Ultrasonic Imaging & Signal Processing

Vol. 4687

PROGRESS IN BIOMEDICAL OPTICS AND IMAGING
Vol. 3, No. 25

Medical Imaging 2002

Ultrasonic Imaging and Signal Processing

Michael F. Insana
William F. Walker
Chairs/Editors

26–28 February 2002
San Diego, USA

Sponsored and Published by
SPIE—The International Society for Optical Engineering

Cooperating Organizations
AAPM—American Association of Physicists in Medicine
APS—American Physiological Society
FDA Center for Devices and Radiological Health (USA)
IS&T—The Society for Imaging Science and Technology
NEMA—National Electrical Manufacturers Association/Diagnostic Imaging and Therapy Systems Division (USA)
RSNA—Radiological Society of North America
SCAR—Society for Computer Applications in Radiology (USA)

Proceedings of SPIE
Volume 4687

Please use the following format to cite material from this book:

Author(s), "Title of paper," in *Medical Imaging 2002: Ultrasonic Imaging and Signal Processing*, Michael F. Insana, William F. Walker, Editors, Proceedings of SPIE Vol. 4687, page numbers (2002).

ISSN 1605-7422
ISBN 0-8194-4432-4

Published by
SPIE—The International Society for Optical Engineering
P.O. Box 10, Bellingham, Washington 98227-0010 USA
Telephone 1 360/676-3290 (Pacific Time) • Fax 1 360/647-1445
http://www.spie.org/

Printed in the United States of America.

Conference Committee

Conference Chairs

Michael F. Insana, University of California/Davis (USA)
William F. Walker, University of Virginia (USA)

Program Committee

Martin E. Anderson, University of Rochester (USA)
Stanislav Emelianov, University of Michigan (USA)
Ernest J. Feleppa, Riverside Research Institute (USA)
Aaron Fenster, John P. Robarts Research Institute (Canada)
and University of Western Ontario (Canada)
David H. Vilkomerson, EchoCath, Inc. (USA)

Session Chairs

1 Signal Processing and Detection
Michael F. Insana, University of California/Davis (USA)

2 Beamforming
William F. Walker, University of Virginia (USA)

3 Array Transducers
John A. Hossack, University of Virginia (USA)

4 RF Analysis
Stanislav Emelianov, University of Michigan (USA)

5 Tissue Elasticity
James F. Greenleaf, Mayo Clinic and Foundation (USA)

6 Doppler
David H. Vilkomerson, EchoCath, Inc. (USA)

7 Vascular Imaging
Jørgen A. Jensen, Technical University of Denmark

8 Segmentation/Classification
Ronald H. Silverman, Weill Medical College of Cornell University (USA)

Introduction

The main topic at this year's Ultrasonic Imaging and Signal Processing conference was *signal processing*. We received a large number of very exciting papers that formed a total of eight sessions over two days. We heard numerous talks on innovative approaches to synthetic beamforming, both on transmission and reception, and novel coded excitation schemes. It now seems possible to image with incredibly high frame rates without suffering from prohibitively low SNR. This capability opens up new opportunities for ultrasonic imaging of very fast physiological processes, such as small-animal echocardiography. In the context of the advances in high-dimensional and high-frequency arrays, signal processing clearly offers exciting new opportunities for biological imaging using ultrasound. Seven sessions offered papers on novel beamforming methods, array transducers, Doppler and vascular imaging, statistical and spectral analyses, and interesting approaches to motion estimation for blood velocity and elasticity imaging.

The conference began with a keynote presentation by Professor Simon Haykin from McMaster University on the application of artificial neural networks to medical imaging. Professor Haykin challenged us to design our signal processing strategies on the physics of the biological systems under investigation. He also suggested that we consider using support vector machines as a tool for feature discrimination, and briefly described its effects on feature dimensionality and detection performance assessment.

The conference closed with a session on segmentation and classification, which complemented papers in several sessions offered by the concurrent Image Processing Conference, SPIE Proceedings Vol. 4684 (2002).

The conference chairs gratefully acknowledge the contributions of the program committee. Their help reviewing abstracts and organizing and chairing sessions helped us provide a strong program. We also thank the SPIE staff for their organization and support.

Michael F. Insana
William F. Walker

Contents

POSTER SESSION

Comparison Between Different Encoding Schemes for Synthetic Aperture Imaging

Svetoslav I. Nikolov and Jørgen A. Jensen

Center for Fast Ultrasound Imaging, Ørsted•DTU, Bldg. 348,
Technical University of Denmark, DK-2800, Kgs. Lyngby, Denmark

ABSTRACT

Synthetic transmit aperture ultrasound (STAU) imaging can create images with as low as 2 emissions, making it attractive for 3D real-time imaging. Two are the major problems to be solved: (1) complexity of the hardware involved, and (2) poor image quality due to low signal to noise ratio (SNR). We have solved the first problem by building a scanner capable of acquiring data using STAU in real-time. The SNR is increased by using encoded signals, which make it possible to send more energy in the body, while preserving the spatial and contrast resolution.

The performance of temporal, spatial and spatio-temporal encoding was investigated. Experiments on wire phantom in water were carried out to quantify the gain from the different encodings. The gain in SNR using an FM modulated pulse is 12 dB.

The penetration depth of the images was studied using tissue mimicking phantom with frequency dependent attenuation of 0.5 dB/(cm MHz). The combination of spatial and temporal encoding have highest penetration depth. Images to a depth of 110 mm, can successfully be made with contrast resolution comparable to that of a linear array image.

The in-vivo scans show that the motion artifacts do not significantly influence the performance of the STAU.

Keywords: synthetic aperture, ultrasound, focusing, imaging, temporal coding, spatial coding, coding

1. INTRODUCTION

Synthetic Aperture Ultrasound imaging (SAU) has been studied for more than two decades now. Various advantages of SAU compared to the conventional ultrasound imaging were explored: (a) simpler electronic front end[1–4]; (b) better image quality[5,6]; (c) fast imaging, applicable for real-time three-dimensional scanning[7,8]; (d) estimation of low blood flow and high frame rate of color flow mapping.[9,10] In spite all of the investigations, such systems are still not in clinical use. Among the problems are: (a) the presence of motion artifacts,[3,4] and (b) a low Signal-to-Noise Ratio (SNR).[7,11]

We have previously shown how to compensate for the motion artifacts,[10,12] and how to avoid their effect in blood flow estimations.[9,10] Various coding schemes have been suggested to increase the SNR: temporal encoding,[13] spatial encoding,[14] and their combination.[15] Some of these methods have been investigated only in simulation, some tried *in-vivo*, but not with SAU. At the Center for Fast Ultrasound Imaging (CFU), a scanner capable of implementing any of the aforementioned encoding schemes for SAU in real time was developed.[16]

The purpose of this paper is to experimentally investigate the increase in SNR using the various spatial and temporal encoding schemes. The comparison is done on data measured on phantoms. The performance of the algorithms is also shown for different imaging situations *in-vivo*.

Further author information: (Send correspondence to Svetoslav I. Nikolov)
Svetoslav I. Nikolov: E-mail: sn@oersted.dtu.dk, Telephone: +45 45 25 37 05, Address: Ørsted•DTU, Bldg. 348, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

Medical Imaging 2002: Ultrasonic Imaging and Signal Processing, Michael F. Insana,
William F. Walker, Editors, Proceedings of SPIE Vol. 4687 (2002) © 2002 SPIE · 1605-7422/02/$15.00

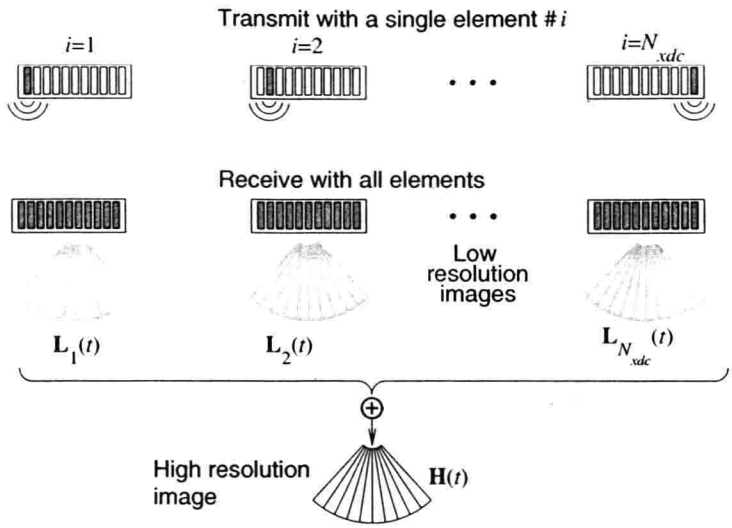


Figure 1: Synthetic transmit aperture imaging.

2. THEORY

In this section the theory behind the SAU imaging is presented. At the present state of technology, the system complexity is not of major concern, and the focus is put on Synthetic *Transmit* Aperture Ultrasound imaging (STAU).

2.1. Synthetic Transmit Aperture Imaging

Synthetic transmit aperture imaging (STAU) is done by sending out a waveform which propagates through all of the scanned region, and whose wavefront is known. Such waves are either spherical or plane. Although plane waves have been also used for STAU imaging,[17] the use of spherical waves has been more widely investigated. A nearly spherical* wave can be created by either transmitting with a single element, or by using multiple elements,[2,7,8] the delays of which are appropriately set. The point to which the origin of such a wave can be traced, will be called a "virtual" source element.[10] In the case that only a single element is used in transmit, the positions of the virtual and real source coincide.

Figure 1 illustrates the process of creating a synthetic aperture image, when only a single element is used in transmission. The wave created by a single element is spherical, and propagates through whole region of interest. The back scattered echoes carry information from all directions, and by applying different delays on the received signals a scan line can be formed in *any* direction. To speed up the acquisition, scan lines in *all* directions are formed, thus creating a whole image. Since the image is focused only in receive, it has a low resolution, hence the name *Low Resolution Image* (LRI). The beamforming of a single line $L_{il}(t)$ is:

$$L_{il}(t) = \sum_{j=1}^{N_{xdc}} a_{lij}(t) r_{ij}(\tau_{lij}(t)), \tag{1}$$

where $r_{ij}(t)$ is the Radio Frequency (RF) signal received by the element j , after transmitting with element i, and $\tau_{lij}(t)$ is the round-trip propagation time from element i to the current focal point and back to element j. It is a function of t, which is the time from the trigger of the emission. The apodization coefficient a_{lij} can also

*Strictly speaking the wave has a complex shape determined by the geometry of the transducer elements.

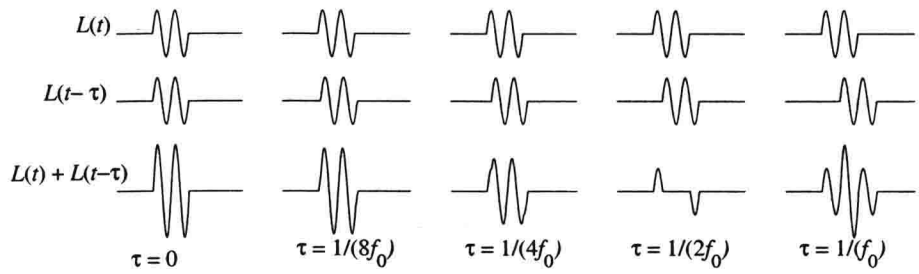


Figure 2: Illustration of several cases of motion artifacts

be a function of t. The subscript l is used to index the scan lines comprising the LRI. LRI in matrix form is written as:

$$\mathbf{L}_i(t) = [L_{i1}(t), L_{i2}(t), \cdots, L_{iN_l}(t)], \tag{2}$$

where N_l is the number of scan lines. The columns of the matrix correspond to different image directions, and the rows correspond to samples in depth. After transmitting with the first element $i = 1$, a second element is used in transmit $i = 2$. The RF signals in $\mathbf{L}_1(t)$ and $\mathbf{L}_2(2)$, although focused at the same points, have different phases due to the different origins of the sources. Summing them is equivalent to focusing in transmit. Because the focusing was done on every point, this is equivalent to dynamic transmit focusing. After using all of the elements across the aperture, a High Resolution Image (HRI) can be made by summing the low resolution ones:

$$\mathbf{H}(t) = \sum_{i=1}^{N_{xdc}} \mathbf{L}_i(t). \tag{3}$$

It is possible to reduce the number of emissions by using only some of the transducer elements.[7,8,14] The number of emissions will be further denoted by N_{xmt}. The resolution of $\mathbf{H}(t)$ is determined by the distance between the outermost elements, and the side- and grating-lobes level by the number of emissions N_{xmt}. The larger the number of emissions N_{xmt}, the lower the side lobe energy is. The acquisition time is, however, increased, and the image is more susceptible to motion artifacts, which are discussed in the next section.

2.2. Motion Artifacts

In order to create a HRI the low resolution images should be summed coherently. The motion of the scatterers between two transmissions prohibits this summation as shown in Fig. 2.2. From the figure it can be seen that there are certain tolerance intervals for the motion. The worst case is when the scatterers move at a distance $\lambda/2$ between every two emissions since that results in destructive interference. If the motion for the N_{xmt} emissions is less than $\lambda/2$, then the sum of the signals will still be larger than the individual RF traces, and the end image will be focused. Assuming a speed of sound $c = 1500$ m/s, center frequency $f_0 = 5$ MHz, a pulse repetition frequency of $f_{prf} = 5$ kHz, and $N_{xmt} = 60$, the maximum velocity for which the time shift is less than half a period is:

$$v_{max} < \frac{f_{prf}c}{4f_0N_{xmt}} = \frac{5 \cdot 10^3 \cdot 15 \cdot 10^2}{4 \cdot 5 \cdot 10^6 \cdot 60} = 6.3 \cdot 10^{-3} \text{ m/s} \tag{4}$$

The major causes of motion are the heart beat, pulsation, and breathing.[18] The motion, apart from the heart, is most pronounced at the walls of the blood vessels. Table 2.2 shows the maximum velocity v_{max} of the vessel walls for some of the blood vessels in the presence of different causes of motion. It can be seen that the peak velocities in the investigated regions are bigger than v_{max} found in (4). In many cases (such as scanning the carotid artery) either a higher f_{prf} or a lower f_0 will be used. The number of emissions N_{xmt} can also be reduced.

Vessel	Scan plane	Motion	v_{max}
Carotid artery	Transverse, scan angle 90°	P, B	$8.9 \cdot 10^{-3}$ m/s
Hepatic vein	Right liver lobe, intercostal scan	B, H (P)	$6.2 \cdot 10^{-3}$ m/s
Hepatic vein	Right liver lobe, intercostal scan	H (P)	$4.2 \cdot 10^{-3}$ m/s
Hepatic vein	Left liver lobe, epigastric scan	H (P)	$10.1 \cdot 10^{-3}$ m/s

Table 1. Maximum velocities of the vessel walls due to pulsation (P), heart beat (H), and breathing (B). Data published by Schlaikjer et. al.[18]

2.3. Signal to Noise Ratio

The SNR is one of the major problems for STAU imaging. In receive all of the transducer elements are used, and the gain in SNR (GSNR), of the beamformed RF line due to the receive focusing is the same as in "conventional" ultrasound imaging. The noise is assumed to be white, and the SNR is increased $\sqrt{N_{rcv}}$ times, where N_{rcv} is number of receiving elements. The degraded SNR arises from using few elements in the transmission. Many authors[2,7,11] suggest the use of multiple elements in transmit, to create a spherical wave. Usually the number of active elements rarely exceeds $N_{act} = 30$. The wave is divergent and its amplitude falls off with $1/r$, even in non-attenuating medium, (r is the traveled distance). In conventional ultrasound scanning, however, a focused beam is sent into the tissue, and a much larger in amplitude signal is returned, even for the same number of transmitting elements. It is clear that to obtain the same SNR of the signal on a single channel, as many as possible of the transducer elements must be used in transmit, and each of them should send a long pulse. To maintain the lateral and axial resolutions, spatial and temporal coding must be used, respectively.

2.3.1. Temporal encoding

Usually matched filtering is applied on the received signal. The peak of the signal returned by a point scatterer is determined by the maximum of the autocorrelation function of the transmitted signal, which is equal to the energy in the signal. The maximum amplitude of the transmitted signal is limited for safety reasons, and one way to increase the transmitted energy is to use either a long pulse, or a sequence of pulses. Using a long pulse, however, decreases the axial resolution, unless some form of modulation is used. The linear frequency modulated (FM) pulse (chirp), is especially suited for ultrasound imaging, since its compression properties are retained in an attenuating media.[13]

One such chirp, and the result of the filtration with a windowed matched filter, is shown in Fig. 3(a). The rising and the falling edges of the chirp are weighted in order to reduce the range side lobes. The matched filter is weighted even more. The weighting of this filter is done in such a way, as to smoothen the spectrum of the received signal.[13]

The gain in SNR (GSNR) due to mathed filtering is:

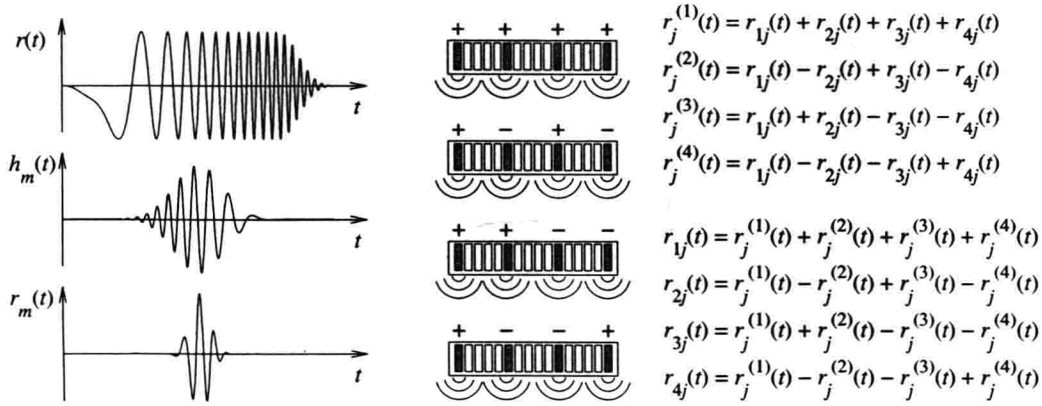
$$GSNR = \frac{SNR_{out}}{SNR_{in}} = TB. \tag{5}$$

The pulse that will be used for the experiments has a 3 cycles, Hamming weighted pulse at 7 MHz. The chirp used in the experiments has a duration $T = 20$ μs and a bandwidth of 7 MHz. The expected gain in signal-to-noise ratio is:

$$GSNR_{chirp} - GSNR_{sin} = 10(\log_{10}(20 \times 7) - \log_{10}(0.45 \times 10)) = 14.9 \text{ dB}. \tag{6}$$

However, the weighting applied on the chirp will reduce the GSNR.

Another class of temporally encoded signals, which are suitable for ultrasound imaging, is the class of binary encoded waveforms such as the Golay codes.[15] The use of Golay pairs, however, requires 2 emissions from every element position, thus increasing the acquisition time.



(a) From top to bottom: the received signal, the compression filter, and the result of the compression.

(b) The apodization in transmit is given with '+' and '-'. $r_j^{(n)}(t)$ is the signal received by element j at emission n. $r_{ij}(t)$ is the part of the signal that is transmitted by element i and received by element j.

Figure 3: Illustration of (a) temporal, and (b) spatial encoding.

2.3.2. Spatial encoding

In order to increase the SNR, as many as possible of the transducer elements must be used in transmit. The idea is to transmit with all of the virtual sources at the same time instead of transmitting with a single virtual source at a time, as shown in Fig. 3(b). The signal received by element j is a linear combination of the signals $r_{ij}(t)$ that would have been received by j, if the transmitting elements i had transmitted one by one. At every emission, the apodization coefficients in transmit are changed. A system of linear equations is made, and when solved the individual components $r_{ij}(t)$ can be found from its solution. In matrix form the signal received by element j at emission n, $n \in [1, N_{xmt}]$ is:

$$\begin{aligned} \vec{r}_j(t) &= \mathbf{Q}\vec{r}_{ij}(t) \\ \vec{r}_{ij}(t) &= \mathbf{Q}^{-1}\vec{r}_j(t). \end{aligned} \tag{7}$$

$\mathbf{Q}$ is the encoding matrix, which consists from the transmit apodization coefficients. The rows of $\mathbf{Q}$ correspond to the emission number n, and the columns to the index of transmitting element i. A suitable encoding matrix $\mathbf{Q}$ is the Hadamard matrix $\mathbf{H}$.[14] The Hadamard matrix is given by its order N, and the lowest order is $N = 2$. The matrix of order 2 is:

$$\mathbf{H}_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \tag{8}$$

The order of the matrix can be only an even number, and the matrix of order $2N$ is found recursively from the matrix of order N by:

$$\mathbf{H}_{2N} = \begin{bmatrix} \mathbf{H}_N & \mathbf{H}_N \\ \mathbf{H}_N & -\mathbf{H}_N \end{bmatrix} \tag{9}$$

The inverse matrix is the matrix itself:

$$\mathbf{H}_N^{-1} = \frac{1}{N}\mathbf{H}_N. \tag{10}$$

The process of coding and decoding using Hadamard matrices for $N = 4$ is given in Fig. 3(b).

2.3.3. Spatial and temporal encoding

There are two types of combinations of spatial and temporal encoding. The first is to use spatial encoding, and instead of a short pulse to use a long linear FM chirp. The process of beam formation is preceded by two pre-processing stages: (1) spatial decoding, and (2) pulse compression.

Parameter name	Notation	Value	Unit
Number of receive elements	N_{rcv}	64	-
Mean frequency	f_0	7.0	MHz
Sampling frequency	f_s	40	MHz
Resolution of A/D converter	-	12	bits
Transducer pitch	d_x	208	μm
Element width	w	173	μm
Element height (elevation)	h	4.5	mm
Elevation focus	F	25	mm
Fractional bandwidth	BW	60	%

Table 2: Parameters of the measurement system.

The second combination is to use spatial encoding in combination with orthogonal temporal signals.[15] This is used in order to decrease the number of emissions. For example, if one has two orthogonal signals $A(t)$ and $B(t)$, then the transmissions in the case of $N_{xmt} = 4$ is:

$$\begin{array}{cccc} A_1(t) & A_2(t) & B_3(t) & B_4(t) \\ A_1(t) & -A_2(t) & B_3(t) & -B_4(t) \end{array}$$

In the first decoding stage, the combination of signals $A_1(t) + B_3(t)$, and $A_2(t) + B_4(t)$ are found by adding and subtracting the received signals. In the second decoding stage $A_1(t)$ and $A_2(t)$ are separated from $B_3(t)$ and B_4 by using cross-correlation (A is orthogonal to B).

3. RESULTS

The various strategies will now be evaluated- Three sets of measurements were done: (a) on a wire, (b) on a phantom. and (c) in-vivo. The measurements on the wire were used to characterize the parameters of the system such as resolution, integrated side lobe to main lobe ratio (ISLMLR), and peak signal to noise ratio (PSNR). The measurements on the phantom are used to demonstrate the improved penetration depth with the use of temporal and spatial encoding, and the measurements in-vivo are used to show that the motion artifacts do not distort the image.

3.1. Experimental Setup

The measurements were done using the experimental system RASMUS,[16] developed at the Center for Fast Ultrasound Imaging. The system has 128 transmit channels and 64 receive channels which can be multiplexed to 128 transducer elements. Up to 4096 different transmit waveforms per channel can be set up in transmit. Each waveform can contain up-to 4096 12 bit samples. The receiver has 128 MB focusing look up tables per channel, and any delay can be set up. Additionally, 16 GB of 12 bit data sampled at 40 MHz can be stored for all the channels, for off-line processing. The transmitting and the receiving units are fully programmable and controlled by software. The parameters of the system are given in Table 2.

3.2. Measurements on a Wire Phantom

In order to characterize the system, a wire phantom was scanned. It consisted of 5 wires, placed along a line at 60° to the surface of the transducer. The width of the wires was 0.25 mm. B-mode images of the wire phantom are shown in Fig. 4. In the rest of the figures and tables, the abbreviation "chirp_" will be used to denote that a linear FM modulated pulse was used in transmit, and "sin_" that the used excitation is a conventional RF pulse. The prefix "had_" will show that the imaging is combined with spatially encoded transmit. The numbers added to the titles of the plots will show the number of transmit events used to form a single image. Acquisitions with 4, 8, 16, 32, and 64 firings were done for each of the transmit schemes.

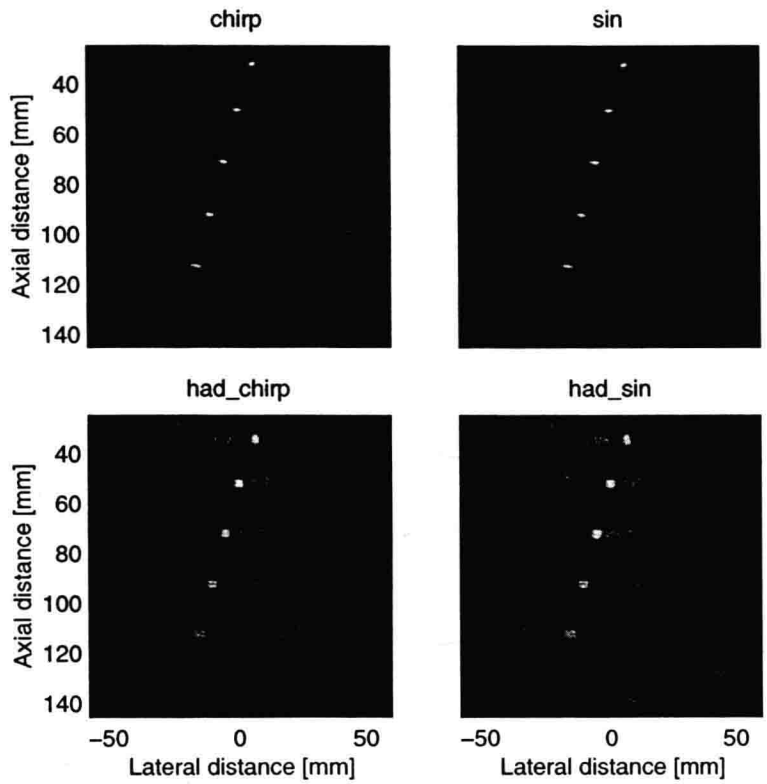


Figure 4: B-mode images of the wire phantom used to characterize the point spread function of the system.

The same FM pulse was used for all of the experiments. The duration of the FM pulse is $T = 20$ μs, and a bandwidth $B = 7$ MHz. The mean frequency of the chirp is 7.0 MHz. The rising and the falling edges of the chirp are weighted in order to suppress the range side lobes,[19,20] which in this case are lower than -55 dB from the peak. The RF pulse that is used for the cases designated with "sin", is a 3 cycles RF pulse at 7 MHz. The envelope of the pulse is weighted with a Hamming window.

For the cases, in which there was no spatial encoding, 25 elements were used to emulate the radiation pattern of a single element in transmit. For the cases with spatial Hadamard encoding, the number of transmit elements used to emulate the radiation pattern of a single one was: 15, 7, 3, 1, 1, for N_{xmt} equal to 4, 8, 16, 32, and 64, respectively.

All of the images in Fig. 4 have a dynamic range of 60 dB, and were acquired with 64 emissions. It can be seen that the image with least imaging artifacts is the image obtained without any kind of encoding. This fact is confirmed by looking at the plots in Fig. 5, and at Table 3.

The lower amount of the artifacts for the case without *any* encoding is due to the lack of range side lobe levels, which are present when FM modulated chirps are used. An indication for this is the slightly lower IMLSLR (higher side lobe energy) when chirps are used. The IMLSLR is lowest for the scans employing Hadamard encoding. In this experiment the use of Hadamard encoding worsened the performance of the system, contrary to all expectations. The only reasonable explanation for this, is that the propagation in the water introduced non-linearities in the response, and in this way the components of the signal could not be separated properly.