy for various transformation modes, and expressing capture range. We compute such a metric called *average distance error* that is the average of displacement error (in Euclidean distance) at the eight vertices of a hypothetical cube centered with the bounding box of a data volume. The side length of the cube was chosen to be 100 mm to maintain average distance error as a normalized metric independent of differences in data dimensions. In clinical practice, the scan depth is typically varied from 100 mm to 160 mm to define the optimal window around the heart. Similarly, the two volumes are not expected to have identical dimensions in multimodality applications. Although the volume size variability was not a concern in the present study because all five data sets had identical spatial dimensions, the above definition maintains generality of this metric across multiple applications and meaningfulness of the results presented here for future use. Average distance error was measured in millimeters.

E. Validation Approach

We tagged, retrospectively, the frames of test data sequences with their cardiac phase and registered identical phase frames from two different data sequences of the same patient. The registration was found visually satisfactory in all cased by the cardiologists; however, a quantitative validation could not be performed because the ground truth was not known.

To determine the accuracy of registration of ultrasound volumes for the four transformation modes, we took a self-validation approach. Transforming the secondary volume in an otherwise registered image pair simulated a pair of misaligned volumes. The goal of registration was then to overcome the user-introduced transformation by applying an exactly

opposite transformation to the secondary volume. Comparing the known transformation with the solution of the registration allowed us to determine the accuracy of registration. For each of the five test data sequences, two adjacent end-diastolic frames (separated in time by 40 ms) were chosen as the primary and the secondary volumes. The proximity (in time) of the two frames and the end-diastolic phase, in which the heart is momentarily stationary, allowed us to assume matching shape and size of the cardiac anatomy in the two frames. It further allowed us to assume uncorrelated electronic noise between the two frames.

F. Capture Range Determination

Our general strategy for computing capture range was to start registration from multiple different starting misalignments and observe whether or not the registration was successful. Capture range was defined as the largest starting misalignment at and below which registration was successful 95% of the time. We describe the process of computing capture range for the rigid-body transformation case in detail below. The same process applies to the other modes.

The starting misalignment values, measured in terms of average distance error, were randomly generated such that their distribution was uniform over a 0- to 80-mm range. The choice of the upper limit (80 mm) was arbitrary, but it was kept large enough to be greater than the expected capture range. A total of 1000 starting misalignment values were generated, and each of the five volume pairs were registered with 200 randomly selected values. Panel (a) of Fig. 2 shows a scatter plot with starting misalignment along the horizontal axis and the residual misalignment, i.e., average distance error upon registration, along the vertical axis for all 1000 trials. There does not exist a standard definition for when to call a registration successful. Lacking such a standard, we decided that it would be reasonable to define a threshold for the residual misalignment. If the residual misalignment is below the threshold, the registration will be considered successful. We experimented with three different thresholds, which were one, two and three times the voxel body diagonal ($[dx^2 + dx^2 + dz^2]^{1/2}$), which measured 3.15 mm. Panel (b) of Fig. 2 shows three curves that plot percentage of successful cases at and below a starting misalignment for the three thresholds mentioned. Quite logically, the higher the threshold, the higher was the success rate of registration achieved from a given starting misalignment. Moreover, the success rate decreased with greater starting misalignment for any given threshold. As mentioned before, the capture range was defined as the abscissa of the intersection of the 95% success rate line with the percent success rate curve. Still, to define capture range uniquely, we needed to choose a single threshold value. We selected twice the voxel body diagonal as the threshold in this study. This choice served as a good trade-off between a very stringent subvoxel accuracy requirement of the first threshold (one times the voxel body diagonal) and a very lenient accuracy requirement otherwise.

G. Determination and Comparison of Registration Accuracy

The accuracy of registration is expected to depend on starting misalignment. To determine accuracy, therefore, it was important that the starting misalignment be kept identical during repeated trials. It was also important that the starting misalignment be kept smaller than the capture range; unsuccessful trials would have corrupted the accuracy determination otherwise. Repeated trials were necessary because a one-to-many mapping exists between a specific misalignment value and geometric transformations. Stated another way, infinitely many geometric transformations between two volumes yield the same average distance error. In practice, starting misalignment (expressed as average distance error) could not be computed directly. Instead, given a transformation mode, the set of associated geometric parameters was generated randomly until it produced an average distance error close to the desired value. The

range of possible values for translation, rotation, scaling and shearing was +/-30 mm, +/-30 degrees, +/-30% and +/-30 degrees, respectively. This indirect method of computing average distance error presented difficulty in generating starting misalignments of a unique value. We, therefore, allowed the starting misalignment to vary in a narrow interval of 20 +/-2 mm. We selected 20 mm as the desired starting misalignment because it was smaller than the smallest capture range, as we will report in the next section. Such a selection allowed for a meaningful comparison of accuracy among transformation modes. The accuracy of a specific transformation mode was the average of residual misalignment values upon 200 trials (5 volume pairs x 40 trials/volume pair). We also computed and report the average error in determining each geometric parameter individually.

V. RESULTS

A. *Capture Range*

Knowledge of capture range was necessary in this study to make a judicious choice of the starting misalignment for accuracy measurements and comparisons. Attempting to register grossly misaligned volumes, quite likely outside the capture range, would not have led to meaningful accuracy results. The capture range for the four transformation modes is presented in Table I. As expected, the capture range became narrower with the complexity of the transformation mode and hence the dimensionality of the search space. However, even in the worst case (24 mm in the full affine transformation mode), the capture range was large enough that a human eye could easily discern the associated misalignment.

B. *Effect of Median Filtering*

We showed earlier in Fig. 1 that median filtering the original data created a smoother mutual information function which likely facilitates optimization. The favorable effect of PV interpolation and intensity quantization on optimization has been presented [14, 18]; we present here data demonstrating the effectiveness of median filtering as a preprocessing step. Specifically, we computed registration accuracy and capture range with and without median filtering for all four transformation modes (Table II). The registration accuracy of median filtered images increased with increasing complexity of transformation, and it was generally better than that of the unfiltered images. Although this trend was not met in the RB+NS transformation mode, we attribute this slight inconsistency (a difference of 0.1 mm in accuracy) to experimental error and perhaps lack of adequate data points. When comparing capture range, median filtering consistently allowed a wider capture range except in the RB+US transformation mode, in which the results were identical. Overall, it could be concluded that median filtering aided the registration process positively and, therefore, it was performed as part of preprocessing in all experiments in this study.

C. *Registration Accuracy*

An example of registration of one of the five data pairs is shown in Fig. 5. For each transformation mode corresponding to a row, two orthogonal cross-sections (central XY and central YZ planes) of the fused volume data are presented. Each cross-section is presented twice, showing the relative orientation of the primary and the secondary volumes before and after registration. The primary volume has been depicted with shades of green using the green channel of the RGB color triplet, whereas shades of magenta (red and blue channels) depict the secondary volume. Shades of gray result upon registration when comparable intensities of green and magenta are fused. A visual approach to evaluate the success of registration, therefore, is to

look for a higher occurrence of gray. Qualitatively, a good matching of anatomical structures is apparent following registration in all transformation modes.

Results of registration on all five volume pairs are shown in Table III. The registration accuracy, averaged for 200 trials (5 volume pairs x 40 trials/volume pair), is presented in the second column. This accuracy decreases with the complexity of transformation. It should be noted that the subvoxel accuracy (less than 3.15 mm, the voxel body diagonal) was achieved in RB and RB+US transformation modes. In the other two modes, the accuracy was only slightly worse than the subvoxel accuracy. The root mean squared (rms) deviation (or estimation error) of each parameter from the expected solution is reported in subsequent columns. The top number in each row is the deviation from the zero transformation, the obvious expected solution. Since there is bound to be some, although very small, mismatch between the two starting frames we used, the zero transformation may not be the ideal solution. As an alternative, we considered the median of all solutions per volume pair (to be called median solution) as the expected solution and measured rms deviation of each geometric parameter from median solution also. The numbers in the bottom of each row designate deviation from the median solution. Not surprisingly, the estimated parameters were closer to the median solution than the zero transformation solution. Barring some exceptions, a general trend was increasing error in the estimation of each transformation parameter with increasing complexity of transformation. d_x (translation along x-axis) was estimated with 0.91 mm rms error in RB mode, whereas the same error increased to 1.45 mm in the AT case. This observation is likely due to the ambiguity arising from a higher number of parameters with more complex transformation modes. A single elementary transformation or a combination of them may approximate another single elementary transformation. As an example, scaling along the z-axis may correct for translational error along the same axis in some parts of the volume. If the scaling is not allowed at all, as would be the case in RB mode, the translation parameter will be more accurately determined.

Anisotropic image sampling may result in greater error in estimating transformation parameters that either cause translation parallel to the axis of lower image resolution or cause rotation in a lower resolution plane. No consistent effect of anisotropic image sampling could be deduced from the numbers we obtained, which may be attributed to the achievement of subvoxel and near subvoxel accuracy.

D. Comparison of Rigid-Body and Full Affine Registration for Recovery of Rigid-Body Misalignment

As shown in the previous subsection, the recovery of rigid-body misalignment through rigid-body registration was more accurate than the recovery of full affine misalignment through full affine registration. In this subsection, we present results on the recovery of rigid-body misalignment through rigid-body as well as full affine registration.

The results presented here were obtained from 200 trials on a randomly selected volume pair. In panel (a) of Fig. 3, we show the scatter plot of residual versus starting misalignment upon both rigid-body registration ('*' points) and full affine registration ('o' points). It is apparent that both modes perform roughly identically below approximately 15 mm of the starting misalignment. Between approximately 15-45 mm of the starting misalignment, the full affine registration starts to incur progressively greater error than the rigid-body registration. Given that the rigid-body transformation is a special case of affine transformation, one may, quite logically, expect the affine registration to produce a result identical to what rigid-body registration produces, with the exception that the scaling and shearing parameters are zero. It appears, however, that the presence of additional geometric parameters causes these parameters to start

assuming nonzero values in the optimization process. To accommodate nonzero scaling and shearing parameters, the translation and rotation parameters veer off from the values they should converge to, thus affecting the overall accuracy of registration. Moreover, the greater the starting misalignment, the more the optimization may be confounded by the presence of unnecessary parameters. Perhaps this effect caused a narrower capture range with higher parameters. The capture range values (panel (b) of Fig. 3) were 45 mm and 22 mm, respectively, for rigid-body and affine registrations. Note also the striking similarity between the recovery of rigid-body misalignment and the recovery of full affine misalignment via full affine registration.

E. Comparison of Rigid-Body and Full Affine Registration for Recovery of Full Affine Misalignment

Recovery of full affine misalignment using rigid-body and full affine transformations is presented in this subsection. As before, the results are a compilation of 200 trials on a randomly selected volume pair. The difficulty of rigid-body registration in overcoming affine misalignment of even small magnitude is evident from the scatter plot in Fig. 4 (note the 'o' points). This is likely due to the inability of translation and rotation parameters in adequately compensating for the scaling and shearing transformations. In contrast, the full affine registration is successful up to a starting misalignment of approximately 25 mm before the residual misalignment starts becoming significant. Panel (b) of Fig. 3 also has the plots showing the success rate of both rigid-body and full affine registrations on full affine starting misalignment. A very narrow capture range of 8 mm for rigid-body registration in this case is evident. The capture range for the affine registration was found to be 25 mm.

F. Registration at Multiple Data Resolutions and Execution Times

Multiresolution approach is a well-known strategy to expedite voxel similarity-based image registration [16]. Although we did not perform multiresolution image registration in the present study, we show data on the feasibility of such a registration for the 3D ultrasound images. In Table IV, the accuracy and capture range are presented for all four transformation modes at three image resolutions – the original image resolution (128 x 128 x 512) and two subsampled versions (64 x 64 x 256 and 32 x 32 x 128). Registration did not succeed for a resolution coarser than 32 x 32 x 128. The results are a compilation of 200 trials on a randomly selected volume pair. The accuracy numbers suggest that a three-level coarse-to-fine registration is possible. Upon registration at the coarsest level, the two subsequent registrations at finer levels should take considerably less time to converge because they would be significantly closer to the solution to start. A drawback is that capture range of the coarsest resolution, modestly smaller than that of the higher resolutions, will decide the capture range of the multiresolution image registration.

The execution time of registration, together with the accuracy and the capture range, determines the clinical usefulness and acceptability of registration-based applications. Although not thoroughly optimized, the execution times for the registration program written in C++ and running on a dual 933 MHz Pentium-III processor personal computer with 1 GB of memory are reported in Table V. A time range is reported because the exact time depends on the specific nature (i.e., the presence and the extent of the elementary transformations) of the starting misalignment. The timings reported are for 64 x 64 x 256 resolution data. Optimization with more parameters required a greater number of mutual information evaluations, proportionately increasing the execution time. The execution time was also approximately proportional to the number of voxels. Therefore, rigid-body registration of the original 128 x 128 x 512 resolution

images took 16-24 minutes. On the other hand, 32 x 32 x 128 resolution images were registered in less than 1 minute in all modes.

VI. DISCUSSION

We have demonstrated that successful registration of ultrasound volumes, despite their characteristic poor image quality, is possible using the mutual information measure of voxel similarity. Furthermore, we have demonstrated that the image registration can recover a global deformation as simple as a rigid-body transformation and as complex as an affine transformation. We discuss the features and performance of our method and the likely clinical applications below.

A successful approach to ultrasound image registration should be fundamentally 3D and compensate for both rigid and nonrigid deformations of the underlying organ. Moreover, it should be robust, accurate and fast and should require minimal user intervention. Our method is 3D and accommodates global nonrigid deformation of an organ between image acquisitions. We note that the shape mismatch between two representations of an organ from two different instants could arise from two sources. The first is the deformation due to patient positioning and the other due to physiological and pathological changes in the organ itself. For successful serial comparison, it is imperative that the registration account for only the patient positioning component, not the physiological and pathological changes. The shape of a deformable soft tissue organ cannot be the same between differing patient positions such as the supine and the prone due to gravity and pressure changes from neighboring organs. Even small variations in a given position (supine, for example) between image acquisitions could contribute to deformation. Although this deformation will be nonrigid generally, it is expected to be global primarily because gravity, the major contributor, is uniform. The physiological and pathological changes such as tissue growth or decay, on the other hand, are expected to be local. Although a totally elastic registration is possible within the mutual information-based image registration framework, a caveat with elastic image registration is the loss of structural changes that are of clinical interest. Constraining deformation to a global transformation during registration, at least in principle, allows serial follow-up without patient positioning errors. An enhancement would exclude any locally diseased regions from the registration process. Meyer *et al.* [20] showed that the affine transformation was adequate for registering images of the breast, a highly deformable organ. A totally elastic registration, however, does make sense in intermodality registration, where visualizing local structural changes are not an issue and the goal is fusion of complementary information.

In general, and as discussed by Carrillo *et al.* [17] also, the accuracy of image registration is difficult to assess for organs that can deform or move with respect to the exterior of the body, which includes virtually all organs imaged with ultrasound. External marker-based approaches used successfully for validating brain image registration do not apply to the case of deformable organs, thus limiting one to comparing internal landmarks, concurrence with experts and recovery of user-introduced transformations. We took the latter approach by registering two neighboring frames from sequences of frames acquired rapidly at 25 Hz. The temporal proximity allowed us to assume equality of the anatomy.

Overall, the accuracy achieved in all four global transformation modes was highly encouraging. In RB and RB+US scaling transformation modes, subvoxel accuracy, defined here as the average registration error being smaller than the maximum voxel dimension, was achieved. In RB+NS and AT modes, the accuracy was only slightly worse than the subvoxel accuracy. Studholme *et al.* [16] reported 3 mm and 4 degrees (7 mm displacement 100 mm away

from the center of axis) as the limits of accuracy when human experts performed a manual registration of MRI and positron emission tomography (PET) brain images. The specified numbers for translational accuracy were comparable to the voxel size of PET, the lower resolution image in the pair. Although no such data are available for ultrasound images of the heart, optimistic limits of human accuracy could be considered equivalent to the voxel size. Using the voxel dimension yardstick, the accuracy is excellent in the RB and RB+US transformation modes and quite acceptable in the RB+NS and AT modes. The accuracy of the last two modes can be improved in absolute terms with the use of the original resolution data; however, the condition for subvoxel accuracy at the higher resolution may still not be met. Furthermore, the accuracy numbers presented by us pertain to the best case scenario because they were obtained from simulated experiments performed under ideal conditions. For that reason, the numbers serve as the upper limit of accuracy expected in a real application.

The capture range data confirmed the expected answer, that the capture range becomes narrower in more complex transformation modes in which a greater number of parameters need to be optimized. However, the encouraging finding was that even the worst-case capture range (24 mm in the full affine mode) was large enough to pose no difficulty in practice. When designing a practical image registration system, it is reasonable to require that the starting misalignment be smaller than the capture range. This requirement becomes easy to meet if the capture range is larger than the accuracy of a coarse registration, performed either manually or automatically, for initial "seeding" of the volumes. With capture range being at least eight times the maximum voxel dimension, initial seeding becomes relatively easier.

We have found the simplex optimization to be most successful with ultrasound data even though Powell's method has typically been used in voxel similarity registration in reported studies [16, 17]. Simplex optimization, however, is computationally intensive. Although the execution time of a few minutes for volumes subsampled by a factor of two was acceptable for most offline application, it does become excessively long (from 16-24 minutes in RB mode to 40-64 minutes in the full affine mode) when registration at the original image resolution is attempted. While developing dedicated registration hardware is possible, the multiresolution strategy is a convenient alternative to speed up execution moderately. One would register images at 32 x 32 x 128, 64 x 64 x 256 and 128 x 128 x 512 resolutions in order. The overall execution time should be lower because fewer than normal iterations will be required at the highest resolution; it would be somewhere between the times taken at the lowest and the highest resolutions. A penalty paid for faster execution is the limited capture range that equals the capture range at the lowest resolution. A limited multiresolution strategy using only the upper two resolution levels, therefore, may provide a compromise between speed and adequately large capture range.

Another coarse-to-fine strategy is to recover a complex deformation (for example, full affine transformation) starting with rigid-body registration followed by progressively more complex transformation modes [20]. An obvious advantage of this strategy is savings in the execution time. A second advantage may be a larger capture range. However, our simulation showed that the capture range shrank significantly when the recovery of affine deformation was attempted using rigid-body registration. This observation seriously weakens the prospect of such a coarse-to-fine registration. Nonetheless, we speculate that a limited coarse-to-fine registration may be advantageous. In such an approach one will, for example, use RB+NS and full affine registration modes to recover full affine deformation. The coarse-to-fine registration strategy did not appear promising from our simulation perhaps because we allowed large scaling and

shearing deformations. In real-life situations, both rigid and nonrigid (arising from scaling, shearing and/or elastic transformation) deformation may be present, but typically the deformation is rigid to a greater degree. A clinical deformation therefore may not exactly mirror our simulated affine deformation. A coarse-to-fine registration approach, therefore, may be more successful for naturally occurring nonrigid deformation, as was the case in [20].

Contrary to the coarse-to-fine strategy is the use of a higher parameter transformation model to recover a simpler deformation. Although there are no apparent disadvantages in terms of accuracy, such an attempt will require longer execution time and have a narrower capture range. We conclude that the best strategy, therefore, is to use *a priori* information to match the underlying deformation as best as possible to the transformation mode used in registration.

Clinical applications of a generalized, accurate and robust 3D image registration technique could be many. At the very least, it would allow comparison, in qualitative terms, of the images of the organs such as the kidney and the liver. As has been shown [19], image quality could be improved by registration and subsequent superimposition of several successive scans. Many cardiac applications are possible as well. One specific application to which we are applying the developed techniques is the alignment of pre- and post-stress ultrasound 3D images of the heart. Once registered, a side-by-side presentation of pre- and post-stress images along any arbitrary orientation is possible, allowing a physician to perform accurate and comprehensive diagnosis.

VII. CONCLUSION

We have demonstrated that mutual information-based registration, originally applied to multimodality registration of brain images, is effective for registration of 3D ultrasound images. Furthermore, given the deformable nature of organs typically imaged with ultrasound, we can apply the same framework for nonrigid deformations such as rigid-body transformation with both uniform and nonuniform scaling and affine transformation. The accuracy of the registration is comparable to the voxel dimension. The capture range, which becomes smaller with the complexity of transformation, is large enough for most practical applications.

APPENDIX
TRANSFORMATION MATRIX FORMULATION

A generalized affine transformation is the product of scaling ($\boldsymbol{S}$), shearing ($\boldsymbol{H}$), rotation ($\boldsymbol{R}$) and translation ($\boldsymbol{D}$) matrices. For the translation vector $\{d_x, d_y, d_z\}$ and the scaling vector $\{s_x, s_y, s_z\}$, the translation and scaling matrices are expressed as

$$\boldsymbol{D} = \begin{vmatrix} 1 & 0 & 0 & d_x \\ 0 & 1 & 0 & d_y \\ 0 & 0 & 1 & d_z \\ 0 & 0 & 0 & 1 \end{vmatrix} \quad \text{and} \quad \boldsymbol{S} = \begin{vmatrix} s_x & 0 & 0 & 0 \\ 0 & s_y & 0 & 0 \\ 0 & 0 & s_z & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix}.$$

The rotation matrix is the product of three matrices representing individual rotation about the x, y and z axes, respectively, by angles ϕ_x, ϕ_y and ϕ_z. $\boldsymbol{R} = \boldsymbol{R}_z * \boldsymbol{R}_y * \boldsymbol{R}_x$, where

$$\boldsymbol{R}_z = \begin{vmatrix} \cos\phi_z & -\sin\phi_z & 0 & 0 \\ \sin\phi_z & \cos\phi_z & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix}, \quad \boldsymbol{R}_y = \begin{vmatrix} \cos\phi_y & 0 & -\sin\phi_y & 0 \\ 0 & 1 & 0 & 0 \\ \sin\phi_y & 0 & \cos\phi_y & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} \quad \text{and}$$

$$\boldsymbol{R}_x = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\phi_x & -\sin\phi_x & 0 \\ 0 & \sin\phi_x & \cos\phi_x & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix}.$$

The shearing matrix is a product of six matrices of the form

$$\boldsymbol{H}_{ab} = \begin{vmatrix} 1 & h_{xy} & h_{xz} & 0 \\ h_{yx} & 1 & h_{yz} & 0 \\ h_{zx} & h_{zy} & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix}$$

where $h_{cd} = \begin{cases} -\tan\theta_{ab}, & \text{if } a = c \text{ and } b = d \\ 0, & \text{otherwise} \end{cases}$, a, b, c and d assume values x, y and z.

Each pair of axes produces two shears, θ_{ab} and θ_{ba}, but only one of them is unique. Consequently, only three out of six shear parameters need to be used for transformation matrix formulation for registration. To prove this, let us consider the transformation $\boldsymbol{H}_3$ achieved with the set of three redundant shears.

$$\boldsymbol{H}_3 = \boldsymbol{H}_{yx} * \boldsymbol{H}_{xz} * \boldsymbol{H}_{zy} = \begin{vmatrix} 1 & 0 & -\tan\theta_{xz} & 0 \\ -\tan\theta_{yx} & 1 & 0 & 0 \\ \tan\theta_{yx} * \tan\theta_{zy} & -\tan\theta_{zy} & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix}.$$

Let us consider a second transformation $\boldsymbol{H}_9$ incorporating scaling, three rotations and the other three shear matrices.

$$\boldsymbol{H}_9 = \boldsymbol{S} * \boldsymbol{H}_{xy} * \boldsymbol{H}_{zx} * \boldsymbol{H}_{yz} * \boldsymbol{R}_x * \boldsymbol{R}_y * \boldsymbol{R}_z =$$

$$= \begin{vmatrix} s_x(\cos\phi_y - (\sin\phi_x \tan\theta_{xy} + \cos\phi_x \tan\theta_{xy} \tan\theta_{yz})\sin\phi_y)\cos\phi_z - s_x(\cos\phi_x \tan\theta_{xy} - \sin\phi_x \tan\theta_{xy} \tan\theta_{yz})\sin\phi_z & \dots \\ s_y(\sin\phi_x + \cos\phi_x \tan\theta_{yz})\sin\phi_y \cos\phi_z + s_y(\cos\phi_x - \sin\phi_x \tan\theta_{yz})\sin\phi_z & \dots \\ -s_z(\cos\phi_y \tan\theta_{zx} - \cos\phi_x \sin\phi_y)\cos\phi_z + s_z \sin\phi_x \sin\phi_z & \dots \\ 0 & \dots \end{vmatrix}$$

$$\begin{vmatrix} \dots & -s_x(\cos\phi_y - (\sin\phi_x \tan\theta_{xy} + \cos\phi_x \tan\theta_{xy} \tan\theta_{yz})\sin\phi_y)\sin\phi_z - s_x(\cos\phi_x \tan\theta_{xy} - \sin\phi_x \tan\theta_{xy} \tan\theta_{yz})\cos\phi_z & \dots \\ \dots & -s_y(\sin\phi_x + \cos\phi_x \tan\theta_{yz})\sin\phi_y \sin\phi_z + s_y(\cos\phi_x - \sin\phi_x \tan\theta_{yz})\cos\phi_z & \dots \\ \dots & s_z(\cos\phi_y \tan\theta_{zx} - \cos\phi_x \sin\phi_y)\sin\phi_z + s_z \sin\phi_x \cos\phi_z & \dots \\ \dots & 0 & \dots \end{vmatrix}$$

$$\begin{vmatrix} \dots & s_x \sin\phi_y + s_x(\sin\phi_x \tan\theta_{xy} + \cos\phi_x \tan\theta_{xy} \tan\theta_{yz})\cos\phi_y & 0 \\ \dots & -s_y(\sin\phi_x + \cos\phi_x \tan\theta_{yz})\cos\phi_y & 0 \\ \dots & -s_z(\sin\phi_y \tan\theta_{zx} - \cos\phi_x \cos\phi_y) & 0 \\ \dots & 0 & 1 \end{vmatrix}$$

Element-by-element comparison of matrices $\boldsymbol{H}_3$ and $\boldsymbol{H}_9$ provides the following set of equations and solutions:

$H_9(2,3) = 0 \quad \Rightarrow \quad \theta_{yz} = -\phi_x;$

$H_9(1,2) = 0 \quad \Rightarrow \quad \theta_{xy} = \arctan(-\tan\phi_z \cos\phi_y \cos\phi_x);$

$H_9(2,1) * H_9(3,2) = H_9(3,1) \quad \Rightarrow \quad \theta_{zx} = \arctan(\tan\phi_y \cos\phi_x);$

$$
\begin{aligned}
H_9(1,1) = 1 \quad &\Rightarrow \quad s_x = \cos\phi_z \,/\, \cos\phi_y; \\
H_9(2,2) = 1 \quad &\Rightarrow \quad s_y = \cos\phi_x \,/\, \cos\phi_z; \\
H_9(3,3) = 1 \quad &\Rightarrow \quad s_z = \cos\phi_y \,/\, \cos\phi_x. \\
H_9(2,1) = -\tan\theta_{yx} \quad &\Rightarrow \quad \phi_z = -\theta_{yx}; \\
H_9(1,3) = -\tan\theta_{xz} \quad &\Rightarrow \quad \phi_y = \arctan(\tan\theta_{xz} / \cos\phi_z); \\
H_9(3,2) = -\tan\theta_{zy} \quad &\Rightarrow \quad \phi_x = \arctan(-\tan\theta_{zy} / (\cos\phi_y \cos\phi_z)).
\end{aligned}
$$

To summarize, three shearing parameters of $\boldsymbol{H}_3$ can be expressed in terms of the other three shearing, three rotation and three scaling parameters.

ACKNOWLEDGMENTS

The authors thank Drs. James Thomas, Takahiro Shiota and Neil Greenberg of the Department of Cardiology at The Cleveland Clinic Foundation for making available the images used in the study, and Ms. Christine Kassuba of the Department of Biomedical Engineering for editorial assistance with the preparation of this manuscript.

REFERENCES

[1] J. B. Maintz and M. A. Viergever, "A survey of medical image registration," *Med. Image Anal.*, vol. 2, pp. 1-36, 1998.

[2] J. C. Gee, "On matching brain volumes," *Pattern Recogn.*, vol. 32, pp. 99-111, 1999.

[3] P. A. Freeborough, R. P. Woods, and N. C. Fox, "Accurate registration of serial 3D MR brain images and its application to visualizing change in neurodegenerative disorders," *J. Comput. Assist. Tomogr.*, vol. 20, pp. 1012-22, 1996.

[4] M. I. Miga, K. D. Paulsen, P. J. Hoopes, F. E. Kennedy, Jr., A. Hartov, and D. W. Roberts, "In vivo quantification of a homogeneous brain deformation model for updating preoperative images during surgery," *IEEE Trans. Biomed. Eng.*, vol. 47, pp. 266-273, 2000.

[5] J. L. Andersson, A. Sundin, and S. Valind, "A method for coregistration of PET and MR brain images," *J. Nuclear Med.*, vol. 36, pp. 1307-15, 1995.

[6] A. Fenster and D. B. Downey, "3-D ultrasound imaging: A review," *IEEE Eng. Med. Biol.*, vol. 15, pp. 41-51, 1996.

[7] S. Berg, H. Torp, D. Martens, E. Steen, S. Samstad, I. Hoivik, and B. Olstad, "Dynamic three-dimensional freehand echocardiography using raw digital ultrasound data," *Ultrasound Med. Biol.*, vol. 25, pp. 745-753, 1999.

[8] O. T. von Ramm, S. W. Smith, and B. A. Carroll, "Advanced real-time volumetric ultrasound scanning," *J. Ultrasound Med.*, vol. 14, p. S35, 1995 [Abstract].

[9] C. R. Hazard and G. R. Lockwood, "Theoretical assessment of a synthetic aperture beamformer for real-time 3-D imaging," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, vol. 46, pp. 972-980, 1999.

[10] G. R. Lockwood, J. R. Talman, and S. S. Brunke, "Real-time 3-D ultrasound imaging using sparse synthetic aperture beamforming," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, vol. 45, pp. 980-988, 1998.

[11] T. Peters, B. Davey, P. Munger, R. Comeau, A. Evans, and A. Olivier, "Three-dimensional multimodal image-guidance for neurosurgery," *IEEE Trans. Med. Imag.*, vol. 15, pp. 121-128, 1996.

[12] R. T. Malison, E. G. Miller, R. Greene, G. McCarthy, D. S. Charney, and R. B. Innis, "Computer-assisted coregistration of multislice SPECT and MR brain images by fixed external fiducials," *J. Comput. Assist. Tomogr.*, vol. 17, pp. 952-960, 1993.

[13] W. M. Wells, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis, "Multi-modal volume registration by maximization of mutual information," *Med. Image Anal.*, vol. 1, pp. 35-51, 1996.

[14] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Trans. Med. Imag.*, vol. 16, pp. 187-98, 1997.

[15] J. West, J. M. Fitzpatrick, M. Y. Wang, B. M. Dawant, C. R. Maurer, Jr., R. M. Kessler, and R. J. Maciunas, "Retrospective intermodality registration techniques for images of the head: Surface-based versus volume-based," *IEEE Trans. Med. Imag.*, vol. 18, pp. 144-150, 1999.

[16] C. Studholme, D. L. Hill, and D. J. Hawkes, "Automated three-dimensional registration of magnetic resonance and positron emission tomography brain images by multiresolution optimization of voxel similarity measures," *Med. Phys.*, vol. 24, pp. 25-35, 1997.

[17] A. Carrillo, J. L. Duerk, J. S. Lewin, and D. L. Wilson, "Semiautomatic 3-D image registration as applied to interventional MRI liver cancer treatment," *IEEE Trans. Med. Imag.*, vol. 19, pp. 175-185, 2000.

[18] V. Zagrodsky, R. Shekhar, and J. F. Cornhill, "Mutual information based registration of cardiac ultrasound volumes," in *Proc. SPIE - Int. Soc. Optical Eng.*, vol. 3979, pp. 1605-1614, 2000.

[19] R. N. Rohling, A. H. Gee, and L. Berman, "Automatic registration of 3-D ultrasound images," *Ultrasound Med. Biol.*, vol. 24, pp. 841-854, 1998.

[20] C. R. Meyer, J. L. Boes, B. Kim, P. H. Bland, G. L. Lecarpentier, J. B. Fowlkes, M. A. Roubidoux, and P. L. Carson, "Semiautomatic registration of volumetric ultrasound scans," *Ultrasound Med. Biol.*, vol. 25, pp. 339-347, 1999.

[21] I. N. Bankman, *Handbook of Medical Imaging: Processing and Analysis*. San Diego, CA: Academic Press, 2000.

[22] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Interpolation artefacts in mutual information-based image registration," *Comput. Vision Image Understanding*, vol. 77, pp. 211-232, 2000.

[23] V. Zagrodsky, R. Shekhar, and J. F. Cornhill, "Multi-function extension of simplex optimization method for mutual information based registration of ultrasound volumes," in *Proc. SPIE - Int. Soc. Optical Eng.*, vol. 4322, pp. 508-515, 2001.

[24] J. A. Nelder and R. Mead, "A simplex method for function minimization," *Comput. J.*, vol. 7, pp. 308-313, 1965.

TABLE I

CAPTURE RANGE AS A FUNCTION OF TRANSFORMATION MODE

Transformation Mode	Capture range (mm)
RB	44
RB+US	43
RB+NS	26
AT	24

RB: Rigid-body; RB+US: Uniform scaling; RB+NS: Nonuniform scaling; AT: Affine transformation

TABLE II

ACCURACY AND CAPTURE RANGE WITH AND WITHOUT MEDIAN FILTERING FOR THE FOUR TRANSFORMATION MODES OF IMAGE REGISTRATION

Transformation Mode	Accuracy (mm)		Capture range (mm)	
	With median filtering	Without median filtering	With median filtering	Without median filtering
RB	1.4	1.4	56	55
RB+US	1.7	2.1	44	32
RB+NS	4.0	3.9	26	26
AT	4.2	4.6	24	19

RB: Rigid-body; RB+US: Uniform scaling; RB+NS: Nonuniform scaling; AT: Affine transformation

TABLE III

EXECUTION TIME FOR 64 X 64 X 256 RESOLUTION DATA

Transformation mode	Number of mutual information evaluations for convergence	Execution time (minutes)
RB	150 – 200	2 – 3
RB+US	180 – 250	3 – 4
RB+NS	220 – 350	4 – 6
AT	280 – 460	5 – 8

RB: Rigid-body; RB+US: Uniform scaling; RB+NS: Nonuniform scaling; AT: Affine transformation

TABLE IV

SUMMARY OF REGISTRATION RESULTS FOR ALL FIVE VOLUME PAIRS. TOP AND BOTTOM NUMBERS IN EACH ROW CORRESPOND TO DEVIATIONS FROM THE ZERO TRANSFORMATION SOLUTION AND THE MEDIAN SOLUTION, RESPECTIVELY

Transformation mode	Accuracy (mm)	d_x (mm)	d_y (mm)	d_z (mm)	ϕ_x (deg)	ϕ_y (deg)	ϕ_z (deg)	s_x (%)	s_y (%)	s_z (%)	θ_{xy} (deg)	θ_{yz} (deg)	θ_{zx} (deg)
RB	2.11	0.91 0.52	0.41 0.38	0.89 0.36	0.60 0.57	0.71 0.66	1.03 1.02						
RB+US	2.19	0.86 0.47	0.39 0.38	0.95 0.40	0.70 0.60	0.73 0.67	0.99 0.92	0.63 0.57					
RB+NS	3.95	1.12 0.80	0.56 0.50	2.01 1.92	0.76 0.74	1.19 1.10	1.10 1.03	2.50 2.20	1.90 1.90	5.20 4.60			
AT	4.08	1.45 1.27	0.61 0.56	1.48 1.42	1.84 1.75	2.21 2.12	1.05 0.94	2.30 2.10	1.90 1.80	4.10 3.60	1.40 1.30	2.20 2.10	2.20 2.10

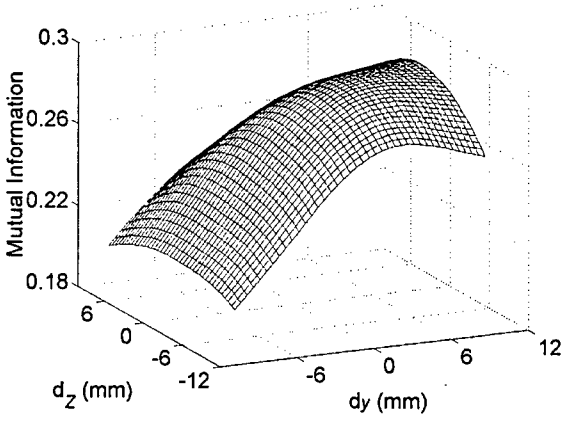
RB: Rigid-body; RB+US: Uniform scaling; RB+NS: Nonuniform scaling; AT: Affine transformation.

TABLE V

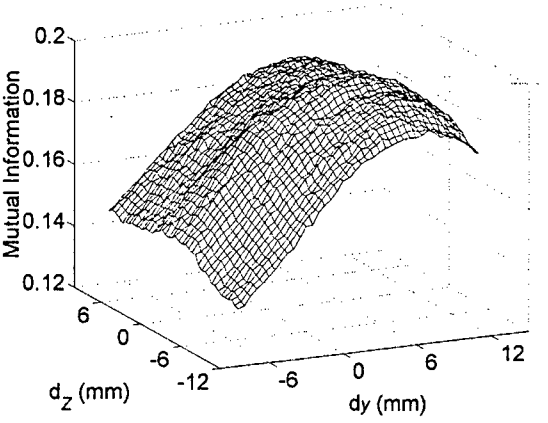
ACCURACY AND CAPTURE RANGE OF REGISTRATION AT MULTIPLE RESOLUTIONS

Transformation mode	Accuracy (mm)			Capture range (mm)		
	128 x 128 x 512	64 x 64 x 256	32 x 32 x 128	128 x 128 x 512	64 x 64 x 256	32 x 32 x 128
RB	1.7	1.4	2.0	56	56	42
RB+US	1.7	1.7	2.9	50	44	37
RB+NS	3.0	4.0	5.3	42	26	16
AT	2.7	4.2	5.2	39	24	17

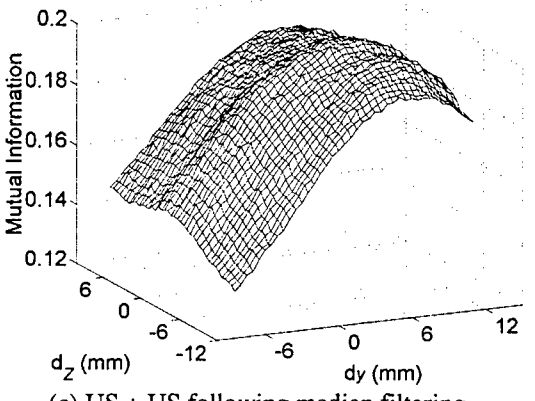
RB: Rigid-body; RB+US: Uniform scaling; RB+NS: Nonuniform scaling; AT: Affine transformation



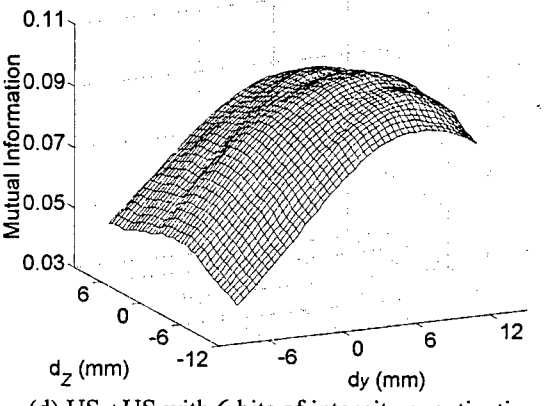
(a) MRI + SPECT



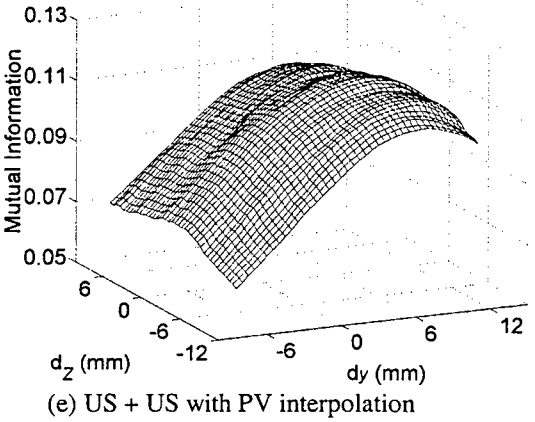
(b) US + US



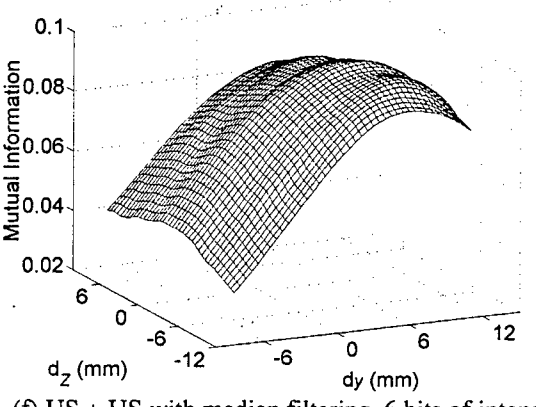
(c) US + US following median filtering



(d) US +US with 6 bits of intensity quantization



(e) US + US with PV interpolation



(f) US + US with median filtering, 6 bits of intensity quantization and PV interpolation

Fig. 1. Mutual information as a function of misalignment for an MRI and SPECT image pair (panel (a)) and a 3D ultrasound (US) image pair (panel (b)) with no preprocessing. Panels (c), (d) and (e) show the mutual information surface plots, following median filtering, 6 bits of intensity quantization, and trilinear partial volume distribution (PV) interpolation, respectively. Note the inherent smoothness of the mutual information surface for the non-ultrasound image pair in panel (a) and the roughness of the same for the ultrasound image pair in panel (b). The combined effect of median filtering, intensity quantization and PV interpolation on the smoothness of the mutual information surface is shown in panel (f).

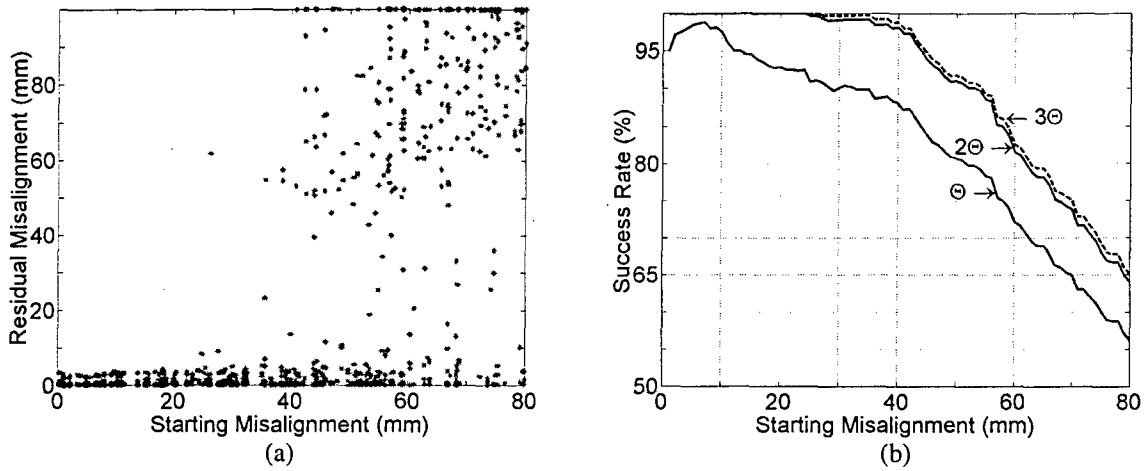
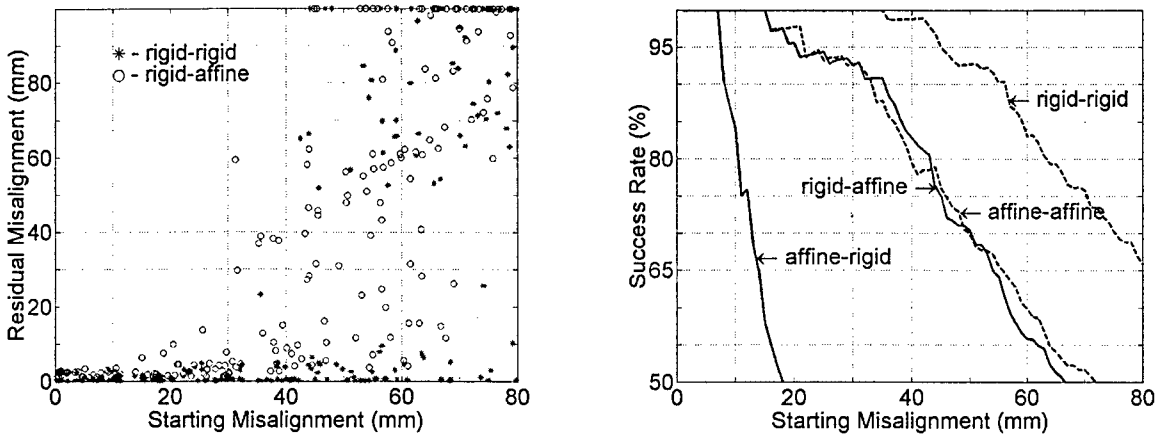


Fig. 2. Results of capture range determination experiments: (a) Scatter plot of residual misalignment against starting misalignment, and (b) Percent success rate plotted against starting misalignment for three different accuracy thresholds, Θ, 2Θ, and 3Θ, where Θ equals the voxel body diagonal.



(a) (b)

Fig. 3. (a) Scatter plot showing residual and rigid-body starting misalignments following rigid-body ('*' points) and full affine ('o' points) registrations. (b) Success rate plots for all four combinations of rigid-body and full affine misalignments through rigid-body and full affine registrations. The first word in the legends corresponds to the type of misalignment, whereas the second word indicates the type of registration.

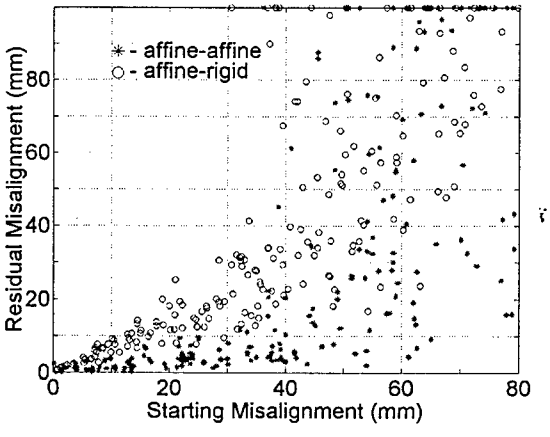


Fig. 4. Scatter plot showing residual and full affine starting misalignments following rigid-body ('o' points) and full affine ('*' points) registrations.

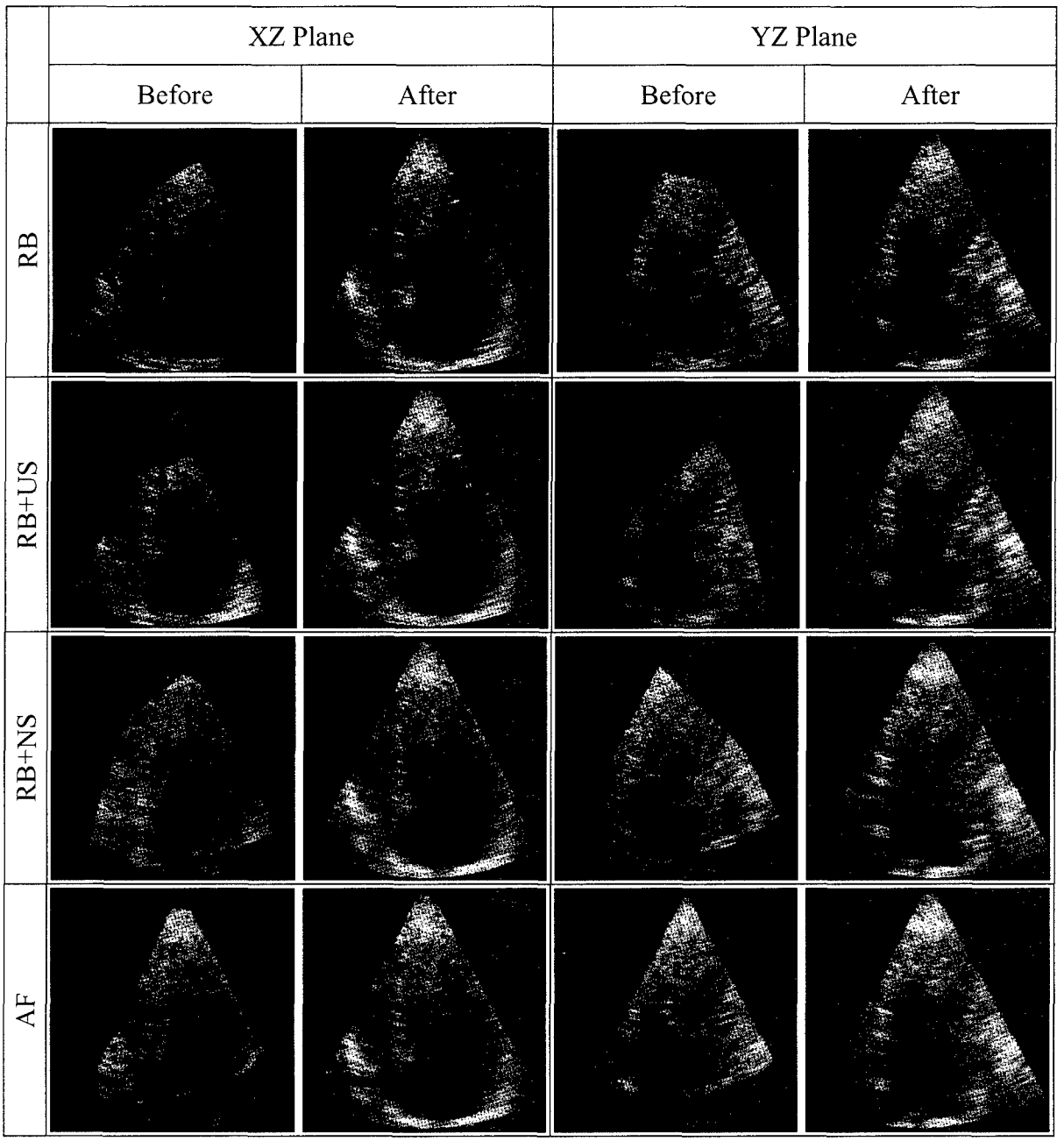


Fig. 5. Fused primary and secondary volumes before and after registration for all four transformation modes. Each row has a pair of fused images before and after registration belonging to XY and YZ planes of the 3D data.

Appendix 6: Submitted manuscript - Radhika Sivaramakrishna, Kimerly A. Powell, Michael L. Lieber, William A. Chilcote and Raj Shekhar, "Texture analysis of lesions in breast ultrasound images," submitted to Journal of Digital Imaging.

Texture analysis of lesions in breast ultrasound images

Radhika Sivaramakrishna*, Kimerly A. Powell*, Michael L. Lieber**, William A. Chilcote***, Raj Shekhar*

*Agnes Christine Roberts Breast Imaging Laboratory, Department of Biomedical Engineering
**Department of Biostatistics and Epidemiology
***Breast Center, Department of Radiology
9500 Euclid Avenue, The Cleveland Clinic Foundation, Cleveland OH 44195

***Acknowledgment of Support*:**

This research was supported by the Department of Defense Grant DAMD17-99-1-9034.

Author Information:

Radhika Sivaramakrishna, Ph.D., Project Staff, Department of Biomedical Engineering, The Cleveland Clinic Foundation.

Kimerly A. Powell, Ph.D., Assistant Staff, Department of Biomedical Engineering, The Cleveland Clinic Foundation.

Michael L. Lieber, M.S., Biostatistician, Department of Biostatistics and Epidemiology, The Cleveland Clinic Foundation.

William A. Chilcote, M.D., Staff Physician, Breast Center, Department of Radiology, The Cleveland Clinic Foundation.

Raj Shekhar, Ph.D., Project Staff, Department of Biomedical Engineering, The Cleveland Clinic Foundation.

Address for correspondence:

Raj Shekhar, Ph.D.
Department of Biomedical Engineering - ND20
Lerner Research Institute
The Cleveland Clinic Foundation
9500 Euclid Avenue, Cleveland OH 44195
Telephone: (216) 445-3246
Fax: (216) 444-9198
Email: shekhar@bme.ri.ccf.org

Texture analysis of lesions in breast ultrasound images

ABSTRACT

The focus of our research was to investigate the use of Haralick's texture features and posterior acoustic attenuation descriptors extracted from two-dimensional (2D) breast ultrasound (US) images for the characterization of breast masses as either cysts or benign or malignant solid masses. The study database consisted of 71 breast US images contained 24 cyst, 21 benign solid mass and 26 malignant solid mass cases confirmed by biopsy. All lesions were manually segmented on the images. 28 Haralick's descriptors were evaluated at 5 different neighborhood sizes to obtain a total of 140 descriptors. Two posterior acoustic attenuation descriptors were also evaluated. Stepwise logistic regression was used to determine the best subset of descriptors at each neighborhood. Separate models were constructed for distinguishing cysts from noncysts and benign solid masses from malignant lesions. C-statistics were used to compare models employing varying neighborhood sizes. In the task of differentiating cysts from noncysts, the best model was the one at a neighborhood of four pixels, and contained *Mean of Sum Average* and *Range of Sum Entropy* Haralick's descriptors and the 2nd posterior acoustic attenuation descriptor with a c-statistic value of 0.954. For separating benign and malignant solid masses, the best model was the one at a neighborhood of 16 pixels, and contained the *Range of Correlation* Haralick's descriptor and the 2nd posterior acoustic attenuation descriptor with a c-statistic value of 0.886. In conclusion, computerized analysis of US images has the potential to increase the specificity of breast sonography.

Keywords: Breast cancer, Breast ultrasound, Breast lesion discrimination, Texture analysis, Image analysis

INTRODUCTION

Breast cancer is the most common cancer found in women today. In the year 2001, it is estimated that there will be approximately 192,000 new cases of invasive breast cancer and about 40,600 deaths in the United States.[1] Breast cancer is most effectively treated when detected at an early stage,[2] and x-ray mammography is currently the primary imaging technique for the detection and diagnosis of breast lesions.[3] However, mammographers miss about 10% of all cancers, especially those in dense breasts.[4] It is estimated that approximately two-thirds of these missed cancers are detected retrospectively by radiologists.[5] In addition, about two-thirds of lesions sent to biopsy turn out to be benign. This high miss rate and low specificity is due to the low conspicuity of mammographic lesions and the noisy nature of the images, as well as the overlying and underlying structures that obscure features of interest in the projection radiographs.

Breast sonography is an important adjunct to diagnostic mammography and has primarily been used to distinguish between mammographically identified cystic and solid masses. The accuracy rate of breast ultrasound (US) has been reported to be 96% to 100% in the diagnosis of simple benign cysts, and lesions so characterized do not require further evaluation.[6] However, breast ultrasound has not been widely used for the characterization of benign from malignant masses seen on ultrasound due to the considerable overlap in the sonographic appearance of these masses. Thus, when a palpable or mammographically suspicious mass cannot be ruled out as a cyst in an ultrasound examination, a biopsy is usually ordered. Patient trauma and the cost of these unnecessary surgical procedures have prompted many researchers to investigate the characterization of solid breast masses as benign or malignant. Several sonographic features have emerged as potential indicators of malignancy and others as potential indicators of benign

masses.[7,8] Benign features include hyperechogenicity, ellipsoidal shape, mild lobulation, and a thin, echogenic pseudocapsule. Malignant features include spiculation, angular margins, marked hypoechogenicity, posterior acoustic shadowing, and a depth-to-width ratio greater than 0.8.

By combining several ultrasonic characteristics, Stavros *et al.*[9] achieved a specificity of 98.4% and a sensitivity of 68.7% on a dataset of 750 solid breast nodules. However, the sonographic evaluation used is more complex than what is traditionally performed at most breast imaging centers. Doppler US evaluation of breast masses has also been reported with promising results.[10] However, color Doppler US imaging is not traditionally used for breast lesions. Moreover, since Doppler US imaging is based on the vascularity of the lesion, and several benign lesions demonstrate vascularity, this technique is inherently limited.

Computer-aided characterization of breast masses has been reported in the literature. Giger *et al.*[11] used features related to lesion margin, shape, homogeneity, and posterior acoustic attenuation to distinguish between benign and malignant US lesions. Accurate identification of lesion margin and shape is generally difficult in ultrasound images.

Sahiner *et al.*[12] used spatial gray level dependence features on three-dimensional (3D) breast US images to characterize between benign and malignant lesions. However 3D breast US is not traditionally used in most breast imaging centers.

The focus of the research presented in this paper was to investigate the use of Haralick's texture features and posterior acoustic attenuation descriptors, features describing attenuation posterior to the lesion, extracted from 2D breast US images for the characterization of breast masses as either cysts or benign or malignant solid lesions.

MATERIALS AND METHODS

A. STUDY SAMPLE AND CASE SELECTION

Direct-digitally acquired US images of regions of interest containing mammographically identified lesions were used for this study. These US images were from patient studies performed at the Breast Center within the Division of Radiology at The Cleveland Clinic Foundation. Only one US image per patient was selected, and only those images that did not contain overlaid cursors and in which pathology was clearly available were used for the study. The final data set comprised a total of 71 cases (24 cysts, 21 benign solid masses and 26 malignant solid masses). All lesions were manually segmented on the images by a trained observer. Typical ultrasound images corresponding to each of the three cases are shown in Figure 1.

B. TEXTURE DESCRIPTORS

28 Haralick's descriptors for five different pixel-pair distances (1, 2, 4, 8, 16 pixels) for spatial gray-level dependence (SGLD) matrix computation and two posterior acoustic attenuation descriptors were calculated for each image. Haralick's texture features are based on calculating the SGLD matrix that characterizes the spatial distribution of gray levels in the region of interest (ROI).[13] An element at location (i, j) of the SGLD matrix signifies the joint probability density of the occurrence of gray levels i and j in a specified orientation θ and specified distance d from each other. Thus for different θ and d values, different SGLD matrices result. Various features can be derived from these SGLD matrices. Usually θ is restricted to values of 0°, 45°, 90° and 135°, and d is restricted to integral multiples of pixel size. For every (θ, d) pair, 14 features can be derived from the SGLD matrix. For a fixed d, four values are obtained for each feature corresponding to the four values of θ. The mean and range of these four values, comprising a

total of 28 features are then calculated for the 14 texture features that are used for texture analysis. Since the dimension of the SGLD matrix is define by the number of distinct possible gray levels in the image, quantization is performed to a manageable number of gray levels to reduce the size of the SGLD matrix and computation in turn. This quantization is specified in terms of the number of bits used in the calculation. In the case of US images, the inherent gray level resolution is 8 bits, and hence this full resolution was used for the evaluation of Haralick's features. The 28 Haralick's descriptors were calculated at five different neighborhood sizes (1, 2, 4, 8 and 16 pixels) to make a total of 140 Haralick's descriptors.

Benign lesions are often associated with posterior enhancement, malignant lesions with posterior shadowing and simple cysts with relative posterior hyperechogenecity.[7,8] Therefore, two posterior acoustic attenuation descriptors were also calculated for each image. The first descriptor was determined as the difference between the average gray level within the lesion ROI with the average gray level in a 32 x 32 pixel region posterior to the lesion ROI. The second descriptor was determined as the difference between the average gray level in the 32 x 32 pixel region posterior to the lesion ROI and the average of the average gray levels in the two adjacent 32 x 32 pixel regions on the left and right of this posterior region.

C. STATISTICAL ANALYSIS

Stepwise logistic regression was used to determine the best subset of predictors (from the 28 Haralick's descriptors and two attenuation descriptors) for each neighborhood. Separate models were constructed for (1) distinguishing cysts from noncysts (solid benigns and malignants) (n = 71) and (2) solid benigns from malignants (n = 47). C-statistics were used to compare models among neighborhoods. The c-statistic is a nonparametric estimate of the area under the receiver operating characteristic (ROC) curve. The best model for each neighborhood

was evaluated in the form of logit(p) values. From these logit(p) values, the specified predicted probabilities for a single unknown case could be obtained from *logit(p)* as $p = \frac{exp(logit(p)}{(1+exp(logit(p))}$. In the first task of differentiating cysts from noncysts, the model estimated the probability that the case was a noncyst; a higher probability using the equation indicated a greater chance of the case being a noncyst. In the second task of differentiating solid benigns from malignants, the model estimated the probability that the case was benign. Therefore, the lower the probability, the greater was the chance of the case being malignant. The probability of the case being malignant could be obtained as one minus the estimated probability of its being benign.

RESULTS

Tables I and II give the best model in terms of *logit(p)* equations for the two tasks of differentiating cysts from noncysts and solid benigns from malignants, respectively. The 140 Haralick's descriptor values have been given variable names X1 to X140 and the two attenuation descriptors have been given variable names X141 and X142. The actual Haralick descriptors and attenuation descriptor in the models have been enumerated below each table.

Based on c-statistics values, in the task of differentiating cysts from noncysts, the best model was the one using a neighborhood size of 4 pixels, and included the Haralick's descriptors *Mean of Sum Average* and *Range of Sum Entropy* and the 2nd posterior acoustic attenuation descriptor. The c-statistic of 0.954 suggests that, if we chose at random a cyst case and a noncyst case, 95.4 % of the time, the model would assign a higher probability of the instance being a cyst to the "actual" cyst case than to the noncyst case.

In the task of differentiating solid benigns from malignants, the best model was the one using a neighborhood size of 16 pixels and included the Haralick's descriptor *Range of Correlation* and the 2nd posterior acoustic attenuation descriptor. The c-statistic of 0.886 indicates that, if we chose at random a malignant case and a solid benign case, 88.6 % of the time the model would assign a higher probability of the instance being benign to the actual benign case than to the malignant case.

DISCUSSION

Breast sonography is an important adjunct to diagnostic mammography. However, it has primarily been used to resolve cysts from noncysts. In this paper, we have reported the feasibility of using computer-aided analysis using textural descriptors to characterize an unknown mammographically identified lesion viewed on breast US. Haralick's descriptors and posterior acoustic attenuation descriptors were used to develop discriminatory models to first identify cysts from noncysts, and within noncysts, to separate benign from malignant lesions. In our study, all of Haralick's descriptors were considered for a wide range of neighborhood sizes, and the best Haralick's descriptors (along with posterior acoustic attenuation descriptors) were identified based on stepwise logistic regression.

The models in our study were developed from individual patients, i.e., no multiple lesions from the same patient were used. Moreover, we have taken care to see that the maximum number of descriptors making up each model was in proportion to the number of cases used in order to avoid the problem of overmodeling and to improve the chances of building generalizable models.

Although the use of shape and margin descriptors[11] would improve the performance of these models, computing such descriptors would require accurate delineation of the mass

background that is a challenging task on a breast US. Hence, we have focused on surface descriptors that are less sensitive to accurate boundary identification.

Computerized analysis allows for an objective assessment of posterior acoustic shadowing, which is an important factor in both the discrimination of cysts from noncysts, and solid benigns from malignants. However, in breast US patient examinations, frequently, overall gain adjustments are performed by the technologist on a per-case basis to achieve the best possible visual image. Since shadowing depends to some extent on these gain settings, these parameters should be set more automatically than they are at present.

Future studies will consist of evaluating and validating the models developed in our research on unknown cases. We will also explore other surface descriptors like the Laws descriptors[14], fractal descriptors[15] and geometric surface roughness descriptors[16].

In conclusion, computerized analysis of breast US images has the potential to increase the specificity of breast sonography. The models developed in this study could be used as part of computer-aided-diagnosis system to characterize unknown breast US lesions.

REFERENCES

1. American Cancer Society: Cancer facts and figures, 2001.

2. Sivaramakrishna R, Gordon R: Detection of breast cancer at a smaller size can reduce the likelihood of metastatic spread. A quantitative analysis. Acad. Radiol. 4:8-12, 1997.

3. Moskowitz M: Breast imaging. Philadelphia: W. B. Saunders, 1995.

4. Jackson VP, Hendrick RE, Feig SA, et al: Imaging of the radiographically dense breast. Radiology 188:297-301, 1993.

5. Giger ML: Computer-aided diagnosis. *In* RSNA Categorical Course Phys, 1993, pp 283-298.

6. Jackson VP. The role of US in breast imaging. Radiology 177:305-311, 1990.

7. Tohno E, D.O C, Sloane JP: Ultrasound diagnosis of breast diseases. Edinburgh, Scotland: Churchill, Livingstone, 1994, pp 50-73.

8. Fornage BD, Lorigan JG, Andry E: Fibroadenoma of the breast: sonographic appearance. Radiology 172:671-675, 1989.

9. Stavros A, Thickman D, Rapp C, et al: Solid breast nodules: use of sonography to distinguish between benign and malignant lesions. Radiology 196:123-134, 1995.

10. Huber S, Delorme S, Knopp M: Breast tumors: computer-assisted quantitative assessment with color Doppler US. Radiology 192:797-801, 1994.

11. Giger ML, Al-Hallaq H, Huo Z, et al: Computerized analysis of lesions in US images of the breast. Academic Radiology 6:665-74, 1999.

12. Sahiner B, LeCarpentier GL, Chan H-P, et al: Computerized characterization of breast masses using three-dimensional ultrasound images. Proc SPIE Image Processing 3338, 1998.

13. Haralick RM, Shanmugam K, Dinstein I: Textural features for image classification. IEEE Trans. System Man Cybernetics SMC-3:610-621, 1973.

14. Laws K: Rapid texture identification. Proc. SPIE-IPMG 238:376-380, 1980.

15. Barnesly M: Fractals Everywhere. Toronto: Academic Press, 1988.

16. Mann JA, Rains EM, Woyezynski WA: Measuring the roughness of interfaces. Chemometrics Intel. Lab. Systems 12:169-180, 1991.

(a) Cyst (b) Benign solid mass (c) Malignant solid mass

Figure 1. Typical ultrasound images corresponding to each of the three categories studied. The suspected lesion is the darker closed area in each image.

Table I: Distinguishing cysts from noncysts

Neighborhood	Best model (as chosen by stepwise selection)	C-statistic
1	logit(p) = 2.44 + 0.07*(X7) – 19.07*(X26) – 0.08*(X142)	0.943
2	logit(p) = 2.80 + 0.05*(X35) – 11.38*(X45) – 0.08*(X142)	0.943
4	**logit(p) = 1.33 + 0.07*(X63) – 70.81*(X79) – 0.08*(X142)**	**0.954**
8	logit(p) = -0.30 + 0.05*(X91) – 0.07*(X142)	0.931
16	logit(p) = -0.71 + 7.11*(X115) + 0.05*(X119) – 0.08*(X142)	0.944

X7: Mean of Sum Average (neighborhood: 1)

X26: Range of Difference Entropy (neighborhood: 1)

X35: Mean of Sum Average (neighborhood: 2)

X45: Range of Correlation (neighborhood: 2)

X63: Mean of Sum Average (neighborhood: 4)

X79: Range of Sum Entropy (neighborhood: 4)

X91: Mean of Sum Average (neighborhood: 8)

X115: Mean of Sum Average (neighborhood: 16)

X119: Mean of Difference Variance (neighborhood: 16)

X142: 2nd Posterior Acoustic Attenuation Descriptor

Table II: Distinguishing malignants from benigns

Neighborhood	Best model (as chosen by stepwise selection)	C-statistic
1	logit(p) = -1.16 + 0.01*(X22) + 0.05*(X142)	0.838
2	logit(p) = 2.64 – 20.48*(X56) + 0.07*(X142)	0.853
4	logit(p) = -3.62 + 8.36*(X70) + 0.07*(X142)	0.848
8	logit(p) = -2.29 + 55.47*(X110) + 0.05*(X142)	0.855
16	**logit(p) = -2.32 + 11.33*(X129) + 0.07*(X142)**	**0.886**

X22: Range of Sum Variance (neighborhood: 1)

X56: Range of Information Measure of Correlation 2 (neighborhood: 2)

X70: Mean of Information Measure of Correlation 2 (neighborhood: 4)

X110: Range of Difference Entropy (neighborhood: 8)

X129: Range of Correlation (neighborhood: 16)

X142: 2nd Posterior Acoustic Attenuation Descriptor

DEPARTMENT OF THE ARMY
US ARMY MEDICAL RESEARCH AND MATERIEL COMMAND
504 SCOTT STREET
FORT DETRICK, MARYLAND 21702-5012

REPLY TO
ATTENTION OF:

MCMR-RMI-S (70-1y) 28 July 03

MEMORANDUM FOR Administrator, Defense Technical Information Center (DTIC-OCA), 8725 John J. Kingman Road, Fort Belvoir, VA 22060-6218

SUBJECT: Request Change in Distribution Statement

1. The U.S. Army Medical Research and Materiel Command has reexamined the need for the limitation assigned to technical reports written for this Command. Request the limited distribution statement for the enclosed accession numbers be changed to "Approved for public release; distribution unlimited." These reports should be released to the National Technical Information Service.

2. Point of contact for this request is Ms. Kristin Morrow at DSN 343-7327 or by e-mail at Kristin.Morrow@det.amedd.army.mil.

FOR THE COMMANDER:

Encl

PHYLIS M. RINEHART
Deputy Chief of Staff for
Information Management

ADB233865
ADB265530
ADB244706
ADB285843
ADB240902
ADB264038
ADB285885
ADB274458
ADB285735
ADB286597
ADB285707
ADB274521
ADB259955
ADB274793
ADB285914
ADB260288
ADB254419
ADB282347
ADB286860
ADB262052
ADB286348
ADB264839
ADB275123
ADB286590
ADB264002
ADB281670
ADB281622
ADB263720
ADB285876
ADB262660
ADB282191
ADB283518
ADB285797
ADB269339
ADB264584
ADB282777
ADB286185
ADB262261
ADB282896
ADB286247
ADB286127
ADB274629
ADB284370
ADB264652
ADB281790
ADB286578
ADB264750
ADB282776
ADB286264
ADB260563
ADB277918
ADB286365
ADB275327
ADB286736
ADB286137
ADB286146
ADB286100
ADB286266
ADB286308
ADB285832