Synthetic aperture tissue and flow ultrasound imaging

Svetoslav Ivanov Nikolov

August 31st, 2001

Center for Fast Ultrasound Imaging
Technical University of Denmark

SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
AT
THE TECHNICAL UNIVERSITY OF DENMARK
AUGUST 2014

Signature of Author

To my parents

Contents

Abstract

This Ph.D. project was carried out at the Center for Fast Ultrasound Imaging, Technical University of Denmark. The goal was to improve existing imaging techniques in order to make them suitable for real-time three-dimensional ultrasound scanning. This dissertation focuses on the synthetic aperture imaging applied to medical ultrasound. It is divided into two major parts: tissue and blood flow imaging.

Tissue imaging using synthetic aperture algorithms has been investigated for about two decades, but has not been implemented in medical scanners yet. Among the other reasons, the conventional scanning and beamformation methods are adequate for the imaging modalities in clinical use - the B-mode imaging of tissue structures, and the color mapping of blood flow. The acquisition time, however, is too long, and these methods fail to perform *real-time* three-dimensional scans. The synthetic transmit aperture, on the other hand, can create a B-mode image with as little as 2 emissions, thus significantly speeding-up the scan procedure.

The first part of the dissertation describes the synthetic aperture tissue imaging. It starts with an overview of the efforts previously made by other research groups. A classification of the existing methods is made, and a new imaging technique, the "recursive ultrasound imaging" is suggested. The technique makes it possible to create a new image after every emission. This opens further the possibility for visualizing the blood flow. Various aspects of the scan procedure are considered, among them: the use of sparse one- and two-dimensional arrays; the use of multiple elements in transmit to create virtual sources of ultrasound; the use of virtual sources of ultrasound to improve the resolution of the images in the elevation plane; the use of temporal and spatial encoding to increase the signal to noise ratio. In many of the mentioned areas, the author presents the existing state of the art, and adds his personal contributions.

The second part describes *blood flow* estimation using synthetic aperture techniques. It starts by introducing the velocity estimator based on the time shift measurement of the received signals. This estimator fails to estimate the velocity when applied on the radio frequency signals formed by synthetic aperture techniques. The failure is caused by the motion artifacts, and the second part continues by developing a new model for them. Based on this model a novel motion compensation scheme is presented. The velocity can successfully be estimated from the motion compensated images. The standard deviation and the bias are both within 2 %. The estimation of blood flow using synthetic transmit aperture ultrasound is further extended by developing a scheme of how to modify the existing blood flow estimators. In the new approach images n and $n+N$, $n+1$ and $n+N+1$ are cross correlated, where N is the number of emissions for one image. These images experience the same phase distortion due to motion and therefore have a high correlation without motion compensation. The estimate of the cross-correlation

is improved by averaging the estimates obtained from the pairs of frames $[n, n+N]$, $[n+1, n+1+N]$, and so on up to $[n+N-1, n+2N-1]$. The advantage of the approach is that a color flow map can be created for all directions in the image simultaneously at every emission, which makes it possible to average over a large number of lines. This makes stationary echo canceling easier and significantly improves the velocity estimates. Only 8 emissions per plane are necessary to create the color flow map. Scanning 12 cm in depth, up to 800 planes can be obtained, making it possible for real-time three-dimensional tissue and blood-flow imaging.

Acknowledgements

First of all I would like to thank to my advisor Professor Jørgen Arendt Jensen for making this project possible, and for his attention and guidance through the years of my Ph.D. study. I was truly fortunate to have the opportunity to work under his advice. His knowledge and experience helped me to determine which path of research was worth pursuing, and which not, thus giving me the confidence that all my efforts were doomed to success. I am also grateful for his idea for recursive ultrasound imaging, which proved to be a very fruitful field. On a personal plan I have learnt from him a lot - how to set my priorities, how to organize my time, and how to carry out independently my research. Last, but not least, I am thankful for introducing me to world of Linux and the GNU software, of which I became a passionate advocate.

I would like to express my appreciation to *all* my colleagues at the Biomedical Engineering Group at Ørsted•DTU for the good working atmosphere, for the high level of communication, and for the free flow of ideas here. In particular I would like to thank:

- Peter Munk for his large knowledge of ultrasound and signal processing, for the numerous pleasant, fruitful and inspiring conversations about science, life and the universe, for his useful comments and advice on the contents and organization of the thesis. Without him and his help in everyday matters, my life in Denmark would not have been the same.
- Malene Schlaikjer for proof reading this thesis, for many useful comments on papers and presentations, for helping me with many issues related to my life in Denmark, and simply for being a good and helpful colleague.
- Jens E. Wilhjelm, for giving the XTRA system at my disposal, for showing and helping me how to use the positioning system, for many useful comments on the papers and presentations I had during the years.
- Thanassis Misaridis for the set of scripts designing optimized pulse compression filters, for discussions on coded excitations, and other technical and non-technical matters, for useful comments on my papers, and for being a good and pleasant office neighbor.
- Borislav Tomov for proof reading several chapters of the thesis, for helping to debug parts of the system software for RASMUS, for staying up-to-date with the latest software releases, for being a good officemate and friend.
- Kaj-Åge Henneberg for showing me how to use LaTeX notations in XFig and some other useful “tricks”, and for the helpful comments on my presentations.

- Henrik Pedersen for his help in developing the software for RASMUS, and for helpful comment on presentations and papers.

- Morten Høgholm Pedersen for helping me to acquire some in-vivo data.

I would like to thank Dimitar Lazarov. Without his encouragement and support I would have never started this project. His work with 3D computer graphics inspired me to use OpenGL for scan-conversion, for which I am also grateful.

During the years at DTU I had the pleasure to work with some very good students, to some of which I would like to thank:

- Juan Pablo Gómez Gonzaléz for his hard and excellent work on the 3D visualization program.

- Kim Gammelmark for the great job of calibrating the XTRA system and for acquiring some of the data I have used in this dissertation.

- Rasmus T. Bjerngaard for setting up the flow rack.

I would like also to thank:

- Henrik Laursen for being very helpful and eagerly trying new ideas, software distributions and packages.

- Kjeld Martinsen for making the wire phantom used in the 3D scans.

Last, but not least I am greatly thankful to my girlfriend Madalina Breten, for all the care, understanding, and support she has given me during the intensive years of my Ph.D. study, and for proof reading parts of this dissertation and making useful comments about its contents.

The thesis is typeset using LaTeX, and the drawings are done in XFig. Most of the tools used in my everyday work were a product of the movement for open source software.

This work was supported by grant 9700883 and 9700563 from the Danish Science Foundation, by B-K Medical A/S, and by grant EF-782 from the Danish Academy of Technical Sciences.

Preface

This dissertation marks the end of my Ph.D. study which began on September 1st, 1998. The study was carried out at the Center for Fast Ultrasound Imaging (CFU), the director of which is my advisor Prof. Jørgen Arendt Jensen. CFU is based in the department Ørsted•DTU, at the Technical University of Denmark. Initially the goal of the project was to develop and test a method for real-time 3D imaging based on the "exploso-scan" - a real-time scanning procedure developed at Duke University and for 10 years the only real-time 3D scanning technique. This approach involves the transmission of a wide ultrasound beam followed by the reception of multiple parallel beams, which are formed within the transmitted beam. The imaging relies on the use of 2D matrix arrays. Because not all of the elements can be accessed due to technological limitations, the major challenge is to select the active elements in such a way as to minimize the artifacts introduced by doing so. The initial efforts in the design of sparse 2D arrays resulted in a presentation at the 25th anniversary of the Danish Biomedical Engineering society[1].

A turning point in my research was the article by G. R. Lockwood and colleagues, "Real-time 3-D ultrasound imaging using sparse synthetic aperture beamforming" [1] which I read in the spring of 1999. In this article it is shown how a 3-D volume can be scanned plane by plane using synthetic *transmit* aperture. Usually 5000 transmissions followed by receptions are possible for a normal scan situation. In the paper it was shown how with only 5 emissions a plane could be scanned. If the volume is divided into 50 planes, then the number of volumes scanned per second is 20. All this is achieved without compromise in the image quality, (as the one existing in the "exploso-scan"), and is realizable with a conventional phased array and only 64 channels. At that time, the building of the experimental scanner, which now is known as RASMUS[2], was well under way, and the budget allowed only for 64 digital receive channels. These factors combined weighed down the pair of scales in favor of the synthetic aperture focusing. The field of synthetic aperture *medical* ultrasound imaging has never left the area of research. Apart from the technical issues, one of the hold-backs of implementing this modality in commercial scanners was its perceived inapplicability for blood flow imaging. Led by the principle "never say that this cannot be done", the firm believe that this was the way to go, and encouragement by my advisor and colleagues (mainly Peter Munk) I boldly plunged into this exciting, and new to me (and to the other members of CFU) field.

A modern scanner is unimaginable without the capability to visualize the blood flow. The algorithms for velocity estimation operate on data sequences formed from the samples at the

[1] The talk was a part of a contest for student presentations, and won the third prize.

[2] RASMUS stands for Remotely Accessible and Software-configurable Multi-channel Ultrasound System.

same spatial position acquired at successive emissions. In conventional scanners the data used for velocity estimations is acquired by transmitting several times in the same direction. This was not possible using the existing synthetic aperture algorithms. At least several emissions were necessary before a scan line was formed. The idea was to modify the way the synthetic transmit aperture was implemented and to create a new image at every emission. My advisor Jørgen Jensen suggested a recursive algorithm for the beamformation, and I worked out the details of implementing it. The image is updated after every emission. This is done by adding the newly acquired information to the beamformed image and discarding the old information acquired several emissions ago. I suggested several more versions of the recursive imaging, but they remained unexplored since my efforts were later shifted towards designing motion compensation algorithm and a new velocity estimator.

In 1999 RASMUS was still on the design table, and another system was needed for the experimental verification of the method. The experiments were supposed to answer several questions: whether the time-domain focusing algorithms using linear interpolation could produce high-resolution images; whether the recursive procedure worked; how many emissions were necessary to obtain images with low grating lobes; what the correspondence between the simulations and the measurements was. Such a system, the XTRA system, was available at CADUS[3]- and kindly given at my disposal by Dr. Jens Wilhjelm, the director of CADUS. The experiments went as planned, and the results were in good correspondence with the simulations.

Most of the year 1999 was spent for teaching and for developing the system software and test programs for RASMUS. This software was developed, and almost finalized without any actual hardware present. In time it became apparent that the initial plans for developing and producing the system were too ambitious, and that the system would be delayed. More measurements were needed to test the possibility of estimating the blood flow using recursive imaging. A special course, in which I was a co-tutor, was organized for Kim Gammelmark, who calibrated the XTRA system and did an excellent job of scanning a tissue mimicking phantom at 65 separate positions. The raw channel data from 64 channels for 13 emissions at every positions were stored. The abundance of data allowed for different imaging strategies to be tested using off-line beamforming. The attempts to estimate the velocity using recursive ultrasound imaging on the acquired data, however, failed.

The major cause for the failure were the motion artifacts. They were caused by the relative motion between the phantom and the change in position of the transmitting element. The next year, 2000, was spent searching for algorithms for motion compensation.

Two articles played a decisive role in the way my efforts took: "Motion estimation using common spatial frequencies in synthetic aperture imaging" by Bilge, Karaman and O'Donnel [2], and "Sparse array imaging with spatially-encoded transmits" by Chiao, Thomas, and Silverstein [3]. The former showed that the images obtained using the same pairs of transmit and receive elements exhibit higher correlation, and that it was possible to estimate the gross tissue motion. The latter discussed how to use the same transmit elements at every emission combined with spatial encoding. The ideas from these two articles were further developed by me resulting in a combined motion compensation and velocity estimation scheme. The velocity is estimated from motion compensated data. The estimated velocity is used for motion compensation and so on. The motion compensation was applied on data obtained with and without spatial encoding. The performance was successfully tested on the data measured by Kim. Both the standard deviation, and the bias were around 2 %. There was one "but" - the data contained

[3]CADUS stands for Center for Arteriosclerosis Detection with Ultrasound. The center is based at Ørsted•DTU.

no noise. Noise increases the uncertainty of the estimates. The approach heavily relies on the success of both, the velocity estimation and motion compensation. The failure of any would result in the failure of the other.

Another event of great importance to me was the lecture given by Martin Anderson (co-authored by Gregg Trahey): “The *k*-space analysis of ultrasonic flow estimation”. It was delivered at the International Summer School on Advanced Ultrasound Imaging, held at the Technical University of Denmark from July 5 to July 9, 1999. In this lecture an attempt to characterize some transient effects in ultrasound was made. This gave me the inspiration to develop a model explaining the motion artifacts.

The result of all these efforts came in the spring of 2001. Apart from the simple model of the motion artifacts, a modified velocity estimator was developed. The estimator can be applied on the beamformed data from the recursive ultrasound imaging. The estimator needs no motion compensation and is capable of estimating blood profiles. Even more, a frame of how to modify most time-domain algorithms for velocity estimation was developed.

All the three years were marked by the continual effort to bring the experimental system RASMUS to life, mainly developing the software and debugging the hardware. The joint efforts of the group (in the last year mostly of Borislav Tomov and Peter Munk) were rewarded with success in May 2001, when the first *in-vivo* images were obtained. The only unusual thing was that these were not conventional B-mode images as everyone expected, but images obtained using synthetic transmit aperture. Only two emissions were used in transmit. The transmissions were carried out using multiple elements and frequency modulated pulses. The beamformation was done off-line, using a standard PC and the ”Beamformation toolbox”. The speed was satisfactory, approximately one frame per second. The frequency modulated pulses were designed using a script by Thanasis Misaridis, a fellow Ph.D. student whose project involved the use of coded excitations.

Since my years as an undergraduate student I have had fondness for 3D graphics. My hopes of making some advanced visualization and 3D volume rendering did not come true, but some work in the field was done. A visualization module for RASMUS was needed and the big question was whether an average PC could cope with the task. In 1995, I saw a demo by Dimitar Lazarov, a friend of mine, who showed to me how morphing could be implemented by mapping textures to polygons and changing the shape of the polygons. Later he made a screen saver using OpenGL. The screen saver was simple - triangles rotating on the screen with colors blending from one to another. From the screen saver I saw how the blending of the colors could be realized in OpenGL with very few commands. The hardware acceleration resulted in a breakneck frame rate. This gave me the idea to use OpenGL and hardware acceleration instead of optimizing the existing algorithms for scan-conversion. The acquired data is treated as a texture which is mapped onto a number of rectangles. The shape of the rectangles is distorted to match the geometry of the scanned region. The rendering and the interpolations are performed by the 3D accelerated video card. Based on a small engine written by me, Juan Pablo Gómez Gonzaléz developed a visualization program for his M.Sc. project.

During the project I worked on a number of problems, some more theoretical, other very practical, but all of them interesting. I tried to summarize the most important results achieved during the past three years in this report. I hope that this brief historical overview of my work will give a better perspective on the contents of the dissertation (as it presently is).

Nomenclature

Abbreviations

ADC	Analog to digital converter
ARUI	Add-only recursive ultrasound imaging
B-mode	Brightness mode
CFM	Color flow mapping
CRUI	Classical recursive ultrasound imaging
CW	Continuous wave
F-number	Ratio between the distance to focus and width of aperture
FM	Frequency modulation
GRUI	Generalized recursive ultrasound imaging
GSAU	Generic synthetic aperture ultrasound (imaging)
GSNR	Gain in signal to noise ratio
IMSLR	Integrated main to side lobe ratio
HRI	High resolution image
LRI	Low resolution image
ML	Main lobe
NDT	Non-destructive testing
PM	Phase modulation
PSF	Point spread function
PSNR	Peak signal to noise ratio
RF	Radio frequency
RMS	Root mean square
RUI	Recursive ultrasound imaging
SAD	Sum of absolute differences
SAR	Synthetic aperture radar
SNR	Signal to noise ratio
SRAU	Synthetic receive aperture ultrasound (imaging)
STAU	Synthetic transmit aperture ultrasound (imaging)
STRAU	Synthetic transmit and receive ultrasound imaging
TGC	Time gain compensation curve

Symbols

a_j	Apodization coefficient of element j

a_{ij}	Apodization coefficient for element j after emission with element i
$a_r(x,y)$	Receive apodization function
$a_t(x,y)$	Transmit apodization function
$a_{t/r}(x,y)$	Transmit/receive apodization function
$\mathsf{A}_t(x,y;z)$	Transmit radiation pattern at distance z from the transducer
$\mathsf{A}_r(x,y;z)$	Receive sensitivity function at distance z from the transducer
$\mathsf{A}_{t/r}(x,y;z)$	Two-way radiation pattern at depth z
B	Fractional bandwidth
c	Speed of sound
d_x	Transducer pitch
D	Transducer width
D_{act}	Width of the active aperture
D_{cluster}	The size of a cluster of elements
E_s	Energy of the signal
$f_{\#}$	F-number
f_0	Center (carrier) frequency
f_i	Instantaneous frequency
$f_{fr\text{ method}}$	Frame rate of the “method”
$f_m(\vec{x})$	Scatterer function
f_{prf}	Pulse repetition frequency
f_s	Sampling frequency
f_x	Spatial frequency along x
f_y	Spatial frequency along y
$\vec{f}$	Force
$\vec{f}_B$	Body force per unit volume
$\vec{f}_S$	Surface force per unit area
F	Distance to a fixed (mechanically) focal point
$g(t)$	RF pulse
h	Height, for transducer this is the width in elevation plane
$h(t)$	Impulse response
$h_m(t)$	Impulse response of a matched filter
h_{pe}	Pulse echo impulse response
$\mathbf{H}_N$	Hadamard matrix of order N
$H^{(n)}(t)$	High resolution scan line formed at emission n
$H_i(t)$	High resolution scan line formed after transmission with element i
$\mathbf{H}^{(n)}(t)$	High resolution image formed at emission n
$\mathbf{H}_i(t)$	High resolution image formed after transmission with element i
k	Wavenumber
k_x	Projection of k on the x axis
k_y	Projection of k on the y axis
k_z	Projection of k on the z axis
L	Length (lateral size) of a synthetic aperture
L_{arc}	Length of an arc
L_{seg}	Length of a segment
$L^{(n)}(t)$	Low resolution scan line formed at emission n
$L_i(t)$	Low resolution scan line formed after emission with element i
$\mathbf{L}^{(n)}(t)$	Low resolution image formed at emission n

$\mathbf{L}_i(t)$	Low resolution image formed after transmission with element i
m	Scaling coefficient
M n	Emission number
n_{int}	Interpolated value of n
N_s	Number of samples
N_c	Number of estimates over which the cross-correlation is averaged
N_l	Number of lines in image
N_{act}	Number of active elements
$N_{parallel}$	Number of parallel beamformers
N_{pos}	Number of positions
N_{rcv}	Number of receive elements
N_{xmt}	Number of transmit elements, number of emissions
N_{xdc}	Number of transducer elements
N_x	Number of elements in x direction
N_y	Number of elements in y direction
$N_{x;\,\mathrm{skip}}$	Number of skipped elements along x
$\vec{n}$	Normal vector
p	Pressure
p_i	Incident pressure
p_t	Transmitted pressure
p_0	Ambient pressure
$p_r(\vec{x},t)$	Pulse echo response
$\mathbf{p}_0$	PSF of a low resolution image at azimuth angle 0
$\mathbf{P}$	PSF of a high resolution image
P_N	Power of the noise
$q_i^{(n)}$	Encoding coefficient applied on the signal of the ith channel
$\mathbf{Q}$	Encoding matrix
r	Radius in polar coordinate system, traveled distance
$r(t)$	Received signal
$r(t,x_i)$	Received signal by the element with spatial location x_i, $r(t,x_i) \equiv r_i(t)$
r_{max}	Maximum distance
r_{min}	Minimum distance
$r_j(t)$	Signal received by element j
$r_j^{(n)}(t)$	Signal received by element j after the nth emission
$r_{ij}(t)$	Signal received by element j after emission with element i
R	Radius
$R_{ss}(\tau)$	Auto correlation function of $s(t)$
$R_{mn}(\tau)$	Cross correlation between signal m and n
$s(t)$	Signal, scan line of an image
S	Closed surface
t	Time relative to the start of emission (fast time)
t_s	Time shift
T_{acq}	Time for acquisition of an image
T_{prf}	Pulse repetition period
T_s	Accumulated time shift
u	Particle velocity
v_n	Normal component of the velocity

$\vec{v}$	Velocity
$\vec{v}_0$	Ambient velocity
V	Volume
w	Width of a transducer element
w_v	Width of a virtual element
$w(t)$	Windowing function in time
$w[n]$	Discrete windowing function
$\vec{x}$	Point in space $\vec{x} = (x, y, z)$
$\vec{x}_i$	Position of the center of element i
$\vec{x}_f$	Focal point
$\vec{x}_c$	Reference point for delay calculation (center focus)
Z	Acoustic impedance
z	Depth

Greek symbols

α	Angle of divergence
β	Angle of rotation, angle between a blood vessel and the z axis
γ	Angle between a blood vessel and the beam
$\delta(t)$	Dirac delta function
Δ_f	RMS bandwidth
Δ_t	RMS duration
Δm	Mass
Δt	Elapsed time; time step
$\Delta\theta$	Sector size
$\delta_{\theta N\mathrm{dB}}$	Angular beamwidth (in azimuth plane) at a level of $-N$ dB
$\delta_{\phi N\mathrm{dB}}$	Angular beamwidth (in elevation plane) at a level of $-N$ dB
$\delta_{xN\mathrm{dB}}$	Beamwidth along x (in azimuth plane) at a level of $-N$ dB
$\delta_{yN\mathrm{dB}}$	Beamwidth along y (in elevation plane) at a level of $-N$ dB
$\delta_{zN\mathrm{dB}}$	Pulse length along z
ε	Difference, sum of absolute differences
η	Lag in space
θ	Azimuth angle
θ_i	Angle of incidence
θ_r	Angle of reflection
θ_t	Angle of transmission
θ_{step}	Step between two scan lines in a phased array image
θ_{min}	The minimum (starting) angle of a phased array image
θ_{max}	The maximum (ending) angle of a phased array image
κ	Compressibility
λ	Wavelength
ξ	Spatial lag
ρ	Density
ρ	Normalized correlation
ρ_0	Ambient density
σ	Standard deviation
Σ	Aperture area
τ	Delay, lag

τ_{ij}	Delay of the signal received by element j after emission with i
ϕ	Elevation angle
φ	Phase
Φ	Velocity potential
ω	Angular frequency

Other symbols

$g(t) \rightleftharpoons G(f)$	Fourier transform pair
$\vec{\cdot}$	Vector of $\cdot$
$\hat{\cdot}$	Estimate of $\cdot$
$\underset{\sim}{\cdot}$	Complex value of $\cdot$
$\underset{\sim}{\cdot}^*$	Complex conjugate of $\underset{\sim}{\cdot}$
$\angle(\vec{x},\vec{y})$	Angle closed by $\vec{x}$ and $\vec{y}$
$\cdot^{(n)}$	Pertaining to the nth emission
$\lvert\cdot\rvert$	Euclidean norm of the vector
$*$	Convolution
$\underset{t}{*}$	Convolution in time
$\underset{x}{*}$	Convolution in space
$\Delta\cdot$	Fraction of $\cdot$
$\bar{\cdot}$	Mean of $\cdot$
$\mathcal{F}\{\cdot\}$	Fourier transform of $\cdot$
$\mathcal{R}[\beta]\{\cdot;\vec{x}_0\}$	Rotation of $\cdot$ with angle β round a pivotal point $\vec{x}_0$
$\mathcal{T}[\Delta\vec{x}]\{\cdot\}$	Translation with $\Delta\vec{x}$ of $\cdot$

List of Figures

List of Tables

CHAPTER

ONE

Introduction

1.1 Incentive for the work

Before we answer the question ”why three-dimensional ultrasound imaging” we should give an answer to the question ”why ultrasound imaging ?” in the first place.

Throughout the years the ultrasound imaging has established itself as one of the main imaging modalities because:

- it is considered to be relative safe, since there have not been any reports showing the opposite;
- it involves no ionizing radiation;
- it is non-invasive and most of the examinations do not distress the patient;
- it does not require special facilities as the X-ray, CT and MRI imaging require. Portable ultrasound scanners exist, enabling the use of ultrasound even in field conditions.
- it displays the dynamics of anatomy due to its real-time capabilities, and gives a real-time view of the functions of the organs;
- it allows for the estimation of such vital parameters as the function of the organs.

The portability and the real-time display are the two key factors for the success of the ultrasound imaging. At present day most ultrasound scanners are capable of displaying only two dimensional cross-sections of the anatomy, and the clinician must reconstruct mentally the three-dimensional structure of the organs under investigations.

Add-ons are used for ultrasound scanners, enabling off-line reconstruction of the 3D images. Although this method gives images with high quality, it loses one of the main advantages of ultrasound, namely the real-time display of the function of organs.

The real-time capabilities are preserved in scanners by using 2-D matrix arrays, which allow for real-time 3D imaging. This, however is accompanied by an immense increase in the complexity and cost of the ultrasound scanner, which cancels the advantage of being relatively cheap and accessible. As it can be seen, there exists a trade-off between scanner complexity and real-time capability. This trade-off comes from the way the ultrasound data is collected and processed: the volume is divided in planes, the plane in lines, and the lines are acquired one-by-one.

An alternative approach exists - instead of sending a pulse and receiving an echo from a single direction, the ultrasound energy can be spread into the whole region of investigation and consecutively processed using *synthetic aperture focusing*. This approach is capable of decreasing the acquisition time without trade-off in resolution. Until recently it was perceived as inapplicable for flow imaging and was not regarded as a viable alternative to the established imaging algorithms. This dissertation deals with some of the weaknesses of the synthetic transmit aperture imaging and tries to develop methods for estimating the blood velocity.

1.2 Contributions of the dissertation

The primary contributions of this dissertation are in the design of recursive algorithm for synthetic aperture imaging, an unified view of the virtual source element, the exploration of the effects of motion on the different acquisition methods, the use of recursive imaging for motion compensation and velocity estimation at the same time, and the use of virtual source element for 3D imaging.

Some work on 2D sampling patterns for matrix transducers has also been done and is presented in the thesis. Practical aspects of how to visualize 3D data using current mass-production computer hardware are also considered in the report.

Most of the work has also been presented at conferences and in articles. These are:

1. **S. Nikolov**, and J. A. Jensen: Velocity estimation using synthetic aperture imaging. The paper will be presented at the IEEE Ultrasonics Symposium in Atlanta, October 2001, and will be published in the proceedings of the conference.

2. **S. Nikolov**, J. P. G. Gonzaléz, and J. A. Jensen: Real time 3D visualization of ultrasonic data using a standard PC. Presented at the Ultrasonics International '01 Holland. The paper will be published in Ultrasonics.

3. **S. Nikolov**, K. Gammelmark and J. A. Jensen: Velocity estimation using recursive ultrasound imaging and spatially encoded signals, Proceedings of the IEEE Ultrasonics Symposium in Puerto Rico, October 2000.

4. **S. Nikolov** and J. A. Jensen: 3D synthetic aperture imaging using a virtual source element in the elevation plane, Proceedings the IEEE Ultrasonics Symposium in Puerto Rico, October 2000.

5. J. A. Jensen and **S. Nikolov**: Fast simulation of ultrasound images, Proceedings of the IEEE Ultrasonics Symposium in Puerto Rico, October 2000.

6. **Svetoslav Nikolov** and Jørgen Arendt Jensen: Application of different spatial sampling patterns for sparse-array transducer design, Presented at the Ultrasonics International '99 and 1999 World Congress on Ultrasonics, Denmark, 1999, published in Ultrasonics February, 2000.

7. Jørgen Arendt Jensen, Ole Holm, Lars Jost Jensen, Henrik Bendsen, Henrik Møller Pedersen, Kent Salomonsen, Johnny Hansen and **Svetoslav Nikolov**: Experimental ultrasound system for real-time synthetic imaging, Proceedings of the IEEE Ultrasonics Symposium, October, 1999.

8. **Svetoslav Nikolov**, Kim Gammelmark and Jørgen Arendt Jensen: Recursive Ultrasound Imaging, Proceedings of the IEEE Ultrasonics Symposium, October, 1999.

Two patent applications are made:

- Jørgen Arendt Jensen and Svetoslav Ivanov Nikolov. Recursive Ultrasound Imaging, International Application Number PCT/DK00/00245.
- Svetoslav Ivanov Nikolov and Jærgen Arendt Jensen. Velocity Estimation Using Synthetic Aperture Imaging.

A talk describing the exploso-scan and presenting some original 2D matrix arrays was given at the 25th Anniversary of the Danish Biomedical Engineering Society, November 1998, Copenhagen. The title of the talk was: "An approach for three-dimensional real-time ultrasound imaging".

Some of the more relevant software written through the course of my Ph.D. study is:

- The system software for the RASMUS system. This project was by far the largest for the last three years. It includes low-level Linux kernel modules (drivers), TCP/IP communications, custom made remote procedure calls and connectivity with Matlab. The software features a client/server architecture at all levels. The documentation in terms of users' and developers' guides exceeds 300 pages. The total number of files (including the documentation) is more than 1000.
- PCI toolbox. This is software for testing any PCI device. It consists of custom made kernel module and a library with unified API. The functions can be also called directly from Matlab.
- Beamformation toolbox. This is a C library implementing time-domain beamformation. Most of the procedures are optimized for speed and realize only linear linear interpolation. High-order interpolation using filter banks is also available to the user. There is a Matlab interface, which closely resembles the interface of the main simulation program Field II.
- Color flow mapping toolbox. This is a combination of a C library and Matlab scripts for creating color flow maps. The flow is found either by estimating the phase or the time shift of the signal. Also frequency based velocity estimator is available.
- Libraries for storing envelope detected data and interpreting initialization files. The latter is shared across several projects - from the system software for RASMUS to the 3D display program. The former has a Matlab interface and is also used in the program for visualization 3D data.
- Parts of a the visualization program by J. P. G. Gonzaléz: the scan conversion engine, the support for initialization files and the loading of envelope detected data. Also the program was ported to Linux by me and some bugs were removed.
- A collection of Matlab scripts for synthetic aperture imaging, motion compensation and velocity estimation.

Each of the individual projects is accompanied by a user's guide, which is bundled with the distribution of the software package.

Some programs not directly related to the scientific part of the project were also made, such as the collection of scripts for 2D and 3D vector graphics. These scripts were used to draw most of the figures in the dissertation.

1.3 Organization of the thesis

The thesis is primarily devoted to the synthetic aperture focusing algorithms. Before diving into the essence of the work, the conventional ultrasound scanning methods are overviewed and the foundations of acoustics is given in Chapters 2 and 3, respectively. Among the conventional methods and trends, the use of standard PC components is considered and the OpenGL scan-conversion algorithm presented.

The rest of the dissertation is divided into two parts: “Synthetic aperture tissue imaging” and “Synthetic aperture flow imaging”. In the first part an attempt is made to introduce some classification based on which aperture is synthesized: the transmit, the receive or both. The basic principle of synthetic aperture focusing is given in Chapter 4, and its variations are presented in Chapter 5. Further, the idea of creating an output image at every emission using recursive ultrasound imaging is introduced in Chapter 6. In this part of the thesis some of the problems of the synthetic transmit aperture imaging are outlined such as a relatively large number of emissions and low signal to noise ratio. Several solutions exist to the latter problem such as the use of multiple elements in transmit. This idea is further developed into the unifying concept of the *virtual* ultrasound sources, which is given in Chapter 7. The chapter finishes by presenting the use of virtual sources to increase the resolution in the elevation plane for 3D synthetic aperture imaging with a linear array. The number of emissions is decreased by using sparse synthetic arrays as shown in Chapter 8. The design of 1D and 2D sparse arrays is outlined, and the performance of several 2D designs is compared. This part of the dissertation finishes with a presentation of spatial and temporal encoding for increasing the signal to noise ratio.

The second part of the dissertation is devoted to the velocity estimation using synthetic aperture imaging. It starts with a description of velocity estimation by measuring the time shift of the signals in Chapter 10. This velocity estimator is chosen because it performs well with short transmit pulses. The velocity cannot be estimated using the images created with the recursive ultrasound imaging. The cause of the failure are the motion artifacts, a model of which is developed in Chapter 11. Based on this model a motion compensation scheme is designed. Chapter 12 presents the results of the motion compensation, and shows that the velocity can be estimated using the compensated images. In order to eliminate the motion compensation as a step in the flow imaging, the cross-correlation velocity estimator is modified in Chapter 13 using the model of the motion artifacts. This part of the dissertation ends with a discussion of how other velocity estimators can be modified for use with synthetic aperture imaging.

The Appendix contains the papers published in the course of the Ph.D. study and a description of some of the equipment used for the experiments.

CHAPTER

TWO

Ultrasonic imaging

This chapter briefly introduces the ultrasound scanning as a medical imaging modality. It gives a basic overview of the process of collecting and displaying ultrasound data, and shortly presents the existing imaging techniques. Here the term *imaging* covers the viewing of both, anatomical structures, and their function[1]. The division between one-dimensional (1D), two-dimensional (2D), and three-dimensional (3D) imaging is based on the number of spatial dimensions [2] viewed on the screen. The application of off-the-shelf computer components and the transition from hardware to software solutions is shortly discussed. The chapter finishes with the presentation of a software display module developed at the Center for Fast Ultrasound Imaging.

2.1 Fundamentals of ultrasonic imaging

Ultrasound usually denotes mechanical vibrations with a frequency of oscillations above 20 KHz, the highest audible frequency. The term *ultrasound imaging* covers the area of remote sensing, where mechanical vibrations with known parameters are generated, sent through a medium, and consecutively recorded. The changes of the parameters introduced during the propagation are used to characterize the medium.

In medical applications ultrasound usually means longitudinal waves with frequency ranging from 1 MHz up to 50 MHz. The diagnostic ultrasound as an imaging modality has some advantages compared to other medical imaging modalities. It is regarded as relatively safe, since there have not been any reports showing the contrary[4]. It involves no ionizing radiation, it is non-invasive and most of the examinations do not distress the patient. It does not require special facilities as the X-Ray, CT and MRI imaging require and portable instruments as the one shown in Figure 2.1 are available. Ultrasound allows the visualization of soft tissues. Although visually with lower quality than, say a CT scan, the ultrasound imaging has the advantage of displaying the dynamics of the anatomical structures.

The development of ultrasonic imaging has resulted in the following modalities: intensity mapping, pulse-echo mapping and phase-amplitude techniques. The pulse-echo mapping is the basis for the visualization of anatomical structures and further we will be concerned mainly with it.

[1] The function can be exhibited through the velocity of blood, the elastic properties of the tissue, etc.

[2] For marketing purposes some companies refer to the real-time 3D scanning as 4D scanning

Figure 2.1: Portable ultrasound scanner. The total weight of the device is 2.4 kg. (Courtesey of SonoSite)

2.2 Ultrasonic morphological imaging

The pulse-echo technique emerged during the World War I and was applied in sonar systems. Further, the technique evolved, and from the 1950s [5] the ultrasound was used for medical purposes.

The pulse-echo imaging can briefly be described as follows. An ultrasound pulse is sent into the tissue in a given direction. As the pulse propagates through the tissue it passes through some inhomogeneities, or reflective surfaces. Each inhomogeneity, depending on its size, causes part of the pulse energy to be *scattered* or *reflected* back to the transducer. The transducer receives this echoes, and converts them into electrical signal (voltage). Since the inhomogeneities are related to the transition between different tissues and to the type of the tissue, a map of the tissue layout can be made from the received signal.

As the technology developed, the ultrasound imaging went from imaging only a single direction (1D), through viewing a whole plane (2D) to displaying the anatomy in its whole (3D).

2.2.1 One-dimensional ultrasonic imaging

A-mode

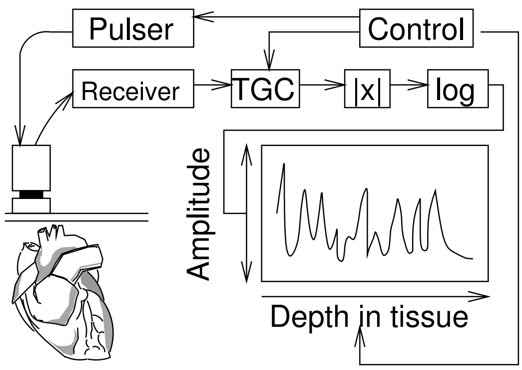


Figure 2.2: Block diagram of a simple A-mode system

Figure 2.2 shows the block-diagram of a simple system working in amplitude mode (A-mode). The pulser generates an electric radio frequency pulse. The transducer converts the electrical pulse into a mechanical vibration. As the pulse propagates through the tissue part of the energy is reflected and scattered back by the present inhomogeneities. The transducer converts the received mechanical wave into a radio frequency (RF) electrical signal, which is sent to the receiver. As the signal propagates further into the tissue it is subject to attenuation and the echoes get weaker with depth, respectively time. The level of the signal is equalized by the the Time Gain Compensation Amplifier (TGC). Although the speed of sound c is different in the different tissues, it is usually assumed to have a mean value of 1540 m/s. From the TGC amplifier, the signal is passed to an envelope detector. The level of the signal depends on the magnitude of the inhomogeneities, and is usually in a narrow region. In order to emphasize the small differences, the signal is logarithmically compressed prior to displaying. For the visualization purposes, a simple analogue oscilloscope can be used. The signal is connected to the Y input. On the X input a generator of triangular signal determines the horizontal offset, respectively the depth into the tissue.

M-mode

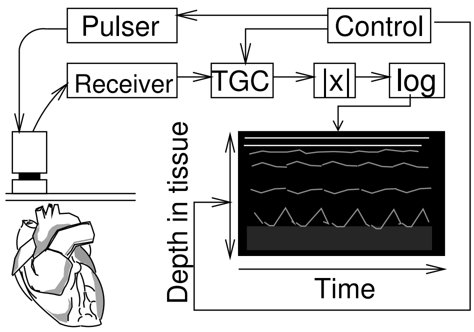


Figure 2.3: Block diagram of a M-mode system

Figure 2.3 shows the block diagram of a system working in a motion mode (M-mode). The process of acquiring and processing the data in the M-mode systems is basically the same as in the A-mode systems. The difference is mainly in the visualization. The logarithmically compressed and envelope detected signal is used to modulate the brightness of the beam on the screen. Thus the vertical screen coordinate y corresponds to depth in tissue, and the horizontal screen coordinate x corresponds to time. The brightness of the signal corresponds to the magnitude of the received echo.

2.2.2 Two-dimensional ultrasonic imaging

B-mode

The systems working in brightness mode (B-mode) are capable of displaying a 2D cross sectional image of the anatomy. As in the M-mode systems the received signal modulates the strength of the electronic beam of a cathode ray tube[3] (CRT). The difference is that the information is not gathered only from the same direction, as in the A- and M- mode systems. In the early B-mode scanners the scanning was performed using a transducer fixed on a mechanical

[3]This does not hold for the liquid crystal displays (LCD). It is more appropriate to say that the strength of the echo is mapped to the brightness of the image.

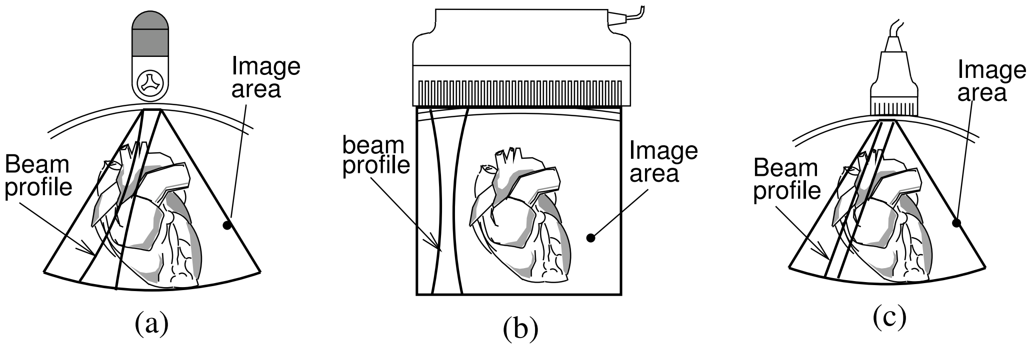


Figure 2.4: Acquiring B-mode images using: (a) mechanically rotated single crystal transducer, (b) linear-array transducer, and (c) phased-array transducer.

arm. The acquired image was static. The physician would sweep the transducer around the patient, and the anatomy would be visualized on the screen. These systems gave nice images, and were valuable in diagnosing the abdomen because they gave a good view of static structures over a large area of the body.

The development of technology led to real-time B-mode imaging. In this case the transducer is either a single mechanically rotated crystal as shown in Figure 2.4(a), or a multi-element transducer. The multi-element transducers can steer the ultrasound beam electronically. The linear array probes part of the tissue by activating groups of elements directly above the tissue as shown in Figure 2.4(b). The phased array uses all of its elements and steers the beam in different directions as shown in Figure 2.4(c).

2.2.3 Three-dimensional imaging

In the last years the 3D ultrasound imaging has become a major area of research, and a strong selling point for many companies. Very few systems are capable of *real-time* three-dimensional scanning [6]. Most 3D systems are standard 2D scanners with a positioning system and a work station added to them as shown in Figure 2.5.

The type of scanning depicted in Figure 2.5 is known as ”freehand 3D scanning” [7, 8, 9]. The volume is scanned plane by plane, and the transducer is moved freely by the physician. Some type of positioning system is used (usually using electromagnetic sensors [10]) with the receiver mounted on the transducer. The position information is fed into the acquisition control module. This module controls a frame grabber that captures the video signal coming out of the scanner. The digitized data is further sent to a graphics workstation where from the planar data and the position information the volume is reconstructed. There are numerous ways to present the acquired information to the user. One of the most popular approaches is to segment the data based on the intensity level. This is especially useful for fetal scans, where the object is surrounded with liquid and the border is easily extracted. The surface is then rendered, shading applied, and the face of the baby is shown to the happy mother.

The freehand scanning systems are, however, expected to be taken over by “real” three-dimensional scanners. There are two possibilities of scanning the volume, either by using matrix 2D arrays capable of electronic beam steering in the azimuth and elevation directions, or by using linear arrays. In the latter case the array can steer the beam electronically only in the azimuth direction, and must be mechanically moved in the elevation. Some of the various types of volumetric data acquisition using linear arrays are shown in Figure 2.6. In the azimuth

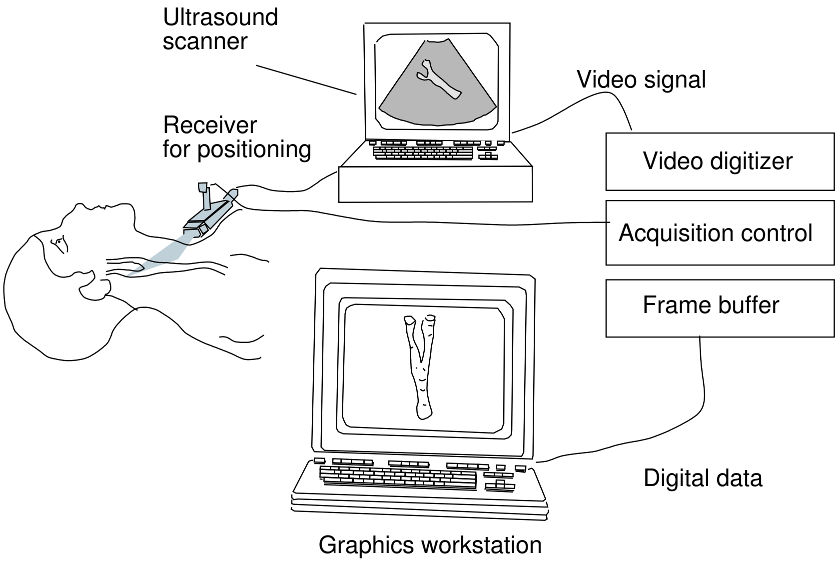


Figure 2.5: Block diagram of a common 3D ultrasound system showing the essential components used to acquire patient data [7].

plane the scans can be either linear or sector. The same applies for the elevation plane, yielding different geometries of the volume acquired. In many applications the data is first scan converted to fill in discrete volume elements on a rectangular grid, and then displayed.

When matrix arrays are involved, the most common format of the acquisition is the combination sector/sector (pyramidal) [6] because of the small size of the arrays employed. In recent years efforts have been also made towards a linear/linear (rectilinear) acquisition [11]. The major challenge for the real-time 3D systems is the speed of acquisition. If the scanned volume consists of 64×64 scan lines, 4096 emissions are necessary. For a typical scan depth of the 15 cm, and average speed of sound of 1540 m/s, only 5000 transmissions/receptions per second are available. This results in a refresh rate of 1 volume per second. Parallel receive beamforming is used to increase the refresh rate .

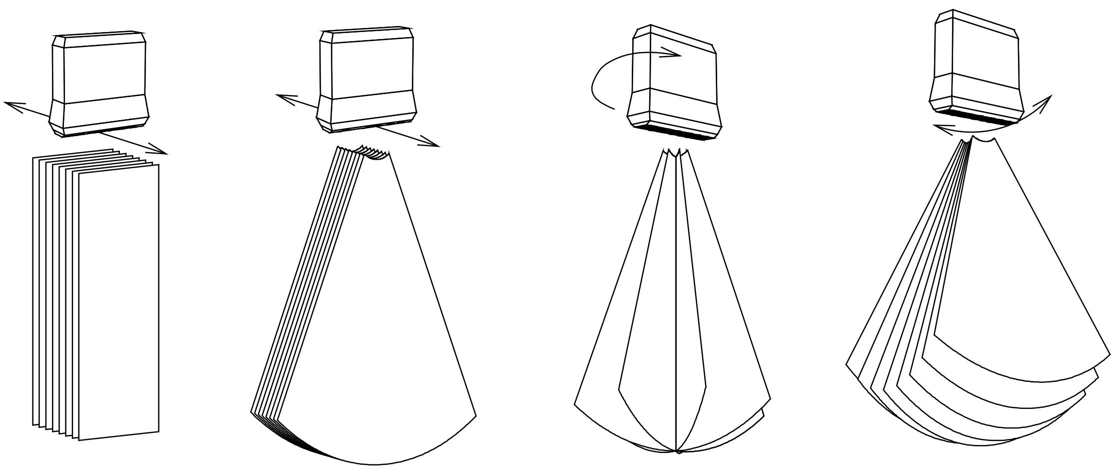


Figure 2.6: Different types of volumetric ultrasound data acquisition. The position and direction of the current plane is determined by a motor drive system.

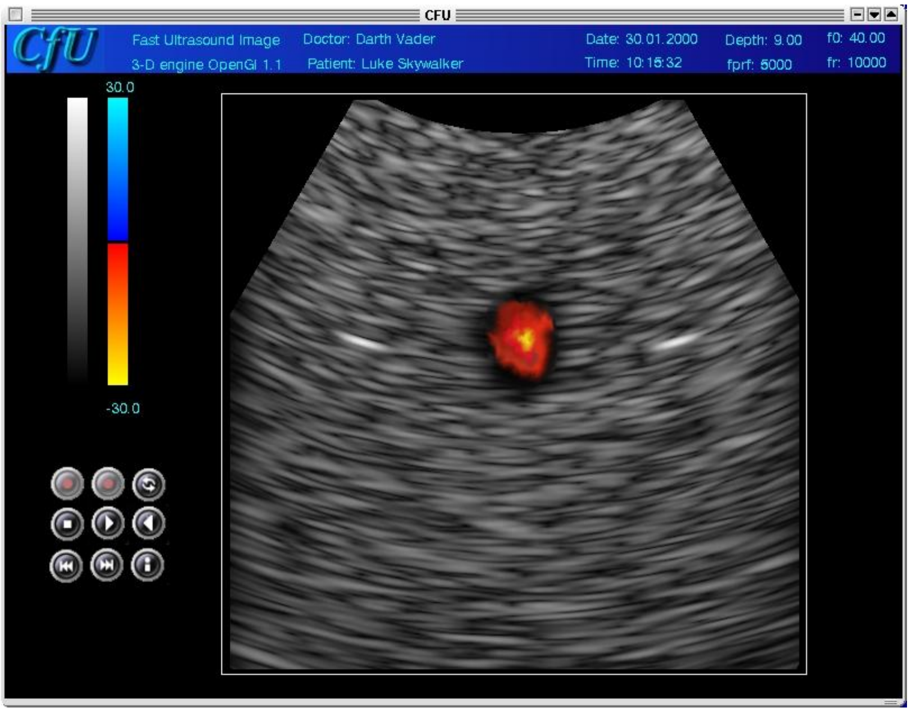


Figure 2.7: Example a color flow map superimposed on B-mode sector image.

2.3 Ultrasonic velocity imaging

By using ultrasound, it is possible to detect the velocity of moving particles. This is especially useful in detecting disorder in the cardio-vascular system. For example, a disturbance of flow in the arteries can be an indicator of stenosis.

In modern scanners, the most used for blood flow estimation is a technique known as "pulsed Doppler". Several lines in the same direction are acquired. The samples from the same depth are used to form a new signal. If the particles at a given depth are motionless, then the samples have the same value. If the particles are moving, then the new signal oscillates with a frequency which is proportional to the velocity of the particles. Estimating the frequency of this signal results in a velocity estimate at the given depth. The frequency can be estimated either by using Fourier transform, or by estimating the change in phase of the signal, or some other method.

The velocity can be estimated at several depths for several imaging directions, and an image of the velocity distribution is made. The magnitude of velocity and the direction of flow are mapped on the screen using different colors. Typically red and blue colors are used to show flow towards and away from the transducer. The brightness of the color corresponds to the magnitude. An example of a color flow map superimposed on a B-mode image is given in Figure 2.7.

2.4 Using off-shelf modules for utrasound scanners

In the last years the digital technology made major inroads in ultrasound scanners. The latest scanners employ digital beamformers, and the post processing is based on digital filters and processing algorithms. One of the advantages is, that some blocks of the ultrasound scanners that were traditionally custom designed can be readily acquired from third party manufacturers, thus reducing the production cost and development time. Another advantage is that the scanners can be reconfigurable [12] using software, rather than hardware. One block that is especially suitable to use off-the-shelf standard components is the display system of an ultrasound scanner.

The visualization of ultrasound data involves (1) logarithmic compression (also known in image processing as gamma correction), (2) scan conversion from polar to Cartesian coordinates in the case of a phased array sector scan, and (3) display and user interface. The most computationally expensive operation is the scan conversion. The number of lines acquired is not sufficient and interpolation of the data must be done [13, 14].

Most of the work on scan conversion algorithms was focused on their optimization regarding computational efficiency [14, 15, 16]. In spite of all the algorithmic optimizations the processors at present day would spend most of the processing time on scan conversion. This problem has been addressed by the vendors of video cards such as ATI and NVidia. They process the data using graphics oriented pipelines and specialized 256 bit graphics processing units with 128 bit wide memory interface. Standardized programming interfaces such as OpenGL exist, allowing with very few commands to define the mapping from scan to screen coordinates. This features of the video cards have been exploited based on an author's idea and algorithm, and the details regarding the user interface were worked out in a Master of Science project by Juan Pablo Gómez Gonzaléz [17, 18].

The program was supposed to be capable of visualizing any of the scan geometries, 2D and 3D, which are shown in Figures 2.4 and 2.6, as well as color flow maps.

2.4.1 Display of phased array sector images

Figure 2.8(a) shows the necessary conversion from polar to Cartesian coordinates. The dots symbolize samples along the scan lines. The lines in the image have a common origin in the (x,y) coordinate system. Every line looks in a different direction determined by an azimuth angle θ_i. The size of the sector in the image is $\theta_{max} - \theta_{min}$. The number of scan lines in the image is N_l. The angular step in the scan process is;

$$\theta_{step} = \frac{\theta_{max} - \theta_{min}}{N_l - 1}. \tag{2.1}$$

The other coordinate in the polar coordinate system is r, corresponding to the scanned depth. It is related to the time t by

$$t = \frac{2r}{c}, \tag{2.2}$$

where t is time from the current pulse emission (sometimes called fast time) and c is the speed of sound.

Neglecting the scaling coefficients and the discrete nature of the image and the acquired data, the relation between the scan and image coordinates is:

$$\begin{aligned} x &= r\cos\theta \\ y &= r\sin\theta \end{aligned} \tag{2.3}$$

The traditional approach is to try to fill the pixel at coordinates (x,y) by making a look up table pointing to the coordinates (r,θ), which correspond to the same spatial position. If the point at

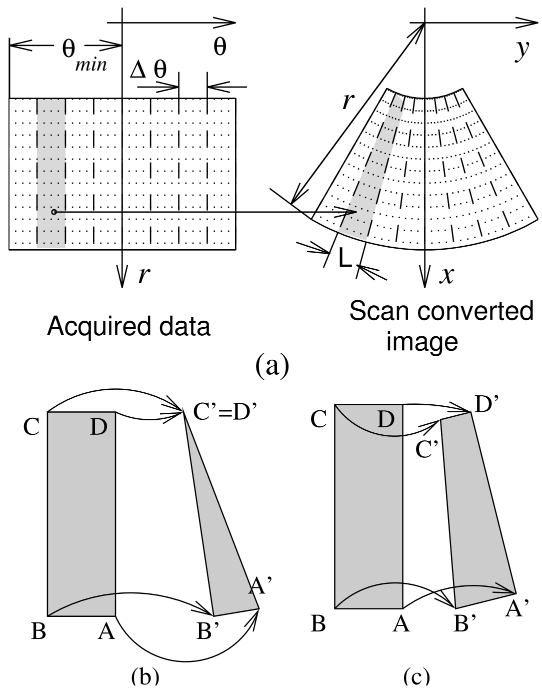


Figure 2.8: Mapping from polar to Cartesian coordinate system. Figure (a) shows the geometry of the acquisition and the display. Figures (b) and (c) show how the data is mapped to the screen when the start depth is 0, and greater than 0, respectively.

coordinates (x,y) lies between two scan lines, or/and between two samples in the scan line, its value is interpolated from the neighboring samples. This is done for every pixel in the image.

The approach undertaken in [18] is instead of mapping all the points, to define the mapping of the whole image. This is done by dividing the sector into tetragons, in other words the arcs are approximated by a collection of segments (Figure 2.8(a)). The same type of segmentation of the data region results in rectangles in the (θ, r) domain. The rectangular regions in the data to be displayed is treated as textures which must be mapped onto polygons displayed on the screen. This mapping can be handled by general purpose texture mapping algorithms which are implemented and accelerated by the hardware. The only necessary step for the developer is to specify the correspondence of the vertexes of the rectangles in the (r, θ) domain and the polygons in the (x,y) domain as shown in Figure 2.8(b) and (c).

The length of the outer arc bounding the scanned region is $L_{arc} = (\theta_{max} - \theta_{min})r_{max}$, where r_{max} is the maximum scan depth. The length of a single segment of the approximating tetragons is:

$$L_{seg} = 2r\sin\frac{\Delta\theta}{2} \tag{2.4}$$

$$\Delta\theta = \frac{\theta_{max} - \theta_{min}}{N_{seg} - 1}, \tag{2.5}$$

where N_{seg} is the number of segments used to approximate the arc. The smaller the difference

$$\varepsilon = L_{arc} - N_{seg}L_{seg}, \tag{2.6}$$

the better the visual appearance of the scan converted image is. The number of segments N_{seg} is generally different than the number of scan lines N_l, and is chosen based on the psychological

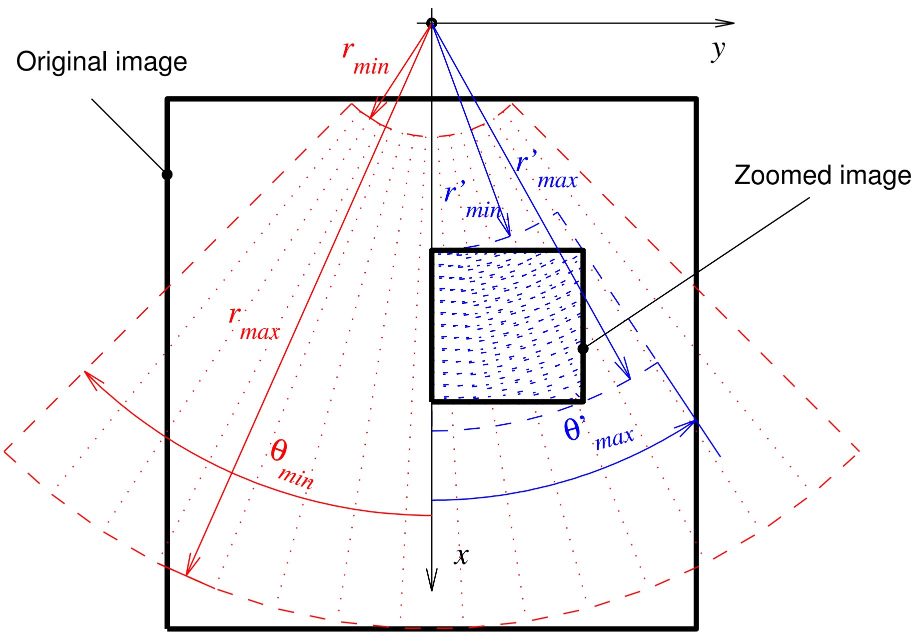


Figure 2.9: Geometry used for zooming in the image. The parameters necessary to find are the minimum and maximum radial distances and the minimum and maximum angles. The number of segments between θ'_{min} and θ'_{max} is kept constant. The clipping can be left to the hardware.

perception of the image displayed. As a rule of the thumb N_{seg} should be at least twice N_l. This ensures that a sample from one scan line is used in several polygons resulting in a gradual transition. A video card based on GeForce 256 can render more than 15 million triangles per second, so the number of used polygons can be quite large. The size of the image shown in Figure 2.7 is 512×512 pixels (this is the size of the active area - without the user interface). The size of the sector is $\theta_{max} - \theta_{min} = \pi/3$. The number of scan lines is $N_l = 64$ and the number of polygons is $N_{seg} = 256$.

The color flow map shown in the figure is independently scan converted. It is superimposed on the B-mode image by using a mask for the regions with flow present. Four bytes per pixel are used in OpenGL. Three of them are used for the red, green and blue components, and the fourth is a transparency coefficient (alpha blending).

One of the advantages of the software implementation is the ease to zoom in the displayed image. The geometry associated with it is illustrated in Figure 2.9. The essential thing to remember is that the visual quality depends on the number of polygons in the *visible part* of the image. Zooming in the image leads to a change in the minimum and maximum displayed angles and radial distances:

$$r_{min} \to r'_{min}$$
$$r_{max} \to r'_{max}$$
$$\theta_{min} \to \theta'_{min}$$
$$\theta_{max} \to \theta'_{max}$$

The arcs defined by the angles θ'_{min} and θ'_{max} are split again into N_{seg} segments. The resulting polygons are clipped with the new coordinates of the displayed image.

Another advantage of the undertaken approach is that more than one views can be displayed

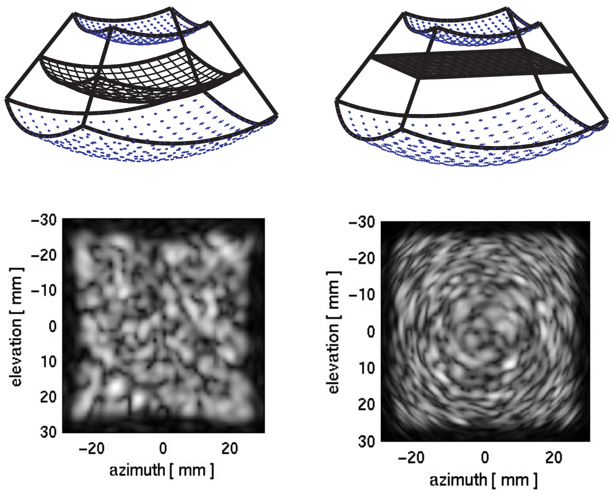


Figure 2.10: Displaying data parallel to the transducer surface: c-scan (left) and a cross-section with a plane (right).

with individual settings and displayed data.

2.4.2 Display of 3D images

The software is targeted at a real-time 3D display for volumetric data acquired at a rate higher than 10 volumes per second. For the case of a pyramidal scan format, the display must map data from polar to Cartesian coordinates.

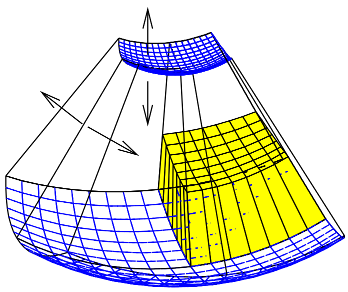
The images acquired from most parts of the body are not suitable for segmentation, surface reconstruction and other advanced visualization techniques. The most robust approach is to show several planes from the volume (multi-planar display). A slightly better approach in terms of orientation in the 3D volume is to display a solid with a shape of the scan format. Such a solid consists of 6 sides, two of which are parallel to the transducer surface. These two sides can either be planes or curved surfaces, where the curvature corresponds to a given depth in the volume (c-scan). The two approaches are illustrated in Figure 2.10. From the figure it can be seen that the image exhibits circular symmetry, which is not characteristic to the human body. Therefore the preferred method to the author was the c-scan depicted in the left side of the figure.

The display was chosen to be a solid with the shape of the scanned volume. The sides of the solid are rendered with the envelope detected data as shown in Figure 2.11(b). The clinician is able to view the data inside the volume by moving in and out the bounding planes as shown with arrows in Figure 2.11(a).

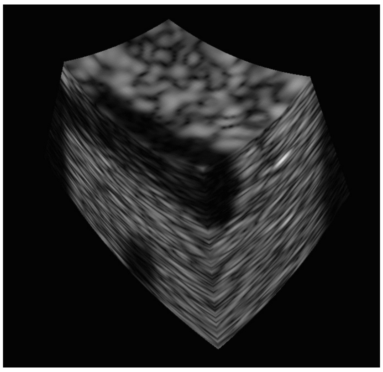
The scan conversion is done in the same way as for the 2D case but with 3D orientation of the polygons.

2.4.3 Performance

The performance in terms of number of frames displayed per second was measured on the following computer configuration: motherboard ASUS K7V with VIA KX133 chipset; 800



(a)



(b)

Figure 2.11: Display and navigation in 3D ultrasonic data

MHz AMD Athlon processor; video card ASUS V6800 with NVidia GeForce 256 GPU and 32 MB DDR. The software was tested under a standard Red Hat 6.2 distribution of Linux. The X server was XFree86, version 4.01 with drivers from NVidia, version 0.94-4.

The tests included the display of a phased array sector image, a linear array rectangular image and a 3D pyramidal image. The size of the input data set for the 2D images was 64 lines of 1024 samples per line. The 3D volume was 64×64 lines with 512 samples per line. The display mode was set to 1280×1024 pixels, 32 bits per pixel.

Table 2.1 shows the speed of display for a window with size 800×600 and 1280×1024 pixels. The sizes of the images are 512×512 and 860×860 pixels, respectively. For the cases of more than one views, the images have the same type of geometry, either sector or rectangular, since these represent the extreme cases in performance. Showing a mixture of geometries at the same time results in performance between the extreme cases.

The speed of display scales with the size of the image that must be rendered and the size of data that must be transfered to the memory of the graphics board. This is seen from the speed of the 3D display, in which the change in the number of displayed polygons almost does not affect the performance.

Phased array image

No of views	Image type	Frames/sec @ 800×600	Frames/sec @ 1280×1024
1	B-mode	360	187
	+CFM	272	146
2	B-mode	337	184
	+CFM	294	162
4	B-mode	208	120
	+CFM	191	111

Linear array image

No of views	Image type	Frames/sec @ 800×600	Frames/sec @ 1280×1024
1	B-mode	643	291
	+CFM	558	251
2	B-mode	602	289
	+CFM	571	276
4	B-mode	457	232
	+CFM	438	224

3D pyramid image

No of polygons	Frames/sec @800×600	Frames/sec @1280×1024
256×256	41	27
64 ×64	43	31

Table 2.1: Measured performance.

CHAPTER

THREE

Foundations of acoustics

For successful processing of ultrasonic information it is important to understand fully the physics behind the acquired signals. Having insight in the measurement situation makes it possible to extract features that are not readily seen from the data. The understanding of the physical laws also gives the possibility to create new measurement situations.

This chapter considers the fundamentals of medical ultrasound. It starts with the wave equation governing the propagation of ultrasound in human tissue. Then solutions to the wave equation are given for different coordinate systems. Most of the results in this dissertation are obtained through simulations using the program Field II [19]. The program is based on the calculation of spatial impulse responses and therefore this method for solving the wave equation is also considered. Further the approximations to the solution of the radiation problem known as the Fresnel and Fraunhofer approximations are presented to the reader, followed by an introduction to delay and sum beamforming. The chapter ends with the introduction of the frequency domain representation of the ultrasound systems, known also as *k*-space representation.

3.1 The wave equation

The opening paragraph is drawn from Insana and Brown [20]:

> Fluids have elasticity (compressibility κ) and inertia (mass density ρ), the two characteristics required for wave phenomena in a spatially distributed physical system whose elements are coupled. Elasticity implies that any deviation from the equilibrium state of the fluid will tend to be corrected; inertia implies that the correction will tend to overshoot, producing the need for a correction in the opposite direction and hence allowing for the possibility of propagating phenomena - acoustic pressure waves.

The tissue is characterized with some *ambient* parameters such as pressure p_0, density ρ_0 and velocity $\vec{v}_0$. There are three laws relating these parameters: (1) the conservation of mass; (2) the equation of motion of fluid and (3) the pressure-density relations. Based on these relations the wave equation of sound in tissue is derived. The following derivation of the wave equation is based on [21].

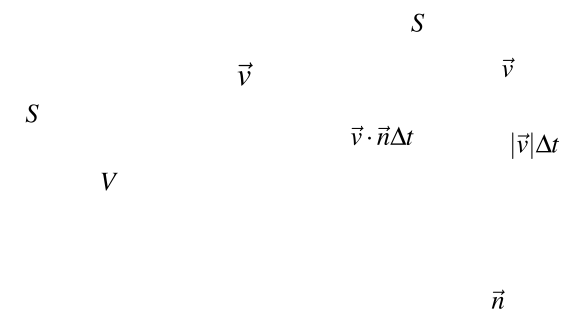


Figure 3.1: On the left: non-moving volume V with moving fluid. The time rate of mass within V is equal to the mass flowing through the surface S. On the right: the mass leaving through the element area ΔS in time Δt is equal to the mass in the slanted cylinder of length $|\vec{v}|\Delta t$, height $\vec{v} \cdot t$ and base area ΔS

3.1.1 Conservation of mass

Consider Figure 3.1. For a fixed volume V inside of a fluid such as air, water etc, the mass m inside V can be taken as the volume integral over of the density $\rho(\vec{x}, t)$, where $\vec{x}$ is a spatial point. The conservation of mass requires that the time rate of change of this mass to be equal to the net mass per unit time entering minus the net mass leaving that volume through the bounding surface S. The mass leaving the volume through a surface area ΔS is:

$$\Delta m = \rho(\vec{x_s}, t)\vec{v}(\vec{x_s}, t) \cdot \vec{n}(\vec{x_s})\Delta S, \tag{3.1}$$

where $\vec{x_s}$ is a point on the surface S, $\vec{v}(\vec{x_s}, t)$ is the fluid velocity and the subscript s means that $\vec{x}$ lies on the surface.

The mass leaving the volume per unit time is the surface integral over the surface S of $\rho\vec{v} \cdot \vec{n}$, and from the law of conservation one gets:

$$\frac{\partial}{\partial t} \iiint_V \rho \, d\vec{x} = -\iint_S \rho\vec{v} \cdot \vec{n} \, d\vec{x_S} \tag{3.2}$$

After applying the Gauss theorem and some mathematical manipulations, the differential equation for for conservation of mass in fluid is obtained:

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot (\rho\vec{v}) = 0 \tag{3.3}$$

3.1.2 Euler's equation of motion for a fluid

The second law of Newton states:

> The acceleration of an object as produced by a net force is directly proportional to the magnitude of the net force, in the same direction as the net force, and inversely proportional to the mass of the object.
>
> $$\frac{\partial \vec{v}}{\partial t} = \frac{\vec{f}}{m} \tag{3.4}$$

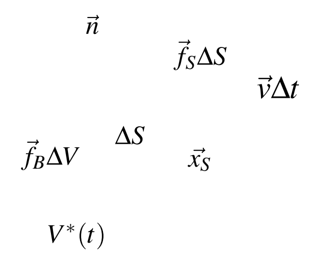


Figure 3.2: Forces acting on a particle occupying volume V^*. Every particle located on on the surface at coordinates $\vec{x}_S$ moves with velocity $\vec{v}(\vec{x}_S, t)$. The acting forces are the surface force per unit area $\vec{f}_S$ and the body force per unit volume $\vec{f}_B$.

This law was applied by Euler for fluids. Consider Figure 3.2. A fluid particle consists of the fluid inside some moving volume $V^*(t)$. Each point of the surface of the volume is moving with a local fluid velocity $\vec{v}(\vec{x}_S, t)$. The Newton's law for the fluid particle can be expressed as:

$$\frac{d}{dt} \iiint\limits_{V^*} \rho \vec{v} \, d\vec{x} = \iint\limits_{S^*} \vec{f}_S d\vec{x}_S + \iiint\limits_{V^*} \vec{f}_B \, d\vec{x} \tag{3.5}$$

Here $\vec{f}_S$ is the apparent surface force per unit area, and $\vec{f}_B$ is the body force, *e.g.*that due to the gravity. The body force is assumed to have a negligible influence. If an *ideal* fluid with no viscosity is assumed, then the surface force is directed normally into the surface S^* and is given by:

$$\vec{f}_S = -\vec{n} p, \tag{3.6}$$

where p is the pressure[1]. The negative sign in the equation follows from the third law of Newton.

3.1.3 Pressure-density relations

A general formula for the relation between the pressure and density can be expressed as:

$$p = p(\rho) \tag{3.7}$$

Many assumptions have been made, but it was Laplace who applied the simple principle that sound propagation occurs with negligible internal heat flow. For a gas, with constant heat coefficients, and for which the pressure is proportional to the density, this principle leads to the relation:

$$p = K\rho^{\gamma} \tag{3.8}$$

3.1.4 Equations of linear acoustics

As previously stated the sound waves represent a propagation of a perturbation of the ambient state of the field. Usually their magnitude is significantly smaller compared to the ambient

[1] $\vec{f}_S$ is defined per unit area.

values characterizing the medium ($p_0, \rho_0, \vec{v}_0$). The ambient variables satisfy the fluid-dynamic equations, but in the presence of disturbance one has:

$$\begin{aligned} p &= p_0 + p' \\ \rho &= \rho_0 + \rho' \\ \vec{v} &= \vec{v}_0 + \vec{v}' \end{aligned} \tag{3.9}$$

where p' and ρ' represent the acoustic contributions to the overall pressure and density fields.

The following assumptions are made:

- The medium is homogeneous. The values of the variables describing the medium are independent of their position.
- The medium is stationary. The properties of the medium are independent of time.
- The ambient velocity $\vec{v}_0$ is zero. There is no flow of matter.
- The process is adiabatic. There is no heat flow in the medium.

Having the above assumptions, the linear approximation of the laws governing the propagation of waves in a fluid can be obtained:

$$\begin{aligned} \text{Conservation of mass} \quad & \frac{\partial \rho'}{\partial t} + \rho_0 \nabla \cdot \vec{v}' = 0 \\ \text{Fluid motion} \quad & \rho_0 \frac{\partial \vec{v}'}{\partial t} = -\nabla p' \\ \text{Adiabatic process} \quad & p' = c^2 \rho' \\ & c^2 = \left(\frac{\partial p}{\partial \rho} \right)_0 \end{aligned} \tag{3.10}$$

In the above equations c is the speed of sound in medium for longitudinal waves.

Further in the thesis the primes will be omitted for notational simplicity.

3.1.5 The wave equation

The propagation of sound is subject to the equations presented in the previous section. Consider the equation of preservation of mass:

$$\frac{\partial \rho}{\partial t} + \rho_0 \vec{\nabla} \cdot \vec{v} = 0$$

By substituting $\rho = p/c^2$ from the pressure-density relations one gets:

$$\frac{1}{c^2} \frac{\partial p}{\partial t} + \rho_0 \vec{\nabla} \cdot \vec{v} = 0$$

Differentiating both sides of the equation with respect to time gives:

$$\begin{aligned} \frac{\partial}{\partial t} \left(\frac{1}{c^2} \frac{\partial p}{\partial t} + \rho_0 \vec{\nabla} \cdot \vec{v} \right) &= 0 \\ \frac{1}{c^2} \frac{\partial^2 p}{\partial t^2} + \rho_0 \vec{\nabla} \cdot \frac{\partial \vec{v}}{\partial t} &= 0 \end{aligned}$$

Substituting the equation of fluid motion in the above result gives the wave equation for pressure:

$$\nabla^2 p - \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2} = 0 \tag{3.11}$$

The equation can be also expressed for the other two variables: ρ and $\vec{v}$. If these are to be calculated, then two additional equations must be solved. To simplify the calculations a new variable, the velocity potential $\Phi(\vec{x},t)$, is introduced. It is not a physical variable, but it can lead to an alternative formulation of the wave equation. In order to use the velocity potential an assumption that the acoustic field has no curl must be made:

$$\overrightarrow{\nabla} \times \vec{v} \equiv 0 \tag{3.12}$$

Usually this condition is satisfied by the sound fields. The velocity potential is defined by:

$$\vec{v}(\vec{x},t) = \overrightarrow{\nabla}\Phi(\vec{x},t) \tag{3.13}$$

The relation between velocity potential Φ and the pressure p is given by:

$$p(\vec{x},t) = -\rho_0 \frac{\partial \Phi(\vec{x},t)}{\partial t} \tag{3.14}$$

The wave equation (3.11) becomes:

$$\boxed{\nabla^2 \Phi(\vec{x},t) - \frac{1}{c^2}\frac{\partial^2 \Phi}{\partial t^2} = 0} \tag{3.15}$$

3.2 Solutions to the wave equation

Two classical methods for solving the wave equation exist, those of d'Lambert and Bernoulli. The former expresses the solution as a sum of two wave fields - one converging and one diverging. The latter assumes that the equation can be written as a product of functions that depend only on one variable.

In the following we will assume that the source is a delta function in space oscillating at a constant angular frequency $\omega = 2\pi f$.

3.2.1 Solution of the wave equation in Cartesian coordinates

According to Bernoulli's method the function $\Phi(\vec{x},t) = \Phi(x,y,z,t)$ must be assumed as a product of 4 functions:

$$\Phi(x,y,z,t) = f(x)g(y)h(z)p(t) \tag{3.16}$$

For simplicity let's initially assume (see Section 2.2 in [22]) that $\Phi(x,y,z,t)$ has a complex exponential form:

$$\dot{\Phi}(x,y,z,t) = \dot{F}\exp\left(j(\omega t - (k_x x + k_y y + k_z z))\right), \tag{3.17}$$

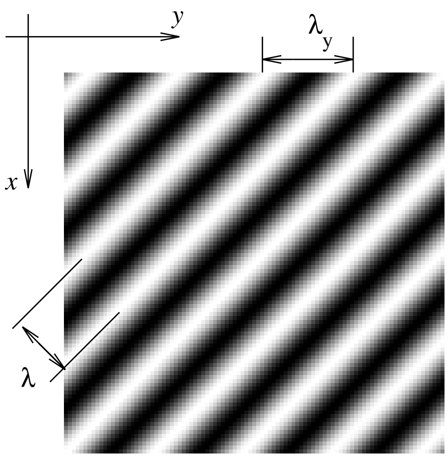


Figure 3.3: A snapshot of pressure distribution of a 2D plane wave. The gray level is proportional to the pressure magnitude.

where $\dot{F}$ is a complex constant and k_x, k_y, k_z and ω are real constants with $\omega \geq 0$. Substituting this form in the wave equation one gets:

$$k_x^2\dot{\Phi}(x,y,z,t) + k_y^2\dot{\Phi}(x,y,z,t) + k_z^2\dot{\Phi}(x,y,z,t) = \frac{\omega^2}{c^2}\dot{\Phi}(x,y,z,t) \tag{3.18}$$

$$\Downarrow$$

$$k_x^2 + k_y^2 + k_z^2 = \frac{\omega^2}{c^2} \tag{3.19}$$

As long as the above constraint is satisfied, signals with the form $\dot{F}\exp\{j[\omega t - (k_x x + k_y y + k_z z)]\}$ satisfy the wave equation (3.15).

This solution is known as a monochromatic plane wave and for a given point in space with coordinates (x_0, y_0, z_0) it becomes:

$$\dot{\Phi}(x_0,y_0,z_0,t) = \dot{F}\exp\left(j[\underbrace{\omega}_{\text{frequency}} t - \underbrace{(k_x x_0, + k_y y_0 + k_z z_0)}_{\text{phase}}]\right) \tag{3.20}$$

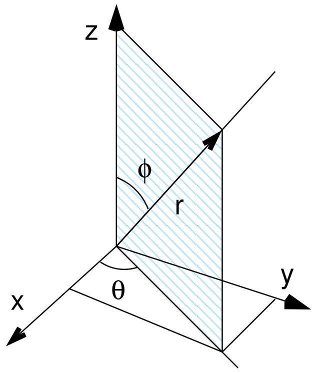
The observable signal at (x_0, y_0, z_0) is a complex exponential with frequency ω. The wave is a *plane wave* because the phase is the same at the points lying on the plane given by $k_x x + k_y y + k_z z = C$, where C is a constant. Figure 3.3 shows a snapshot of a plane wave. The time is frozen ($t = \text{const}$) and a 2D cross-section ($z = \text{const}$) is shown. The gray level in the image is proportional to the pressure. The wave propagates in a plane perpendicular to the z axis at an angle of 45° to the x axis ($k_z = 0$, $k_x = k_y = k\cos\frac{\pi}{4}$). The distance between two points that have the same phase is called *wavelength* and is equal to:

$$\lambda = \frac{c}{f}. \tag{3.21}$$

The wave number and the wavelength are related through the equation:

$$k = \frac{2\pi}{\lambda} \tag{3.22}$$

If the pressure is taken along one of the axis, say y, one will get a one-dimensional pressure distribution with wavelength $\lambda_y = k_y/(2\pi)$.



$$x = r\sin\phi\cos\theta$$
$$y = r\sin\phi\sin\theta$$
$$z = r\cos\phi$$

Figure 3.4: Relation between spherical and Cartesian coordinates

3.2.2 Solution of the wave equation in polar coordinates

The wave equation in spherical coordinates can be written as [22]:

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\dot{\Phi}}{\partial r}\right)+\frac{1}{r^2\sin\phi}\frac{\partial}{\partial\phi}\left(\sin\phi\frac{\partial\dot{\Phi}}{\partial\phi}\right)+\frac{1}{r^2\sin^2\phi}\frac{\partial^2\dot{\Phi}}{\partial\theta^2}=\frac{1}{c^2}\frac{\partial^2\dot{\Phi}}{\partial t^2}. \tag{3.23}$$

where $\dot{\Phi}$ is the velocity potential which is a function of the distance to the origin of the coordinate system r and the azimuth and elevation angles θ and ϕ, and the time, $\dot{\Phi} = \dot{\Phi}(r,\theta,\phi,t)$. The general solutions are rather complicated. A solution of particular interest is when Φ is independent of the angles θ and ϕ:

$$\frac{\partial\dot{\Phi}}{\partial\theta}=\frac{\partial\dot{\Phi}}{\partial\phi}=0. \tag{3.24}$$

The wave equation reduces to :

$$\frac{\partial^2}{\partial r^2}(r\dot{\Phi})=\frac{1}{c^2}\frac{\partial^2}{\partial t^2}(r\dot{\Phi}). \tag{3.25}$$

Solutions to this equation are the *spherical* waves, which in complex form are given by:

$$\dot{\Phi}(r,t)=\frac{\dot{F}}{r}\exp j(\omega t-kr), \tag{3.26}$$

where $\dot{F}$ is a complex amplitude, r is the distance to traveled by the wave and k is the wave number:

$$k^2=\frac{\omega^2}{c^2}. \tag{3.27}$$

The waves can be either diverging (going away from the source) or converging. The diverging wave can be expressed by:

$$\dot{\Phi}(r,t)=\frac{\dot{F}}{r}\exp(j\omega(t-\frac{r}{c})). \tag{3.28}$$

Figure 3.5 shows a 2D snapshot of the pressure distribution. The amplitude of the pressure is normalized. The figure is not exact, since at the source of the spherical wave the radius r tends to zero ($r \rightarrow 0$), and the function should go to infinity.

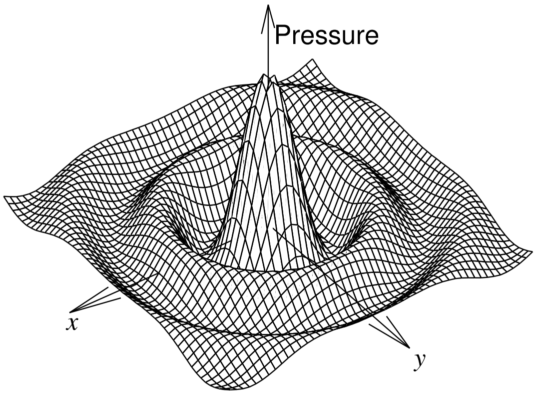


Figure 3.5: A snapshot of the 2D pressure distribution of a spherical wave. There is discontinuity at the center of the coordinate system.

3.3 Theory of radiation and diffraction

The wave equation describes the propagation of waves in the free space. To describe the radiation of ultrasound the source and propagation are linked together through boundary conditions. There are two classical solutions to the problem: the Kirchoff and the Rayleigh-Sommerfeld theories of diffraction [23]. The simulation program Field II used in this thesis is based on a method derived from the Rayleigh-Sommerfeld theory and this will be examined in the next section. Then a brief description of the numerical calculation of the pressure field through the means of spatial impulse responses will be given. Finally some approximation to the solutions will be presented.

3.3.1 Rayleigh integral

Figure 3.6 shows the basic setup of the problem. The aperture is planar and lies on an infinite rigid baffle, on which the velocity to the plane is zeros except for the aperture. The problem is to find the field in a point with spatial coordinates $\vec{x}_1$. The coordinates of a point lying on the aperture is denoted with $\vec{x}_0$. The velocity potential at the point is given by the Rayleigh integral [21, 23]:

$$\Phi(\vec{x}_1,t) = \iint\limits_{\Sigma} \frac{v_n\left(\vec{x}_0, t-\frac{|\vec{x}_1-\vec{x}_0|}{c}\right)}{2\pi|\vec{x}_1-\vec{x}_0|}\, d\vec{x}_0, \tag{3.29}$$

where Σ is the area of the aperture. The Rayleigh integral is basically a statement of Huygens' principle that the field is found by summing the contributions from all infinitely small area elements that make up the aperture. The pressure is given by:

$$p(\vec{x}_1,t) = \frac{\rho_0}{2\pi} \iint\limits_{\Sigma} \frac{\frac{\partial}{\partial t} v_n\left(\vec{x}_0, t-\frac{|\vec{x}_1-\vec{x}_0|}{c}\right)}{|\vec{x}_1-\vec{x}_0|}\, d\vec{x}_0 \tag{3.30}$$

This solution is given for a rigid baffle. There is a solution for a soft baffle which requires that the pressure on the baffle is equal to zero. The solution is the same except for an obliquity term:

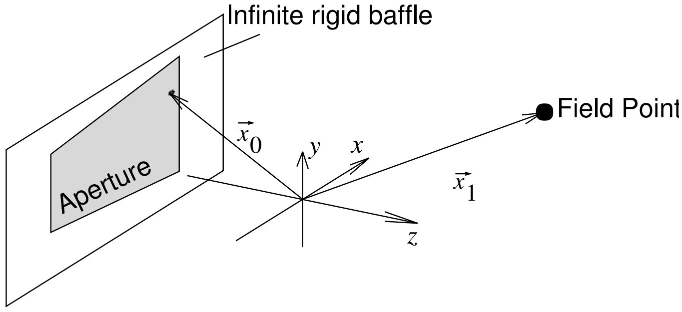


Figure 3.6: Position of transducer, field point and coordinate system.

$$\Phi(\vec{x}_1,t) = \iint\limits_{\Sigma} \frac{v_n\left(\vec{x}_0, t - \frac{|\vec{x}_1-\vec{x}_0|}{c}\right)}{2\pi|\vec{x}_1-\vec{x}_0|} \cdot \cos\vartheta \, d\vec{x}_0, \tag{3.31}$$

where ϑ is the angle between the vector $\vec{x}_1$ and the normal vector to the aperture plane. The pressure is given by:

$$p(\vec{x}_1,t) = \frac{\rho_0}{2\pi} \iint\limits_{\Sigma} \frac{\frac{\partial}{\partial t} v_n\left(\vec{x}_0, t - \frac{|\vec{x}_1-\vec{x}_0|}{c}\right)}{|\vec{x}_1-\vec{x}_0|} \cdot \cos\vartheta \, d\vec{x}_0 \tag{3.32}$$

Usually the true value of the field is found between the values of the two solutions.

3.3.2 Spatial impulse responses

The spatial impulse response method applies a linear system theory to the wave equation and it separates the temporal and spatial characteristics of the acoustical field, which is apprehended as a spatial filter. The following considerations are based on [24]. The assumptions are that the medium is non-attenuating and homogeneous, and that the normal component v_n of the velocity $\vec{v}$ is uniform across the aperture. The excitation can be separated from the transducer geometry by using a convolution:

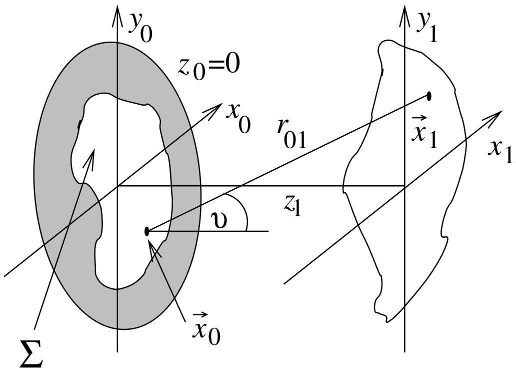
$$\Phi(\vec{x}_1,t) = \iint\limits_{\Sigma} \int\limits_{t} \frac{v_n(\vec{x}_0,t)\delta\left(\vec{x}_0, t - \frac{|\vec{x}_1-\vec{x}_0|}{c}\right)}{2\pi|\vec{x}_1-\vec{x}_0|} \, dt \, d\vec{x}_0, \tag{3.33}$$

where δ is the Dirac delta function. Since the velocity is uniform across the aperture the integral can be written out as:

$$\Phi(\vec{x}_1,t) = v_n(t) \underset{t}{*} \iint\limits_{\Sigma} \frac{\delta\left(t - \frac{|\vec{x}_1-\vec{x}_0|}{c}\right)}{2\pi|\vec{x}_1-\vec{x}_0|} \, d\vec{x}_0, \tag{3.34}$$

where $\underset{t}{*}$ denotes convolution in time. The integral in the equation:

$$h(\vec{x}_1,t) = \iint\limits_{\Sigma} \frac{\delta\left(t - \frac{|\vec{x}_1-\vec{x}_0|}{c}\right)}{2\pi|\vec{x}_1-\vec{x}_0|} \, d\vec{x}_0 \tag{3.35}$$



$$r_{01} = |\vec{x}_1 - \vec{x}_0|$$
$$r_{01} = \sqrt{z_1^2 + (x_1 - x_0)^2 + (y_1 - y_0)^2}$$
$$\cos\vartheta = \frac{z_1}{r_{01}}$$

Figure 3.7: Geometry for Fresnel and Fraunhofer approximations for a radiation problem.

is called the spatial impulse response because it varies with the relative position of the transducer and the point in space. The pressure as a function of time at one given point in space can be found by:

$$p(\vec{x}_1,t) = \rho_0 \frac{\partial v_n}{\partial t} * h(\vec{x}_1,t). \tag{3.36}$$

The use of spatial impulse response is a powerful technique and analytical solutions for several geometries exist such as concave transducers [25], triangle apertures [26, 27], or apertures that can be defined as planar polygons [28].

3.3.3 Fresnel and Fraunhofer approximations

Simplification to the Rayleigh integral can be made by using some geometrical considerations. The derivations are carried out for the continuous wave (CW) case. The Rayleigh integral for the CW case can be found from (3.34) using a Fourier transform and the following relations:

$$h(t) * g(t) \rightleftharpoons H(f) \cdot G(f) \tag{3.37}$$
$$\delta(t-T) \rightleftharpoons j\omega \exp(-j\omega T) \tag{3.38}$$

Further an assumption will be made that the velocity distribution over the aperture can be defined as a separable function:

$$v_n(\vec{x}_0,t) = v_n(t) a(\vec{x}_0), \tag{3.39}$$

where $a(\vec{x}_0)$ is a weighting function of the contribution of the individual points, known also as "apodization function". The apodization function becomes zero outside the aperture limits.

For a single frequency the temporal function $v_n(t)$ will be assumed to be:

$$v_n(t) = v_0 \exp(j\omega t). \tag{3.40}$$

For a single frequency the Rayleigh integral becomes:

$$\Phi(\vec{x}_1,\omega) = v_0 \exp(j\omega t) \frac{jk}{2\pi} \iint\limits_{-\infty}^{\infty} a(x_0,y_0) \frac{\exp(-jkr_{01})}{r_{01}}\, dx_0\, dy_0 \tag{3.41}$$

where r_{01} is the distance between the field point and a point in aperture. Figure 3.7 illustrates the geometry of the radiation problem. The aperture lies on a plane parallel to the (x,y) axes of the coordinate system and a depth $z=0$. The coordinates in this plane are denoted with sub-script $_0$. The radiated field is sought on a plane parallel to the emitting aperture. The distance between the two planes is equal to z_1. The coordinates of the points in the plane where the radiated field is calculated have a sub-script $_1$. So $\vec{x}_0 = (x_0, y_0, 0)$ is a source point, and $\vec{x}_1 = (x_1, y_1, z_1 = \text{const})$ is a field point. The distance between the two points is:

$$r_{01} = |\vec{x}_1 - \vec{x}_0| = \sqrt{z_1^2 + (x_1 - x_0)^2 + (y_1 - y_0)^2}. \tag{3.42}$$

The Fresnel approximation assumes that the changes in the amplitude are not as strongly expressed as the changes in the phase of the field [23]. The distance in the denominator is then approximated:

$$r_{01} \approx z_1. \tag{3.43}$$

The second approximation is done by extending the term r_{01} which is part of the exponent into Taylor series:

$$r_{01} \approx z_1 \left(1 + \frac{1}{2}\left(\frac{x_1 - x_0}{z_1}\right)^2 + \frac{1}{2}\left(\frac{y_1 - y_0}{z_1}\right)^2 \right). \tag{3.44}$$

In this way the spherical waves are substituted with parabolic waves. The velocity potential at a point in the plane $(x_1, y_1, z_1 = \text{const})$ becomes

$$\begin{aligned} \Phi(x_1, y_1, z_1, \omega) = & v_0 \exp(j\omega t) \frac{jk}{2\pi} \frac{\exp(-jkz_1)}{z_1} \\ & \times \iint_{-\infty}^{\infty} \exp\left(-\frac{jk}{2z_1}((x_1 - x_0)^2 + (y_1 - y_0)^2) \right) \, dx_0 \, dy_0, \end{aligned} \tag{3.45}$$

This result is known as the Fresnel diffraction integral. When this approximation is valid the observer is said to be in the region of Fresnel diffraction, or equivalently, in the *near field* of the aperture. This is the usual working region for the ultrasound scanners.

Rearranging the terms in (3.45) one gets:

$$\begin{aligned} \Phi(x_1, y_1, z_1, \omega) = & v_0 \exp(j\omega t) \frac{jk}{2\pi} \frac{\exp(-jkz_1)}{z_1} \\ & \times \iint_{-\infty}^{\infty} a(x_0, y_0) \underbrace{\exp\left(-\frac{jk}{2z_1}((x_1 - x_0)^2 + (y_1 - y_0)^2) \right)}_{h'(x_0, y_0; x_1, y_1)} \, dx_0 \, dy_0 \end{aligned} \tag{3.46}$$

The Fraunhofer assumption is that the distance between the aperture plane and the observation plane is much greater than the dimensions of the aperture function, i.e.:

$$z >> \frac{k \max(x_0^2 + y_0^2)}{2}. \tag{3.47}$$

By developing the term $(x_1 - x_0)^2 + (y_1 - y_0)^2$, the exponential in the integral becomes:

$$h'(x_0, y_0; x_1, y_1) = \exp\left(-jk \frac{x_1^2 + y_1^2}{2z_1} \right) \underbrace{\exp\left(-jk \frac{x_0^2 + x_1^2}{2z_1} \right)}_{\approx 1} \exp\left(jk \frac{x_1 x_0 + y_0 y_1}{z_1} \right). \tag{3.48}$$

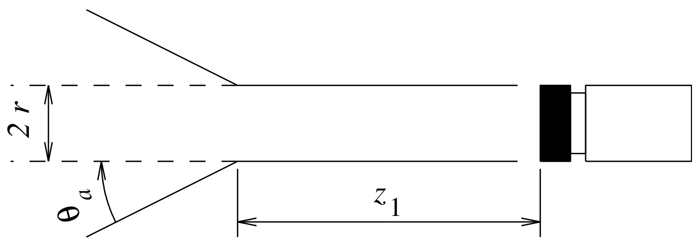


Figure 3.8: Idealized continuous pressure field for flat, round transducer (aperture).

The Fraunhofer approximation becomes (skipping the temporal terms):

$$\Phi(x_1,y_1,z_1) = \frac{jk\exp(-jkz_1)\exp\left(-j\frac{k}{2z_1}(x_1^2+y_1^2)\right)}{2\pi z_1} \times \iint_{-\infty}^{\infty} a(x_0,y_0)\exp\left(jk\left(\frac{x_1x_0}{z_1}+\frac{y_1y_0}{z_1}\right)\right)\,dx_0\,dy_0 \tag{3.49}$$

Aside from the quadratic phase term outside the integral, this result is the Fourier transform[2] of the apodization function $a(x_0,y_0)$ at frequencies:

$$f_x = \frac{x_1}{\lambda z_1} \qquad f_y = \frac{y_1}{\lambda z_1} \tag{3.50}$$

The region in which the Fraunhofer approximation is valid is known as the *far field*. The depth z_1 usually used as the border (see Figure 3.8) of the transition from near to far field is given by the following:

$$z_1 = \frac{r^2}{\lambda}, \tag{3.51}$$

where $2r$ is the lateral size of the aperture. In the near field the main part of the beam is confined to lie within the extent of the transducer surface. In the far field the beam starts to diverge. The angle of divergence is:

$$\theta_a = \arcsin\left(0.61\frac{\lambda}{r}\right). \tag{3.52}$$

The Fraunhofer approximation gives a very nice Fourier relation between the apodization function and the generated field. A similar Fourier relation between the aperture apodization function and the radiated field exists also for the Fresnel zone at the focal plane. The derivation is given in Appendix A. Sometimes it is said that the focusing brings the far field conditions in the near field.

3.4 Propagation in tissue

The methods concerned in this thesis are the pulse-echo methods, meaning that a pressure wave is transmitted into the region under investigation and the received echo is displayed. The

[2]The sign in the exponent term can be either plus or minus depending on the chosen positive direction of the the phase. For a discussion on this subject see *Introduction to Fourier optics* [23].

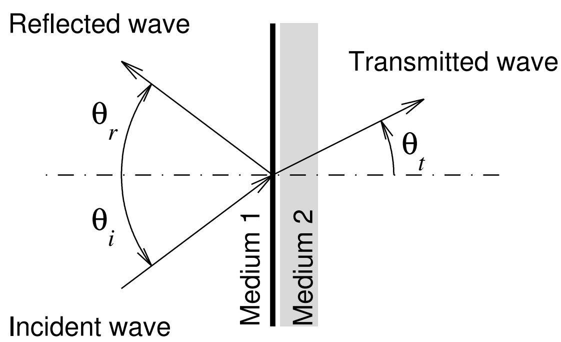


Figure 3.9: Transmission and reflection of a plane wave propagating obliquely to the reflecting surface.

received echo is based on two distinct phenomena: reflection and scattering.

3.4.1 Reflection

Reflection occurs at the border of two regions with different acoustic impedances. For a monochromatic wave the relation between particle velocity u and the pressure p is given through the acoustic impedance Z [21]:

$$u = \frac{p}{Z}. \tag{3.53}$$

For a plane progressing wave the impedance is:

$$Z = \rho_0 c. \tag{3.54}$$

The situation of reflection and transmission of a plane wave propagating obliquely to the reflecting surface is shown in Figure 3.9. If the first medium has a speed of sound c_1 and second medium c_2, then the angles of transmitted and reflected waves are described by Snell's law:

$$\frac{c_1}{c_2} = \frac{\sin\theta_i}{\sin\theta_t}, \tag{3.55}$$

where θ_i and θ_t are the angles of the incident and transmitted waves, respectively. If the angles are measured with respect to the normal vector to reflecting surface then:

$$\theta_r = -\theta_i. \tag{3.56}$$

The pressure amplitude transmission coefficient is given by:

$$\frac{p_t}{p_i} = \frac{2Z_2\cos\theta_i}{Z_2\cos\theta_i + Z_1\cos\theta_t}, \tag{3.57}$$

where Z_1 and Z_2 are the acoustic impedances of the first and the second media, respectively, and p_i and p_t are the respective amplitudes of the pressure.

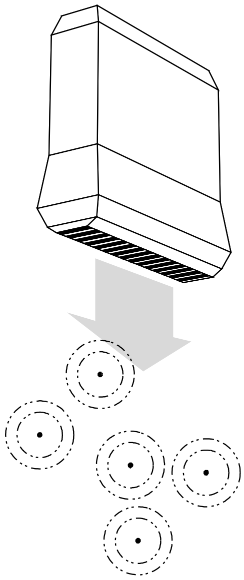


Figure 3.10: Illustration of multiple scattering. Rayleigh scattering is assumed for the point scatterers and Born approximation for the scattering medium.

3.4.2 Scattering

A single scatterer and the point spread function.

The spatial pulse-echo response for a single scatterer corresponds to the convolution of the spatial impulse responses of the transmit and receive apertures [29]. This can be easily justified since a point can be represented as a delta function $\delta(x_p, y_p, z_p)$ and the system is assumed to be linear. In the rest of the dissertation the pulse-echo response of a point will be called *point spread function*. The point spread function can be in space $\mathbf{P}(x_p, y_p, z_p)$, or in space and time. In the latter case this usually represents a collection of the responses along a line of points.

As discussed in Section 3.3, there is a Fourier relation between the apodization of a focused transducer and the point spread function in the focus. The point spread functions must be defined on a spherical surface, so often they will be given as a function of angle.

Multiple scatterers

Figure 3.10 shows the measurement situation. A transducer insonifies a group of scatterers (the operation is shown with a thick gray arrow). The wave reaches the scatterers and they start to vibrate, becoming omnidirectional sources of spherical waves (Rayleigh scattering). The scattering is assumed to be weak, i.e. a wave generated by a scatterer cannot be scattered again by another scatterer (Born approximation [21]). The medium is assumed to be linear and homogeneous, and the total response of the field is the sum of the responses of the individual scatterers. In the thesis yet another approximation will be used - the one of separability of the excitation from the transducer geometry. The received response becomes [30]:

$$p_r(\vec{x},t) = v_{pe}(t) \underset{t}{*} f_m(\vec{x}) \underset{x}{*} h_{pe}(\vec{x},t), \tag{3.58}$$

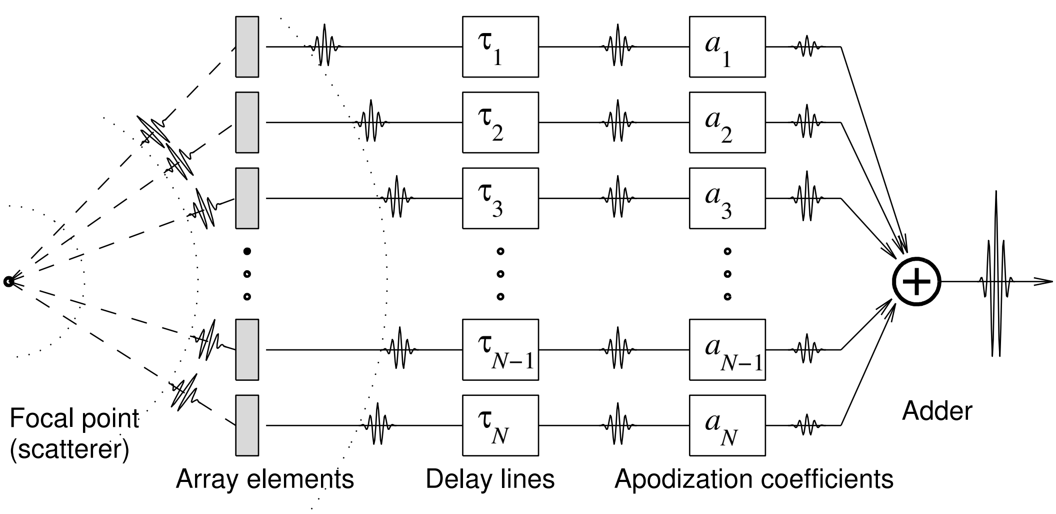


Figure 3.11: Delay and sum beamforming.

where v_{pe} includes the transducer excitation and the impulse responses of the transmitting and receiving apertures, f_m is the scatterer map, and h_{pe} is the pulse echo spatial response. The scatterer function f_m is a function of the two phenomena causing the scattering - the change in density and speed of sound:

$$f_m(\vec{x}) = \frac{\Delta\rho(\vec{x})}{\rho_0} - \frac{2\Delta c(\vec{x})}{c_0}. \tag{3.59}$$

3.4.3 Attenuation

The human tissue is a lossy medium and the ultrasound pulse is attenuated as it propagates in it. This attenuation is not described by the linear wave equation. Most of the energy loss is due to absorption. The general solution is rather complicated, but a rather simple approximation can be made expressing the losses in dB/(MHz·cm) [31].

3.5 Beamforming

This section provides a description of the beamforming used in the modern scanners. The type of beamforming used in them is known also as a time domain beamforming. The frequency domain beamforming methods [22, 32, 33] are neither spread in the medical ultrasound scanners, nor used in the synthetic aperture algorithms presented in the rest of the thesis, and will not be described here.

Figure 3.11 shows a typical delay and sum beamformer [22, 34, 35, 36, 37]. The depicted array is a *linear* array of piezoelectric transducer elements. It transmits a sound pulse into the body and receives echos from the scattering structures within. The transmit and receive signals can be individually delayed in time, hence the term *phased* array. Through the time delays the beam can be steered in a given direction and focused at a given axial distance both in transmit and receive. Figure 3.11 shows focusing during reception. Using simple geometric relations the transducer can be focused at any point. There is a Fourier relation between the aperture weighting function and the point spread function at the focal depth. The point spread function

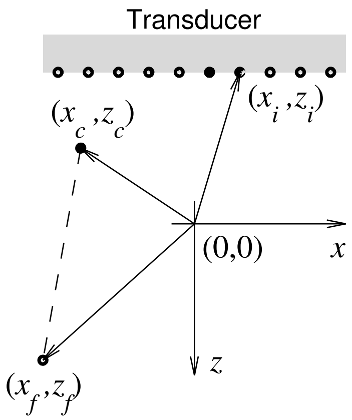


Figure 3.12: 2D geometry for calculating the delay profiles.

determines the imaging capabilities of the system. It consists of a main lobe and side lobes. The narrower the main lobe, the higher the resolution of the system, hence larger arrays are desirable. The side lobes are caused by the finite size of the aperture (the edges). Applying weighting on the aperture function has the same effect as applying weighting on the Fourier transform of a signal - the side lobes decrease. The beamformation procedure becomes (see: Figure 3.11):

$$s(t) = \sum_{i=1}^{N_{xdc}} a_i r_i(t - \tau_i), \tag{3.60}$$

where N_{xdc} is the number of transducer elements, a_i are the apodization coefficients and τ_i are the applied delays. Usually in transmit the focus is fixed. In receive, however, the focus can be changed as a function of time thus "tracking" the current position of the wave front. This is usually denoted as dynamic focusing. The modern digital scanners can change the delays for every sample in the beamformed scan lines $s(t)$. There are different ways to define the scan geometry. The approach adopted in this work and used in the Beamformation Toolbox [38], and in Field II [19] is to define the scan lines using a "focus center" and a "focus point" or a direction in the case of dynamic focusing.

Figure 3.12 shows a 2D geometry for determining the delays τ. The equations will, however, be done for 3D geometry. The assumption is that a plane wave confined in space is transmitted from the center $\vec{x}_c = (x_c, y_c, z_c)$ and propagates along the line defined by the center point and the focal point $\vec{x}_f = (x_f, y_f, z_f)$. The transducer element i with coordinates $\vec{x}_i = (x_i, y_i, z_i)$ transmits sound pulse. The pulse from the transducer element must arrive at the same time with the imaginary pulse transmitted by the center point.

The delay with respect to the trigger is given by:

$$\tau_i = \frac{|\vec{x}_c - \vec{x}_f| - |\vec{x}_i - \vec{x}_f|}{c}. \tag{3.61}$$

Usually the origin of the coordinate system lies in the geometric center of the transducer array. It is possible to have multiple focusing points (focal zones) along the beam. In this case it is necessary to define the starting times at which the next delay profiles will be active. Usually the time is set as the time of flight of ultrasound pulse from $\vec{x}_c$ to $\vec{x}_f$. The fully dynamic focusing implies that the delay profile is recalculated for every point (sample) along the scan line. In this

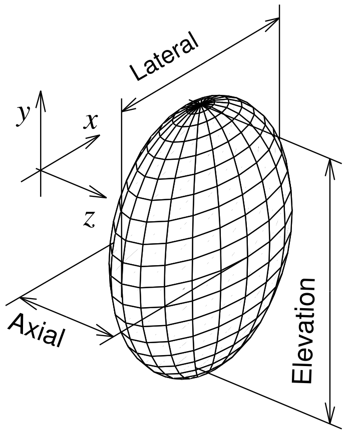


Figure 3.13: The system is characterized by axial, lateral and elevation resolution.

case the scan lines are defined by a center (origin of the line) and a direction determined by the azimuth and elevation angles.

In digital systems the received signal $r(t)$ is sampled at a sampling frequency f_s. The delays that can exactly be generated are multiples of the sampling period $T_s = 1/f_s$. If a delay is needed which is a fraction of T_s, then some interpolation is performed. The nearest neighbor interpolation is the simplest, but is not adequate in most cases [39]. The linear interpolation meets most of the practical needs of B-mode imaging, and is the one implemented in the experimental system RASMUS [40, 41, 42], which was developed at CFU. This is the type of interpolation used in the Beamformation Toolbox when calculation speed is necessary [43]. For some applications the linear interpolation is not adequate. One such application is the motion compensation scheme presented in Chapter 12. There the interpolation is done by splitting the delays into two parts - a coarse delay and a fine delay. The coarse delay is a multiple of the sampling period and is achieved by shifting the signal with a number of samples. The fine delay is implemented by filtering the shifted signal with a FIR fractional delay filter [44]. The filters form a filter bank, whose size depends on how precise the fractional delay should be. It is also possible to achieve fractional delays with theoretically unlimited precision if the filter coefficient is calculated at the moment of beamformation (see http://www-ccrma.stanford.edu/∼jos/resample/). The Beamformation Toolbox provides means for creating a filter bank.

As discussed previously, the focus in transmit is fixed. There exist approaches to compensate for it [45, 46], but they are not what is meant to be “conventional” focusing.

3.6 Resolution

Figure 3.13 shows the three dimensional nature of the resolution cell of an ultrasound system. The axial resolution is inversely proportional to the bandwidth of the system[47]. If the transducer is not 1.5, 1.75 or 2D matrix array [48, 11], the resolution in the elevation plane is determined by the geometry of the transducer. The resolution in the azimuth plane is the only one dependent on the focusing algorithms and is usually the one considered in the thesis. For conventional B-mode scans, the information is considered as gathered only from the $(x-z)$ plane. There are many ways to determine the azimuth resolution [22]. In this thesis the resolution will be stated as the beam width at a given level, most often at -6 dB.

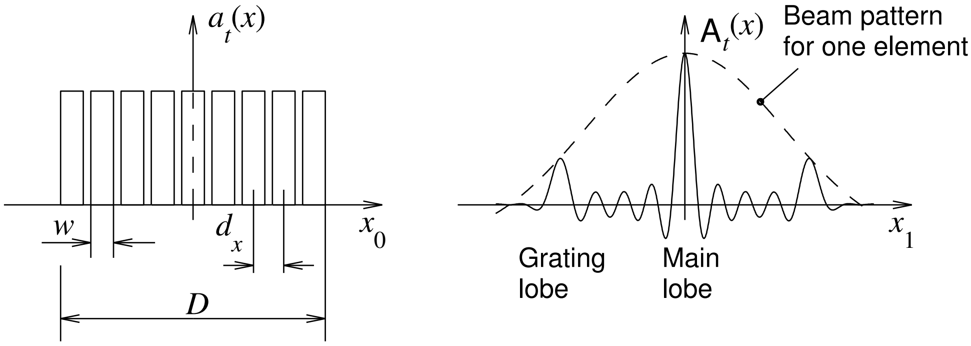


Figure 3.14: (Left) An aperture of size D composed of elements with size w whose centers are positioned at distance d_x apart is modeled as $\Pi(x_0/D)\Pi(x_0/w)*\text{III}(x/d_x)$. The radiation pattern (right) is $\text{sinc}(x_1/(\lambda z_1)w)\cdot[\text{sinc}(x_1/(\lambda z_1)D)*\text{III}(x_1/(\lambda z_1)d_x)]$

The systems are assumed to be linear, and therefore they can be characterized fully by the point spread function. Usually this is the point spread function along the azimuth direction (along x) for linear scans, and along an arc for sector scans. The plots shown will be obtained from the projection of the maximum of the point spread function:

$$\mathsf{A}_{t/r}(x) = \max_z p_r(x,z). \tag{3.62}$$

For the 3D scans, the point spread function will be the projection along the axial direction on a 2D surface. Usually what matters are the relative levels of the point spread function at a given spatial position relative to the level at the focal point. Hence the plots will be normalized, and most often logarithmically compressed to emphasize the details.

3.7 Frequency domain representation of the system

As seen from Section 3.3.2, the linear systems approach can be used in ultrasound systems to calculate the pulse echo response of the system. Frequency domain methods are a powerful tool in engineering, and it turns out that they can be also used for ultrasound scanners.

In Section 3.3.3 and in Appendix A it is shown, that there is a Fourier relation between the apodization function of the aperture and its radiation pattern in the focal plane (at small angles the spherical cap can be approximated with the tangent plane). If the scale factors in (A.9) are skipped, then the radiation patterns of the transmit and receive apertures can be expressed as:

$$\mathsf{A}_t(x_1,y_1;z_f) = \mathcal{F}\{a_t(x_0,y_0)\} \tag{3.63}$$

$$\mathsf{A}_r(x_1,y_1;z_f) = \mathcal{F}\{a_r(x_0,y_0)\} \tag{3.64}$$

where A_t and A_r are the radiation patterns of the transmit and receive aperture, respectively, and a_t and a_r their apodization functions. The focal point lies on a spherical cap at a distance z_f away from the transducer. One of these relations is shown in Figure 3.14. The width of the array is D and is modeled by the function $\Pi(x_0/D)$ where:

$$\Pi(x) = \begin{cases} 1 & -\frac{1}{2} \le x \le \frac{1}{2} \\ 0 & \text{otherwise} \end{cases} \tag{3.65}$$

The elements have width w and each of them is modeled by $\Pi(x_0/w)$. The discrete nature of the array is modeled as a comb function $\text{III}(x_0/d_x)$, which is a sum of offset delta functions:

$$\text{III}(x) = \sum_{n=-\infty}^{\infty} \delta(x-n). \tag{3.66}$$

The apodization function of the array is then:

$$a_t(x_0) = \Pi(x_0/D)[\Pi(x_0/w) \underset{x}{*} \text{III}(x_0/d_x)] \tag{3.67}$$

The radiation pattern of $a_t(x)$ is found through the Fourier transform:

$$\Pi(x_0/D)[\Pi(x_0/w) \underset{x}{*} \text{III}(x_0/d_x)] \rightleftharpoons \text{sinc}\left(\frac{x_1}{\lambda z_1}w\right) \cdot \left[\text{sinc}\left(\frac{x_1}{\lambda z_1}D\right) \underset{x}{*} \text{III}\left(\frac{x_1}{\lambda z_1}d_x\right)\right], \tag{3.68}$$

where the scaling coefficients have been skipped for notational simplicity. From Figure 3.14 it can be seen that the radiation pattern consists of a main lobe, side and grating lobes. The width of the main lobe is inversely proportional to the width of the aperture.

The two-way radiation pattern is the product of the transmit and receive radiation patterns:

$$\mathsf{A}_{t/r}(x_1,y_1;z_f) = \mathsf{A}_t(x_1,y_1;z_f) \cdot \mathsf{A}_r(x_1,y_1;z_f). \tag{3.69}$$

A fictional aperture that would have a radiation pattern equal to the two-way radiation pattern of the system will be called *effective aperture* [49, 50, 51]. Using the properties of the Fourier transform the apodization function of the effective aperture can be found by the spatial convolution of the apodization functions of the transmit and receive apertures:

$$a_{t/r}(x_0,y_0) = a_t(x_0,y_0) \underset{x}{*} a_r(x_0,y_0). \tag{3.70}$$

There is another term, ”co-array” [22, 52], which is by principle the same.

A linear system is characterized in frequency domain by its transfer function $H(\omega)$ which is related to the impulse response of the system $h(t)$ by the Fourier transform [53]:

$$H(\omega) \rightleftharpoons h(t). \tag{3.71}$$

An ultrasound system can be characterized in a similar manner, by taking the Fourier transform of the pulse echo response. This characteristics varies in space and is useful mostly for the far field or for the focus. If the imaging system is considered only in the two-dimensional case, in the azimuth plane, then the 2D Fourier transform is involved. The result is given in spatial frequencies $cycles/m$ or alternatively rad/m. In the latter case the frequencies are k_x and k_z. They are related to the wave number k by:

$$k_x^2 + k_z^2 = k^2. \tag{3.72}$$

The bandwidth in the k_x and k_z domains determine the lateral and axial resolutions of the system, respectively. Because the letter k is used for notation this representation of the system is also known as the k-space [54, 55], or alternatively as wave number representation, angular spectrum [23], or plane wave decomposition [37]. The name angular spectrum comes from the fact that the spatial frequencies are related to the direction of propagation of a plane wave which is defined by an angle θ:

$$\begin{aligned} k_x &= k\sin\theta \\ k_z &= k\cos\theta \end{aligned} \tag{3.73}$$

For the 3D case the azimuth angle also comes into play. The term "k-space " will be used in the rest of the thesis.

In Section 3.3.2 it was assumed that for the time function of the excitation can be separated from the transducer geometry. The Fourier transform of a separable function is also a separable function. The radiation pattern of an aperture is related to the apodization function through the Fourier transform. Taking the Fourier transform of the radiation pattern gives a function which is a scaled version of the aperture function. So the lateral band width of the system $[\min(k_x), \max(k_x)]$ is determined by the spatial extent of the aperture. The larger the aperture, the bigger the bandwidth, the higher the resolution. The same applies for the axial resolution. Its bandwidth is determined by the excitation and the impulse response of the transmit and receive apertures. A higher bandwidth transducer gives a higher axial resolution. The ideal system is therefore an all pass filter.

The assumptions are that linear system theory can be applied for the ultrasound systems [24]. As so the k-space of the system can be increased by synthesizing a large effective aperture using synthetic aperture imaging algorithms which are the topic of the rest of this thesis.

Part I

Tissue imaging

CHAPTER

FOUR

Generic synthetic aperture ultrasound imaging

This chapter introduces synthetic aperture ultrasound imaging (SAU) in its "classical" form, the mono-static synthetic aperture imaging. Historically this modality appeared in the 1950s and was first applied in radar systems. A number of reconstruction algorithms have been developed. Some of them are carried out in frequency domain and are well suited for signal processing using optical elements [23]. This was done in the early years of synthetic aperture radar (SAR), when the limitations were imposed by the computational power available. Others reconstruction algorithms are executed in time domain and are computationally more expensive. They have been introduced in SAR in the recent years [33].

The synthetic aperture focusing in its mono-static form is described in [56], where the performance of the system is discussed. This approach is very suitable in the field of non-destructive testing (NDT), and a frequency based reconstruction was implemented by Mayer et al. in [57] and Busse in [58]. The group of O'Donnell has been using synthetic aperture focusing for intra-vascular imaging [59]. A substantial work in the field of SAU was done Ylitalo [60, 61, 62, 63, 64]. He and his colleagues have considered various geometries of the transducer and investigated the signal-to-noise ratio. They have also implemented a real-time system. Karaman and colleagues have studied the application of synthetic aperture focusing for a cheap ultrasound system [65]. Other groups have also applied synthetic aperture focusing to ultrasound [66, 67].

In the next section a simple model for synthetic aperture imaging is presented. Then some reconstruction methods are outlined and their performance in terms of lateral resolution and signal-to-noise ratio (SNR) is discussed.

4.1 Simple model

4.1.1 Acquisition

In the synthetic aperture imaging a single transducer element is used both, in transmit and receive as illustrated in Figure 4.1. Because of the small element size in the azimuth plane, the transmitted wave has a cylindrical wavefront. It propagates in the whole region of interest and the returned signal carries information from all imaging directions. All of the transducer elements are fired one by one. Every point of the examined tissue is viewed from different angles, and the received signals from the different elements have different phases for the same

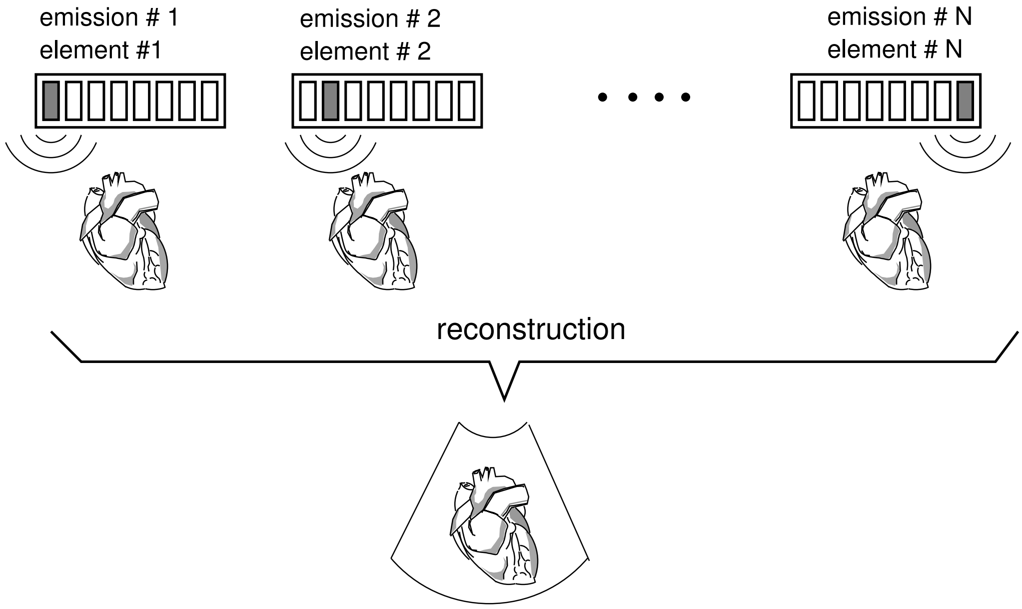


Figure 4.1: Simple model for acquiring synthetic aperture data.

spatial position. After the last emission the signals are processed (the phase information is extracted) and the image of the tissue is reconstructed.

4.1.2 The received signal

In the following some assumptions will be used to simplify the considerations. The region of investigation (the tissue below the transducer) will be a homogeneous, non-dispersive, non-attenuating medium. The electromechanical impulse response of the transducer is a Dirac delta function. Hence the transmitted pressure pulse $p(t)$, and the received signal $r(t)$, are proportional to the generated electrical radio frequency pulse $g(t)$.

Let the elements of a transducer be infinitesimal in the azimuth plane (point sources). The waves that they emit are, in this case, cylindrical. The height of the cylindrical wave is determined by the height of the transducer elements. The considerations can be confined only to the azimuth plane $(x-z)$, and the 2D measurement situation is depicted in Figure 4.2. Consider a point scatterer at spatial coordinates $\vec{x}_p = (x_p, z_p)$. The signal emitted by a transducer element i at coordinates $\vec{x}_i = (x_i, 0)$ propagates as a cylindrical wave. The wave front reaches the point scatterer which in turn becomes a source of a spherical wave. The back-scattered wavefront reaches the transmitting element i at a time instance:

$$\begin{aligned} t_p(\vec{x}_i) &= \frac{2}{c}|\vec{x}_p - \vec{x}_i| \\ t_p(\vec{x}_i) &= \frac{2}{c}\sqrt{z_p^2 + (x_i - x_p)^2}, \end{aligned} \tag{4.1}$$

where c is the speed of sound, and the time t_p is measured from the beginning of the emission.

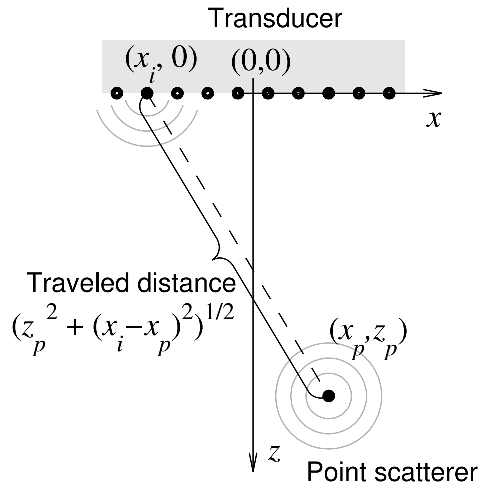


Figure 4.2: Illustration of the synthetic aperture geometry.

If the signal emitted by element i is $g(t)$, then the received signal can be expressed as:

$$r(t,x_i) = \sigma_p g\left(t - \frac{2\sqrt{z_p^2 + (x_i - x_p)^2}}{c}\right), \tag{4.2}$$

where σ_p is the back-scattering coefficient. Assuming that the tissue can be modeled as a collection of point scatterers, and applying Born's approximation (there is no secondary scattering), the signal received by element i becomes:

$$r(t,x_i) = \sum_p \sigma_p g\left(t - \frac{2\sqrt{z_p^2 + (x_i - x_p)^2}}{c}\right) \tag{4.3}$$

The Fourier transform of the generic SAU signal $r(t,x_i)$ with respect to time t is:

$$\dot{R}(\omega, x_i) = \dot{G}(\omega) \sum_p \sigma_p \exp\left(-j2k\sqrt{z_p^2 + (x_i - x_p)^2}\right), \tag{4.4}$$

where $k = \omega/c$ is the wavenumber. As one can see, the SAU signal in the (ω, x) domain is composed of a linear combination of the spherical phase-modulated (PM) signals, that is,

$$\exp\left(-j2k\sqrt{z_p^2 + (x_i - x_p)^2}\right) \tag{4.5}$$

The Fourier transform with respect to space of (4.4) is:

$$\dot{R}(\omega, k_x) = \dot{G}(\omega) \sum_p \sigma_p \underbrace{\exp\left(-j\sqrt{4k^2 - k_x^2}z_p - jk_x x_p\right)}_{\text{Linear phase function of } (z_p, x_p)}, \tag{4.6}$$

for $k_x \in [-2k, 2k]$. k_x is referred to as synthetic aperture frequency domain, or slow-time frequency domain. In the classical SAR imaging systems, from the 1950s, this variable was called the *slow-time Doppler* domain.

4.1.3 Reconstruction

The SAU signal can be rewritten in terms of two variables k_x and k_z:

$$\dot{R}(\omega, k_x) = \dot{G}(\omega) \sum_p \sigma_p \exp\left(-jk_z z_p - jk_x x_p\right), \tag{4.7}$$

where k_z is:

$$k_z = \sqrt{4k^2 - k_x^2}. \tag{4.8}$$

These two variables are also known as SAU *spatial frequency mapping* or *transformation*.

The *ideal* function of the medium is defined in the spatial domain via:

$$f_m(x,z) = \sum_p \sigma_p \delta(x - x_p, z - z_p) \tag{4.9}$$

This function has the following Fourier transform:

$$\dot{F}_m(k_z, k_x) = \underbrace{\sum_p \sigma_p \underbrace{\exp(-jk_z z_p - jk_x k_p)}_{\text{Linear phase function}}}_{\text{Linear combination}} \tag{4.10}$$

The SAU signal is then given by:

$$\dot{R}(\omega, k_x) = \dot{G}(\omega)\dot{F}_m(k_z, k_x), \tag{4.11}$$

where (k_z, k_x) are governed by the SAU spatial frequency mapping. For the reconstruction of $f_m(z,x)$ or $F_m(k_z, k_x)$ from the Fourier transform of the measured signal $R(\omega, k_x)$ one has:

$$\dot{F}_m(k_z, k_x) = \frac{\dot{R}(\omega, k_x)}{\dot{G}(\omega)}. \tag{4.12}$$

This is a theoretical reconstruction since $g(t)$ is usually band limited signal. Moreover $R(\omega, k_x)$ is zero for $|k_x| > 2k$. The practical reconstruction is by time matched filtering, that is:

$$\begin{aligned} \dot{F}_m(k_z, k_x) &= \dot{G}^*(\omega)\dot{R}(\omega, k_x) \\ &= |\dot{G}(\omega)|^2 \sum_p \sigma_p \exp(-jk_z z_p - jk_x x_p), \end{aligned} \tag{4.13}$$

for $k_x \in [-2k, 2k]$.

The reconstruction of the scatter map from the SAU domain signal can be done either in the (k_z, k_x) domain or in the time domain. Several reconstruction algorithms are known, such as reconstruction via spatial frequency interpolation, reconstruction via range stacking, or reconstruction via slow-time fast-time matched filtering. Because the computing power is not of concern in this study, a method known in the SAR systems as *back-projection algorithm*, or conversely ”delay and sum” will be considered in the following.

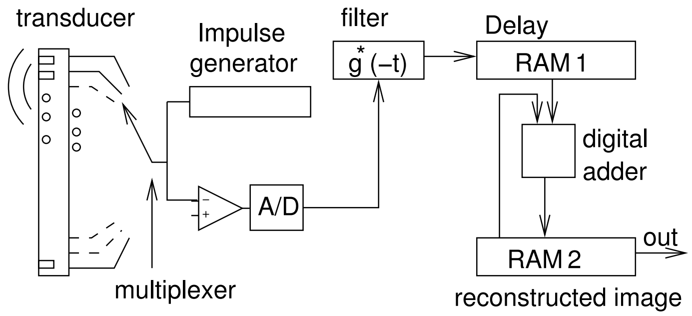


Figure 4.3: Simplified scheme of a synthetic aperture ultrasound system

4.1.4 Delay and sum reconstruction

As it was shown in the previous section, the received signal must be matched filtered in the time domain.

Figure 4.3 shows a simplified SAU system. The signal in receive is sampled and then fed into a matched-filter. The fast-time matched-filtered signal is denoted as:

$$r_m(t, x_i) = \mathrm{Re}\{\dot{r}(t, x_i) \underset{t}{*} \dot{g}^*(-t)\}, \tag{4.14}$$

where $\underset{t}{*}$ denotes convolution in time. The fast-time matched-filtered signal is then stored into a random access memory - RAM 1. From there the samples at the right times are read by a digital adder. The target function at coordinates (z_p, x_p) is then reconstructed by:

$$\begin{aligned}\hat{f}_m(z_p, x_p) &= \sum_i r_m\left(\frac{2\sqrt{z_p^2 + (x_p - x_i)^2}}{c}, x_i\right) \\ &= \sum_i r_m\left(t_p(x_i), x_i\right)\end{aligned} \tag{4.15}$$

where

$$t_p(x_i) = \frac{2\sqrt{z_p^2 + (x_p - x_i)^2}}{c} \tag{4.16}$$

is the round trip delay of the echoed signal to the point scatterer at (x_p, z_p) when the transducer element is at coordinates $(x_i, 0)$. Thus, to form the target function at a given point (x_p, z_p) in the spatial domain, one can coherently add the data at the fast-time samples that corresponds to the location of that point for all elements located at coordinates $(x_i, 0)$.

Usually the samples are interpolated for the right time instances $t_p(x_i)$. If this step is skipped, valuable high-resolution information is lost.

The reconstructed signal is not yet ready for display. It must be further envelope detected, logarithmically compressed, and then sent to the display.

4.2 Performance

The performance of the method will be evaluated in terms of resolution, frame rate and signal to noise ratio.

L is the size of the aperture that can be used to reconstruct one point

L

w

$\theta_a \approx \frac{w}{\lambda}$

z

$L \approx \theta_a z$

Figure 4.4: Illustration of the influence of the element size on the useful aperture. The bigger the width of a single transducer element, the narrower is the beam. The width of the beam determines the width of the useful aperture.

4.2.1 Resolution

Generally the resolution of a system is estimated by its bandwidth. In this case the spatial resolution of the system is estimated by its spatial bandwidth. Remember the frequency representation of the synthetic aperture signal:

$$\dot{R}(\omega, k_x) = \dot{G}(\omega) \sum_p \sigma_p \exp\left(-jk_z z_p - jk_x x_p\right). \tag{4.17}$$

The above signal implies that the array is composed of point-like transducer elements that span $x_i \in (-\infty, \infty)$. The bandwidth is then determined by:

$$k_x \in [-2k, 2k], \tag{4.18}$$

where $k = 2\pi/\lambda$ is the wavenumber. The *theoretical* maximum spatial resolution is given by [33] :

$$\delta_x = \frac{\pi}{2k} \tag{4.19}$$

$$\boxed{\delta_x = \frac{\lambda}{4}} \tag{4.20}$$

There are however practical limitations to the above equation: (1) the elements are not infinitely narrow, and (2) the array is not infinitely large. Let's consider these factors one by one.

Influence of the finite element size

The finite element size in the azimuth direction imposes a general limitation on the maximum size of the aperture. Consider Figure 4.4. The transducer elements are usually planar, and have small width w. Because of their small width it is possible to state that the imaged points are in the far field of the transducer elements. The far field is characterized by a divergence angle θ_a [21] which for a planar radiator is:

$$\theta_a \approx 2 \arcsin \frac{w}{2\lambda} \approx \frac{w}{\lambda}. \tag{4.21}$$

The beamwidth at a given distance z is then determined by [68]:

$$L = 2z \sin \theta_a \approx z \frac{w}{\lambda}, \tag{4.22}$$

which gives a spatial resolution of;

$$\delta_x \leq \frac{w}{2} \tag{4.23}$$

It can be seen that the resolution is depth-independent, provided that the number of transducer elements is infinite, and that there are enough elements to form the aperture for every depth. Usually the element width for phased arrays is $w = \lambda/2$, giving a maximum resolution of $\delta_x = \lambda/2$, which was shown to be also the case for infinitely small elements. The practical limitation for the resolution of a real-life SAU system is therefore not imposed by the element width.

Influence of the finite aperture size

Far more important is the influence of the finite aperture size. Let's consider the radiation pattern of a SAU system. Neglecting the discrete nature of the synthesized aperture [56], the transfer function of the synthetic aperture imaging can be written as:

$$\mathsf{A}_{t/r\ \mathrm{SAU}}(x) = L \frac{\sin\left(kL\frac{x}{z}\right)}{kL\frac{x}{z}}, \tag{4.24}$$

where L is the length of the synthesized aperture, and z is the axial distance at which the point-spread-function is derived. Letting $kL\frac{x}{z} = \zeta$ one gets:

$$\mathsf{A}_{t/r\ \mathrm{SAU}}(\zeta) = L \frac{\sin \zeta}{\zeta} \tag{4.25}$$

A suitable level to search for the resolution δ_x is $2/\pi$, which corresponds to a level -4 dB from the maximum. This level is chosen because is it is easy to find an analytic solution for the resolution. The above equation gets the solution:

$$\frac{\mathsf{A}_{t/r\ \mathrm{SAU}}(\zeta)}{\mathsf{A}_{t/r\ \mathrm{SAU}}(0)} = \frac{\pi}{2} \tag{4.26}$$

$$\mathsf{A}_{t/r\ \mathrm{SAU}}(0) = L \tag{4.27}$$

$$\frac{\sin(\zeta)}{\zeta} = \frac{2}{\pi} \tag{4.28}$$

If ζ is let to be:

$$\zeta = \frac{\pi}{2} = \zeta_0, \tag{4.29}$$

then $\sin \zeta_0 \equiv 1$, and the following equation is valid:

$$\frac{\sin \zeta_0}{\zeta_0} = \frac{\sin \frac{\pi}{2}}{\frac{\pi}{2}} \equiv \frac{2}{\pi}. \tag{4.30}$$

So the -4 dB resolution in the azimuth plane is found when $\zeta_0 = \frac{\pi}{2}$. Going back to the substitution one gets:

$$\frac{\pi}{2} = kL\frac{x_0}{z} \tag{4.31}$$
$$x_0 = \frac{z\pi}{2Lk} \tag{4.32}$$

The beamwidth is twice the distance at which the beam gets the respective level:

$$\delta_{x4\mathrm{dB}} = 2x_0 = \frac{z\pi}{kL}. \tag{4.33}$$

The beamwidth at a different level can be found using some numerical method. It can be seen that the lateral resolution δ_x is proportional to the axial distance z. It is therefore convenient to express the point spread function as a function of the angle $\theta = \arcsin\frac{x}{z}$. In this case:

$$\delta_\theta = 2\arcsin\frac{\delta_x}{2z} \tag{4.34}$$
$$\delta_{\theta 4\mathrm{dB}} = 2\arcsin\frac{\pi}{2Lk}. \tag{4.35}$$

The angular resolution is depth independent for the case of synthetic aperture imaging.

Consider the discrete nature of the synthetic aperture, and let the number of elements in the system be N_{xdc}, and the signal received by element i be r_i.

The reconstructed image is:

$$\hat{f}_m(t,\vec{x}_p) = \sum_{i=1}^{N_{xdc}} r_i(t - 2\frac{|\vec{x}_p - \vec{x}_i|}{c}). \tag{4.36}$$

Using the Fourier relation between the radiation pattern of a single element in its far field, and substituting $\vec{x}_p$ with a spherical coordinate system one gets the angular point spread function of the system. It is in this case given by [56, 65, 68]:

$$\mathsf{A}_{t/r\ \mathrm{SAU};\delta}(\theta) = \frac{\sin(kN_{xdc}d_x\sin\theta)}{\sin(kd_x\sin\theta)}, \tag{4.37}$$

where

$$\theta = \arcsin\frac{z}{x} \tag{4.38}$$

is the angle between the scan-line and the normal vector to the transducer surface. The distance between the centers of two elements (the transducer pitch) is d_x. This function repeats itself with a period which is equal to:

$$\begin{aligned} \pi &= kd_x\sin(\theta) \\ \sin\theta &= \frac{\pi}{kd_x} = \frac{\lambda}{2d_x} \end{aligned} \tag{4.39}$$

The replicas of the main lobe are known as grating lobes. Equation (4.37) does not take into consideration the size of the transducer element, only the finite size and the discrete nature of the aperture.

Comparing it with the two-way radiation pattern of a phased array [24]:

$$\mathsf{A}_{t/r\ \mathrm{ph};\delta}(\theta) = \frac{\sin^2\left(\frac{k}{2}N_{xdc}d_x\sin\theta\right)}{\sin^2\left(\frac{k}{2}d_x\sin\theta\right)}, \tag{4.40}$$

one notices:

- The resolution of a SAU system is bigger than the one of a phased array system for the same number of elements.
- The side-lobe level of a phased array system is lower than that of a SAU system.
- The grating lobes (the repetition of the PSF) appear closer to the main lobe for a SAU system. To push the grating lobes outside the visible region, the pitch of the array for a phased array system must be $\lambda/2$, while for a SAU system it must be $\lambda/4$.

The information necessary to reconstruct the image is carried by the phase *difference* of the received signals by two adjacent elements (in the SAR systems this is known as the slow-time Doppler frequency). When focusing is applied in transmit, at the focal point the wavefront is approximately a plane wave. The signal is scattered back by two neighboring scatterers at the same depth at the same time. The difference of the phase returned by two point scatterers at the focal depth is therefore created on the way back to the transducer. The difference between the phases of the signals received by two elements in a SAU system is based on the two-way propagation and the two scatterers are easier to distinguish based on the measured phase difference[1].

The side lobe level of the SAU system is higher because the amplification at the focal point is determined only by the sum of N_{xdc} channels. In a phased array system, this is a double sum - once for the focusing in transmit and once for the focusing in receive. This means that a normal phased array system will display better the dark regions such as cysts, while the SAU system will display better the strong reflectors.

The combined influence of the element and aperture size

The easiest way to demonstrate the combined influence of the finite aperture and element size is through the k-space representation of the system. In the following only the lateral spatial frequencies will be considered. Figure 4.5 shows how the k-space is built for the cases of phased array and synthetic aperture imaging. Both arrays consist of 5 elements. The phased array system can be represented [54] as a small rectangle convolved with a comb function the length of the array. The aperture function becomes:

$$a_{\mathrm{III};\delta}(x) = \sum_{n=1}^{N_{xdc}} \delta(x - d_x n - \frac{N_{xdc}+1}{2}d_x) \tag{4.41}$$

$$a_{\mathrm{el}}(x) = \begin{cases} 1 & -\frac{w}{2} \le x \le \frac{w}{2} \\ 0 & \text{otherwise} \end{cases} \tag{4.42}$$

$$a_{t\ \mathrm{ph}}(x) = a_{r\ \mathrm{ph}}(x) = a_{\mathrm{el}}(x) \underset{x}{*} a_{\mathrm{III};\delta}(x) \tag{4.43}$$

[1] The lower resolution of a phased array system is determined by the way the information is processed. After all the information is gathered from the same spatial positions.

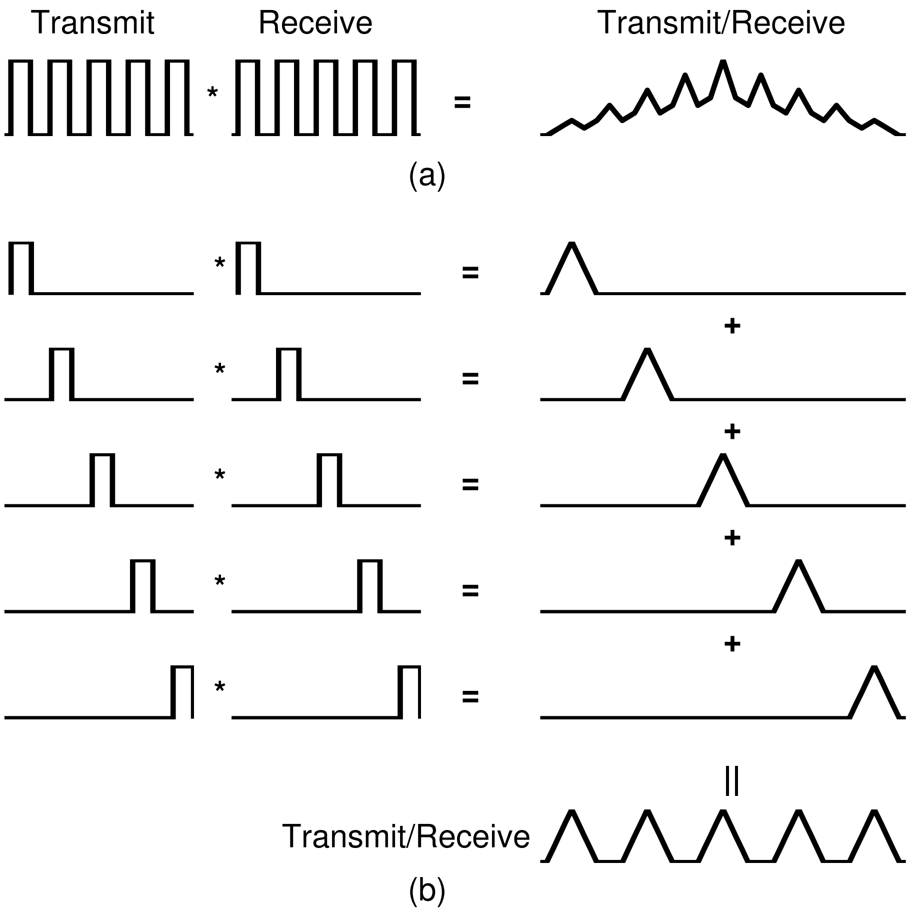


Figure 4.5: Filling in the k-space for (a) phased array imaging, and (b) synthetic aperture imaging.

where $a_{\mathrm{el}}(x)$ is the function describing the finite element size. The lateral k-space representation of the system is equal to the aperture function. The two way k-space representation of the phased array system is the convolution between the transmit and receive aperture functions,

$$a_{t/r\ \mathrm{ph}}(x) = a_t(x) \underset{x}{*} a_r(x) \tag{4.44}$$

$$a_{t/r\ \mathrm{ph}}(x) = a_{\mathrm{III};\delta}(x) \underset{x}{*} a_{\mathrm{III};\delta}(x) \underset{x}{*} a_{\mathrm{el}} \underset{x}{*} a_{\mathrm{el}}(x) \tag{4.45}$$

This operation is shown in Figure 4.5.

The two-way k-space representation of a phased array system consisting of point-like elements is $a_{\mathrm{ph};\delta} = a_{\mathrm{III};\delta} * a_{\mathrm{III};\delta}$, and is related to $\mathsf{A}_{t/r\ \mathrm{ph};\delta}(x)$ through the Fourier transform. It is a digital squared sinc function whose period is determined by the transducer pitch d_x. The width of the function is inversely proportional to the size of the array. This function is plotted in Figure 4.6(a). The pulse-echo response of a single element with width w is:

$$\mathsf{A}_{t/r\ \mathrm{el}}(\theta) = \left(\frac{\sin\left(\frac{k}{2} w \sin\theta\right)}{\frac{k}{2}\sin\theta} \right)^2 . \tag{4.46}$$

Due to the Fourier properties of the convolution the resulting point spread function of the phased array system is given by:

$$\mathsf{A}_{t/r\ \mathrm{ph}}(\theta) = \mathsf{A}_{t/r\mathrm{ph};\delta}(\theta) \cdot \mathsf{A}_{t/r\ \mathrm{el}}(\theta) \tag{4.47}$$

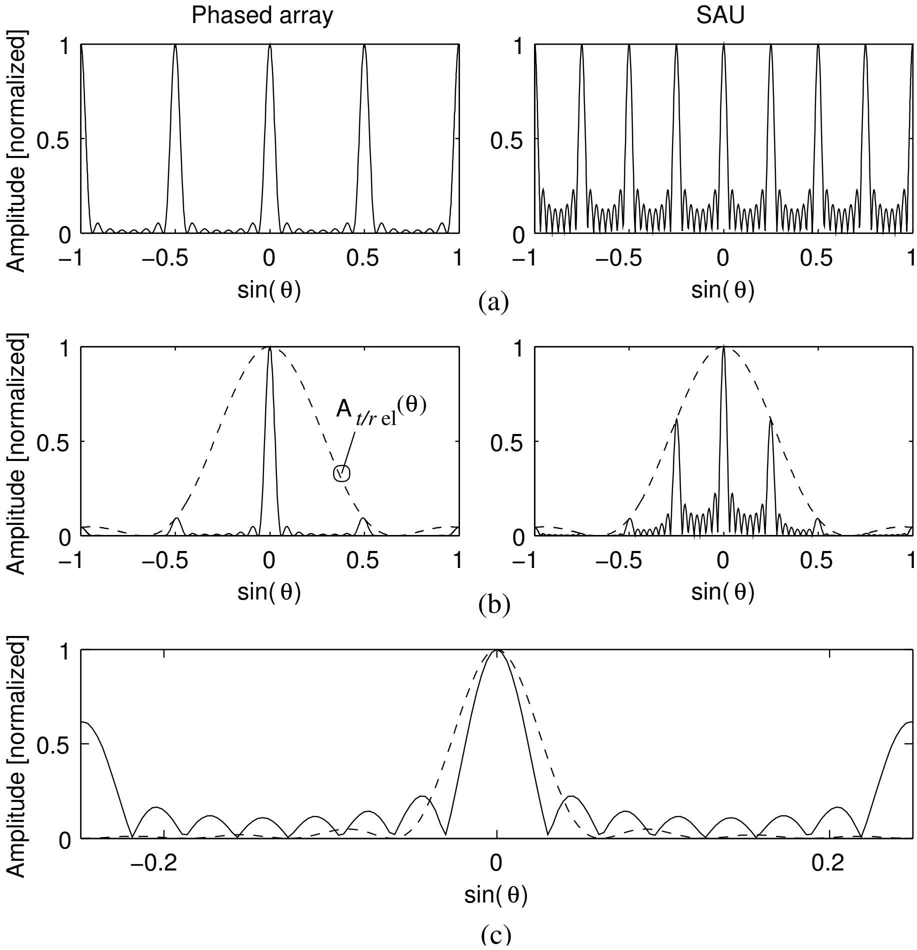


Figure 4.6: Creating the two-way radiation pattern. Sub-plot (a) shows the radiation pattern when the elements are δ functions, and (b) when the elements have finite size. Sub-plot (c) shows a comparison between the radiation patterns of a phased array imaging (dashed) and of a synthetic aperture imaging. The distance between the elements d_x is $d_x = 2\lambda$, and the width of the elements is $w = 1.5\lambda$

and is shown in Figure 4.6(b).

In synthetic aperture imaging the transmission and reception are performed only by a single element. The radiation pattern for the transmit element n is given by:

$$a_{n\ r\ \mathrm{SAU}}(x) = a_{n\ t\ \mathrm{SAU}}(x) = a_{\mathrm{el}} * \delta(x - d_x n - \frac{N_{xdc}+1}{2} d_x). \tag{4.48}$$

The transmit receive radiation pattern is then:

$$a_{n\ t/r\ \mathrm{SAU}}(x) = \underbrace{a_{\mathrm{el}}(x) \underset{x}{*} a_{\mathrm{el}}(x)}_{=a_{t/r\ \mathrm{el}}(x)} \underset{x}{*} \delta\left(x - 2(d_x n - \frac{N_{xdc}+1}{2} d_x)\right). \tag{4.49}$$

Because the system is linear, the k-space representation of the final image is the sum of the

k-space representations of each of the emissions:

$$a_{t/r\ \mathrm{SAU}}(x) = \sum_{n=1}^{N_{xdc}} a_{t/r\ \mathrm{el}}(x) \underset{x}{*} \delta\left(x - 2(d_x n - \frac{N_{xdc}+1}{2} d_x)\right) \tag{4.50}$$

$$a_{t/r\ \mathrm{SAU}}(x) = a_{t/r\ \mathrm{el}}(x) \underset{x}{*} \sum_{n=1}^{N_{xdc}} \delta\left(x - 2(d_x n - \frac{N_{xdc}+1}{2} d_x)\right). \tag{4.51}$$

The radiation pattern $\mathsf{A}_{\mathrm{SAU};\delta}$ of the sum of delta functions is given by (4.37) and is plotted in Figure 4.6(a). Using the properties of the Fourier integral, the two way radiation pattern becomes:

$$\mathsf{A}_{t/r\ \mathrm{SAU}}(\theta) = \mathsf{A}_{t/r\ \mathrm{el}}(\theta) \cdot \mathsf{A}_{t/r\ \mathrm{SAU};\delta}(\theta). \tag{4.52}$$

The transmit-receive radiation pattern of the array is weighted with the radiation pattern of the single element. This operation is shown in Figure 4.6(b).

Generally the size of the transducer element decreases the levels of the grating and side lobes, giving a higher contrast resolution in the system.

4.2.2 Frame rate

The frame rate of a synthetic aperture ultrasound system is given by:

$$f_{fr\ \mathrm{SAU}} = \frac{f_{prf}}{N_{\mathrm{firings}}}, \tag{4.53}$$

where f_{prf} is the pulse repetition frequency and N_{firings} is the number of emissions necessary to acquire the data for a single frame. Because the acquisition is done using the transducers elements in transmit and receive one by one, the number of firings is equal to the number of transducer elements, i.e.:

$$N_{\mathrm{firings}} = N_{xdc}. \tag{4.54}$$

The frame rate then is:

$$f_{fr\ \mathrm{SAU}} = \frac{f_{prf}}{N_{xdc}}. \tag{4.55}$$

For a phased array the frame rate is determined by the number of lines N_l in the image and is:

$$f_{fr\ \mathrm{ph}} = \frac{f_{prf}}{N_l} \tag{4.56}$$

The number of lines in an image is determined by the lateral resolution of the system. For a phased array system one can use the approximate formula [65]:

$$N_l \geq 1.5 N_{xdc} \tag{4.57}$$

The frame rate for a phased array system becomes:

$$f_{fr\ \mathrm{ph}} = \frac{f_{prf}}{1.5 N_{xdc}} \tag{4.58}$$

Comparing the frame rates for phased array imaging and synthetic array imaging, one notices the higher frame-rate of synthetic array imaging, which for the same number of elements is:

$$\frac{f_{fr\ \mathrm{SAU}}}{f_{fr\ \mathrm{ph}}} \geq 1.5 \tag{4.59}$$

There exists a catch. Each of the lines in the phased array imaging is beamformed after a single emission. The influence of tissue motion is in this case virtually non-existent. The information necessary to beamform the same line for synthetic array imaging is gathered for a time span N_{xdc}/f_{prf}. In this case the tissue motion can be substantial, impeding the coherent summation of the RF data. This can be overcome by using motion compensation. Motion compensation strategies will be discussed in Chapter 12

4.2.3 Signal-to-noise ratio

The signal-to-noise ratio depends in general on the transmitted energy into the tissue. What we are concerned here is the *gain* in the signal-to-noise ratio of the beamformed signal, compared to the signal-to-noise ratio of a single received signal. The gain in the signal-to-noise ratio is defined as:

$$GSNR = 10\log_{10}\frac{SNR}{SNR_0}, \tag{4.60}$$

where SNR_0 is the signal-to-noise ratio of a single element. For a synthetic array aperture this ratio is given by

$$GSNR = 10\log_{10} N_{xdc}, \tag{4.61}$$

where N_{xdc} is the number of elements in the array.

CHAPTER

FIVE

Variations of SAU

The previous chapter (Chapter 4) described the Generic Synthetic Aperture Ultrasound Imaging (GSAU), which is a direct copy of the *Airborne* Synthetic Aperture Radar. This type of synthetic aperture imaging is natural for the radar systems, since in most cases only a single antenna is available. In ultrasound scanners, on the other hand, all of the transducer elements are accessible at the same time, and there is no reason not to use them, unless some restrictions on complexity of the hardware are imposed. This is equivalent to having the aircraft at all imaging positions at the same time, giving much higher flexibility in the focusing algorithms as compared to the SAR systems.

Many researchers have tried to exploit this flexibility and a number of methods are described in the literature. Generally in ultrasound imaging two apertures can be defined: transmit and receive apertures. This chapter introduces a classification of the synthetic aperture algorithms depending on which of the two apertures is being synthesized. Three cases are possible: (1) synthetic receive aperture, (2) synthetic transmit aperture, and (3) synthetic transmit and receive apertures.

5.1 Synthetic receive aperture

5.1.1 Simple model

This type of synthetic aperture imaging was studied by the ultrasound group at Duke University and was presented in a series of papers by Nock and Trahey [69, 70, 71]. In 1994 Walker and Trahey presented some experimental results the IEEE Ultrasonics symposium [72]. One of the systems,the XTRA system, which is used for some of the measurements in the following, is also implemented as a synthetic receive aperture system (see Appendix J). In all of the listed cases the reason for using a synthetic receive aperture was to reduce complexity of the hardware. The complex part is usually assumed to be the receiver. If the system is digital and uses N_{act} of the transducer elements at the same time, this means building N_{act} receive channels. Each of the channels usually consists of analog amplifier, matched filter, analog to digital converter, and delay lines capable of dynamic focusing. The idea behind the receive synthetic aperture imaging is, that only one or some of the transducer elements can be used in receive. This results in a small active receive aperture. Changing (multiplexing) between different elements in receive, allows for a bigger receive aperture to be synthesized. The approach is illustrated in Figure 5.1. In transmit all of the elements focus along one line. In receive only one transducer element is active. Its data is memorized. Then the transmission is repeated in the same direction. Again only a single element is active in receive. The procedure is repeated

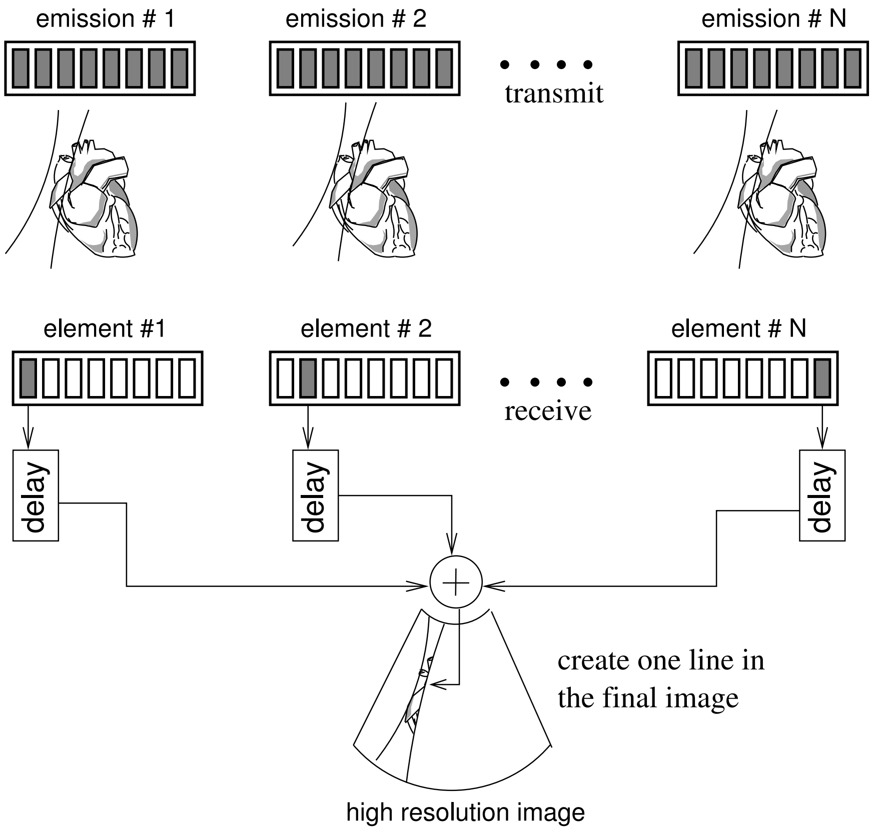


Figure 5.1: Receive synthetic aperture imaging.

until all of the N_{xdc} transducer elements have been used in receive.

5.1.2 Performance

The performance is evaluated by the resolution, the achieved frame rate and the signal-to-nose ratio.

Resolution

The resolution of the system is determined by its *k*-space bandwidth. The spatial bandwidth is given by the effective aperture (see Section 3.7). Because of its nature, the *k*-space representation of the SRAU is equal to the *k*-space of the normal phased array imaging and so is the the resolution.

Frame rate

The frame rate is determined by the number of lines per image N_l, and the number of emissions N_{firings} necessary to acquire a single line:

$$f_{fr\ \text{SRAU}} = \frac{f_{prf}}{N_l \cdot N_{\text{firings}}}, \tag{5.1}$$

where f_{prf} is the pulse repetition frequency.

The simple model assumed that the reception was performed only by a single element, i.e. $N_{\text{firings}} = N_{xdc}$. It is possible that more than one element is active in receive, i.e. a small sub-aperture is active. In many applications only some of the transducer elements are used to beamform a line. Let the total number of elements used in the synthetic receive aperture be N_{rcv}, and the number of active elements in receive be N_{act}. If no element is used twice, then the total number of emissions per line is:

$$N_{\text{firings}} = \left\lceil \frac{N_{rcv}}{N_{act}} \right\rceil \tag{5.2}$$

The frame rate becomes:

$$f_{fr\ \text{SRAU}} = \frac{f_{prf} \cdot N_{act}}{N_{\text{lines}} \cdot N_{rcv}} \tag{5.3}$$

The synthetic receive aperture imaging method has a frame rate f_{frame} which is a N_{firings} times lower than the one of the phased array imaging:

$$\frac{f_{fr\ \text{SRAU}}}{f_{fr\ \text{ph}}} = \frac{1}{N_{\text{firings}}} \tag{5.4}$$

When $N_{act} = N_{rcv}$ only one transmission is necessary per scan line. The method becomes equivalent to the normal phased or linear array imaging.

The frame rate for the generic synthetic aperture ultrasound, given in Chapter 4 is:

$$f_{fr\ \text{GSAU}} = \frac{f_{prf}}{N_{rcv}}, \tag{5.5}$$

The ratio between the frame rates between the generic synthetic aperture imaging and the synthetic receive aperture imaging is:

$$\frac{f_{fr\ \text{GSAU}}}{f_{fr\ \text{SRAU}}} = \frac{N_l \cdot N_{\text{firings}}}{N_{rcv}} \tag{5.6}$$

Signal-to-noise ratio

The signal-to-noise ratio for the synthetic receive aperture imaging is as the signal-to-noise ratio for the phased array imaging. The gain in SNR compared to the SNR of of the signal of a single element is given by[65]:

$$\frac{\text{SNR}}{\text{SNR}_0} = 10 \log N_{xdc}^3 \ \text{dB}, \tag{5.7}$$

assuming that the number of transmitting elements is equal to the number of receiving elements, and is equal to the total number of transducer elements N_{xdc}.

5.1.3 Implementation

The implementation of a synthetic receive aperture system is illustrated in Figure 5.2. In receive only a single element is multiplexed to the receive amplifier. The amplification can be changed as a function of time (TGC). The signal is digitized and stored in RAM. The implementation of

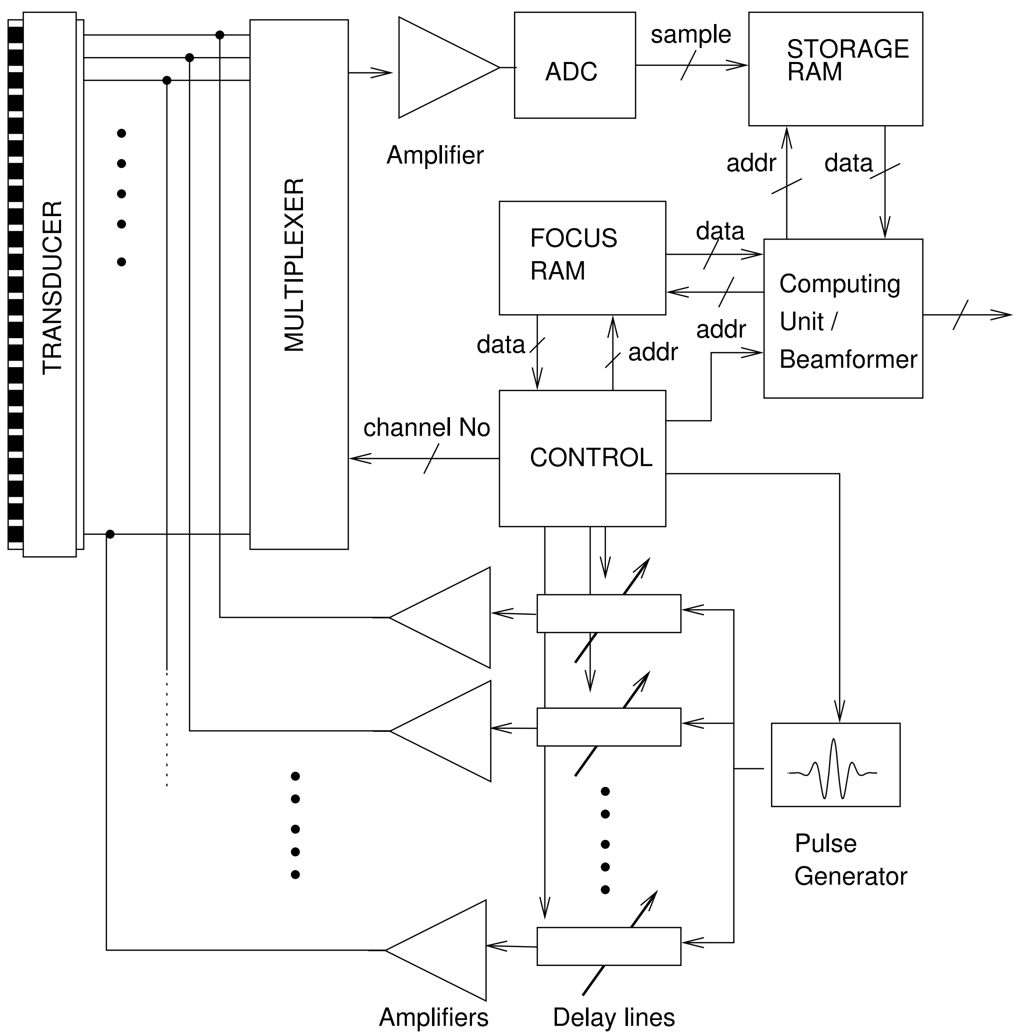


Figure 5.2: An example of the necessary units of a synthetic receive aperture ultrasound imaging system

the RAM can be in two pages, each capable of storing one RF line of data. In order to maintain real-time processing, while the sampled data is fed into one of the pages, the data from the other is read and processed by the beamformer. The beamformer performs dynamic focusing.

The transmitter unit consists of a pulse generator, delay lines and amplifiers. The delay lines realize a fixed focus and their delays change only from line to line.

The role of the CONTROL is to synchronize the work of all the units, to notify the beamformer which receive channel is active and which line is being beamformed. It is also responsible for setting the delays in transmit.

5.2 Synthetic transmit aperture

The synthetic receive aperture imaging has the advantage over the generic synthetic aperture imaging of a higher signal-to-noise ratio and lower sidelobe level. It possesses one serious disadvantage, namely long acquisition time - from all of the methods considered so far, it is the slowest. A way to increase the frame rate is offered by the *synthetic transmit aperture*

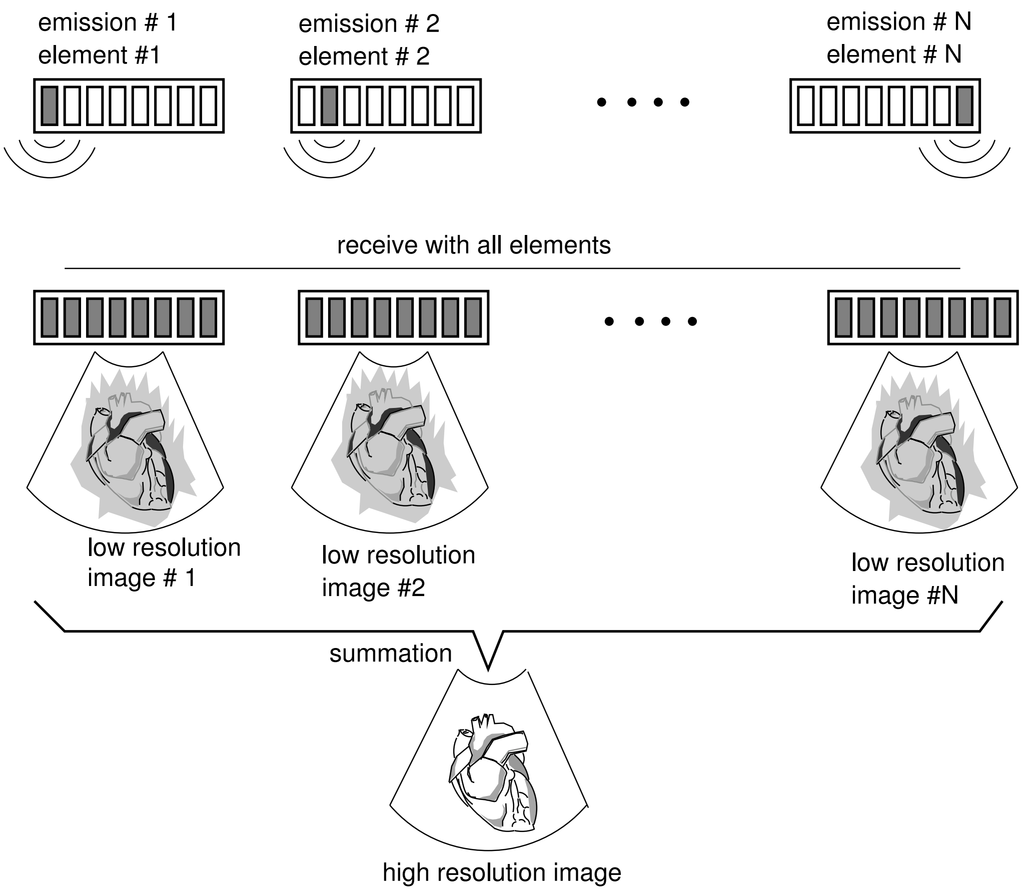


Figure 5.3: Simple model for synthetic transmit aperture ultrasound imaging.

ultrasound imaging. In the following we will consider a simple model and the performance of the method

5.2.1 Simple model

The simple model for the synthetic transmit aperture ultrasound imaging (STAU) is given in Figure 5.3.

In transmission only a single element is used. It creates a cylindrical wave (in the elevation plane the shape of the wavefront is determined by the height of the transducer and the presence/absence of an acoustic lens) which covers the whole region of interest. The received echo comes from all imaging directions, and the received signals can be used to create a whole image - in other words all of the scan lines can be beamformed in parallel. The created image has low-resolution because there is no focusing in transmit, and therefore in the rest of this report it is called a low-resolution image (LRI). After the first LRI, $\mathbf{L}_1(t)$, is acquired another element transmits and a second LRI, $\mathbf{L}_2(t)$, is created. After all of the transducer elements N_{xdc} have transmitted, the low resolution images are summed and a high-resolution image (HRI) is created. In this case only N_{xdc} emissions are necessary to create a high-resolution image. It is possible to use only some of the transducer elements to speed up the acquisition process [1, 73, 74]. The result is a sparse synthetic transmit aperture. The technique was thoroughly studied by Hazard and Lockwood [75] who are currently developing a system for real-time synthetic transmit aperture imaging. The design of sparse arrays (synthetic and real) is considered

$\vec{x}_j$ $\vec{x}_i$

x

Origin of the scan line $\vec{x}_c$

Transmitting element i with spatial position $\vec{x}_i$

Receiving element j at $\vec{x}_j$

Scan line $L_i(t)$

Focal point at position $\vec{x}_f$

z

Figure 5.4: The geometry associated with the delay calculation.

in greater detail in Chapter 8. The synthetic transmit aperture technique was used by Bae and colleagues [76] to increase the resolution of the acquired B-mode images using the standard focused transmissions. This approach will be more thoroughly considered in Chapter 7.

Let $L_i(t)$[1] be a single low resolution line in the image, obtained after the emission with element $i \in [1, N_{xdc}]$. The beamformation procedure for this line can be expressed as:

$$L_i(t) = \sum_{j=1}^{N_{xdc}} a_{ij}(t) r_{ij}(t - \tau_{ij}(t)), \tag{5.8}$$

where t is the time measured from the trigger of the emission, $r_{ij}(t)$ is the signal received by the element j. τ and a are the applied delay and apodization coefficients, respectively. These two coefficients depend on which line in the image is being beamformed, and the time necessary for the wavefront to propagate from the transmitting element i to the focal point and back to the receiving element. Generally τ and a are also functions of time in the case of dynamic focusing and apodization.

Figure 5.4 shows the geometry used to calculate the delays. The current focal point is $\vec{x}_f$.

The time necessary for the emitted wave to propagate from the emitting element i to the focal point f and back to the receiving element j is given by:

$$t_{ij} = \frac{1}{c}\left(|\vec{x}_f - \vec{x}_i| + |\vec{x}_f - \vec{x}_j|\right), \tag{5.9}$$

where $\vec{x}_i$, $\vec{x}_j$ and $\vec{x}_f$ are the positions of the emitting, receiving and the focal point, respectively. The time that is necessary for a wave to propagate from the origin $\vec{x}_c$ to the focal point $\vec{x}_f$ and back is given by:

$$t_{cc} = \frac{2}{c}\left|\vec{x}_f - \vec{x}_c\right| \tag{5.10}$$

[1] For notational completeness $L_i(t)$ must be $L_{il}(t)$, where l is the line number. This, however, just makes the notations more complicated without having any influence on the results.

The delay τ_{ij} necessary to apply on the signal received by the element j after transmitting with element i is:

$$\tau_{ij} = t_{ij} - t_{cc} \tag{5.11}$$

After the low-resolution scan lines have been all beamformed, the high-resolution ones, $H_i(t)$, can be beamformed by using a straight forward sum:

$$H_i(t) = \sum_{i=1}^{N_{xdc}} L_i(t) \tag{5.12}$$

Equation (5.12) shows the formation of a single scan line. One LRI $\mathbf{L}_i(t)$ can be represented as a matrix, the columns of which are the low resolution scan lines $L_{il}(t)$, where l $(1 \geq l \geq N_l)$ is the index of the scan line. The same applies for the high resolution images $\mathbf{H}(t)$. The operation (5.12) can then be expressed in a matrix form:

$$\mathbf{H}(t) = \sum_{i=1}^{N_{xdc}} \mathbf{L}_i(t). \tag{5.13}$$

The high resolution image is obtained by summing the low resolution images. This operation is illustrated in Figure 5.3.

5.2.2 Performance

The synthetic transmit aperture ultrasound imaging covers the whole k-space as the synthetic receive aperture imaging does. The resolution is therefore the same as for the case of synthetic receive aperture imaging. The number of emissions necessary to scan the image is:

$$N_{\text{firings}} = N_{xdc}, \tag{5.14}$$

and the frame rate becomes:

$$f_{\text{frame}} = \frac{f_{prf}}{N_{xdc}}. \tag{5.15}$$

It gives the possibility for much faster imaging, especially if only some of the elements are used in transmit as discussed in Chapter 8. The signal-to-noise ratio and the penetration depths are, however, worser. The pressure amplitude is N_{xdc} times smaller for the STA imaging, and the gain in signal-to-noise ratio is [65]:

$$\frac{\text{SNR}}{\text{SNR}_0} = 10\log(N_{xdc}^2/2). \tag{5.16}$$

This signal-to-noise ratio is mostly "academic". Due to scattering and attenuation, after some depth there is not enough signal left in order to cause an echo detectable by the receiver. Means to increase the transmitted power are needed, and these include the use of virtual ultrasound sources (see Chapter 7) and spatial and temporal encoding (see Chapter 9).

5.3 Synthetic transmit and receive apertures

Figure 5.5 (taken from [65]) shows different combinations of active transmit and receive elements. The use of a single element in both, transmit and receive is what in this thesis is called

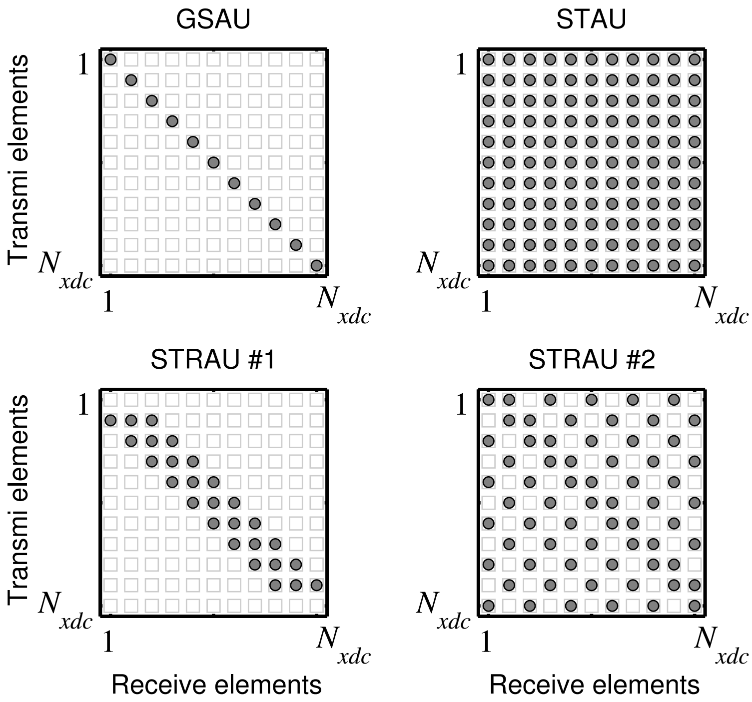


Figure 5.5: Active transmit/receive element combinations for different types of synthetic aperture ultrasound imaging. Each row corresponds to one transmit event. The transmission is done with the element whose index is equal to the row number. The filled squares symbolize an active element in receive.

generic synthetic aperture imaging (top left plot). Using all of the elements in receive is called synthetic transmit aperture imaging (top right plot). The cases in which more than one elements (but not all) are used in receive are labeled as synthetic transmit and receive imaging (the bottom row). These types of imaging have been the topic of several papers [67, 77, 78, 79] because of supposedly simple receiver design.

The bottom row of Figure 5.5 shows two different cases for transmit and receive synthetic aperture imaging. In the first case the elements that are neighboring to the transmit element are used in receive. Using neighboring elements fills in the gaps in the effective aperture (see Figure 4.5 on page 43) thus reducing the grating lobes level. The second case (see the bottom right plot in Figure 5.5) is supposed to generate all combinations of spatial frequencies (determined by the transmit receive pairs) without creating the same spatial frequency twice. For example transmitting with element i and receiving with element j is the same as transmitting with element j and receiving with element i. The effective apertures of the four methods and their two-way radiation patterns are illustrated in Figure 5.6. The transducer pitch is assumed to be half a wavelength ($\lambda/2$), and the elements have omnidirectional radiation pattern. The radiation patterns were obtained using the Fourier relation between them and point-spread function at the focus. The redundancy in the effective aperture caused by the multiple receive elements results in decreased side and grating lobe levels, and in an increase in signal-to-noise ratio.

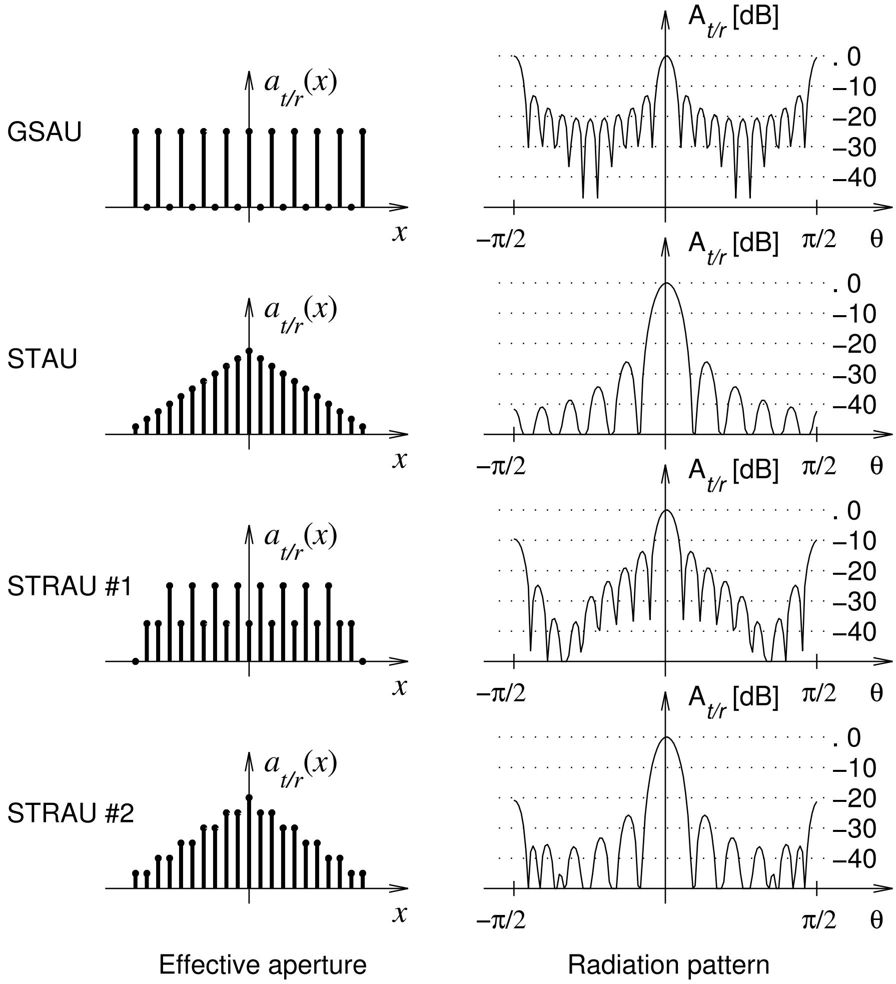


Figure 5.6: The effective apertures and the two-way radiation patterns of four variations of synthetic aperture imaging. The distance between the elements is $d_x = \lambda/2$ and the size of the elements is $w \to 0$.

5.3.1 Performance

The frame rate is determined by the number of emissions and the number of parallel receive beamformers. Assuming that all of the lines in the image can be beamformed in parallel, then the frame rate is:

$$f_{fr\ \mathrm{STRAU}} = \frac{f_{prf}}{N_{xdc}}. \tag{5.17}$$

The signal-to-noise ratio and the penetration depth are worser than those of the SRA and STA imaging modalities. The gain in signal-to-noise ratio compared to the signal-to-noise ratio of the signal received by a single element is:

$$\frac{\mathrm{SNR}}{\mathrm{SNR}_0} = 10\log(N_{xdc}N_{rcv}), \tag{5.18}$$

where N_{rcv} is the number of receive elements used at every reception and N_{xdc} is the total number of elements.

CHAPTER

SIX

Recursive ultrasound imaging

One of the goals of ultrasound scanning is to estimate the velocity of the blood. No matter which method for blood velocity estimation is used, samples from the same position in the tissue are needed at close time instances.

In the previous chapter the two major types of synthetic aperture ultrasound imaging were presented: (1) synthetic receive aperture and (2) synthetic transmit aperture. While the former method results in a somewhat simplified front-end electronics, the latter has the advantage of shorter acquisition time, which can be expressed as follows:

$$T_{acq\ \mathrm{HRI}} = T_{prf} \cdot N_{xdc} \cdot \frac{N_l}{N_{parallel}}, \tag{6.1}$$

where T_{prf} is the pulse repetition period, N_{xdc} is the number of transducer elements, N_l is the number of lines in an image, and $N_{parallel}$ is the number of parallel beamformers. Even if at every emission all the lines in the image were beamformed in parallel, i.e. N_l were equal to $N_{parallel}$, the time necessary to create a high-resolution image would still be N_{xdc} times longer than the pulse repetition interval. A way to create a new high-resolution image at every emission was suggested by us in [80], by using a recursive procedure for synthetic transmit aperture ultrasound imaging. For short it will be called *Recursive Ultrasound Imaging*.

This chapter introduces a simple model for recursive ultrasound imaging, discusses its performance and gives some implementation issues.

6.1 Simple model

Consider the imaging situation shown in Figure 6.1. The depicted transducer has $N_{xdc} = 3$ elements. In normal synthetic transmit aperture, the transducer transmits $N_{xdc} = 3$ times, using a single element $i \in [1,3]$ in transmission. At every emission a low resolution image $\mathbf{L}_i(t)$ is created. Summing coherently the low resolution images, a high resolution image $\mathbf{H}(t)$ is created as described in Chapter 5. After the high-resolution image is created, all the information is discarded and the process is repeated all over again.

If the medium is stationary the order in which the elements transmit does not matter. In other words, whether the order of transmission was $\{[i=1],[i=2],[i=3]\}$ or $\{[i=2],[i=3],[i=1]\}$ the result of summing the low-resolution images would yield the same high-resolution one. Going back to Figure 6.1 it can be seen that a high resolution image can be created at emission $i=3$, combining the low-resolution images from $\mathbf{L}_1(t)$ to $\mathbf{L}_3(t)$, which were obtained by exciting consecutively the elements from 1 to 3. Low resolution image $\mathbf{L}_4(t)$ was obtained

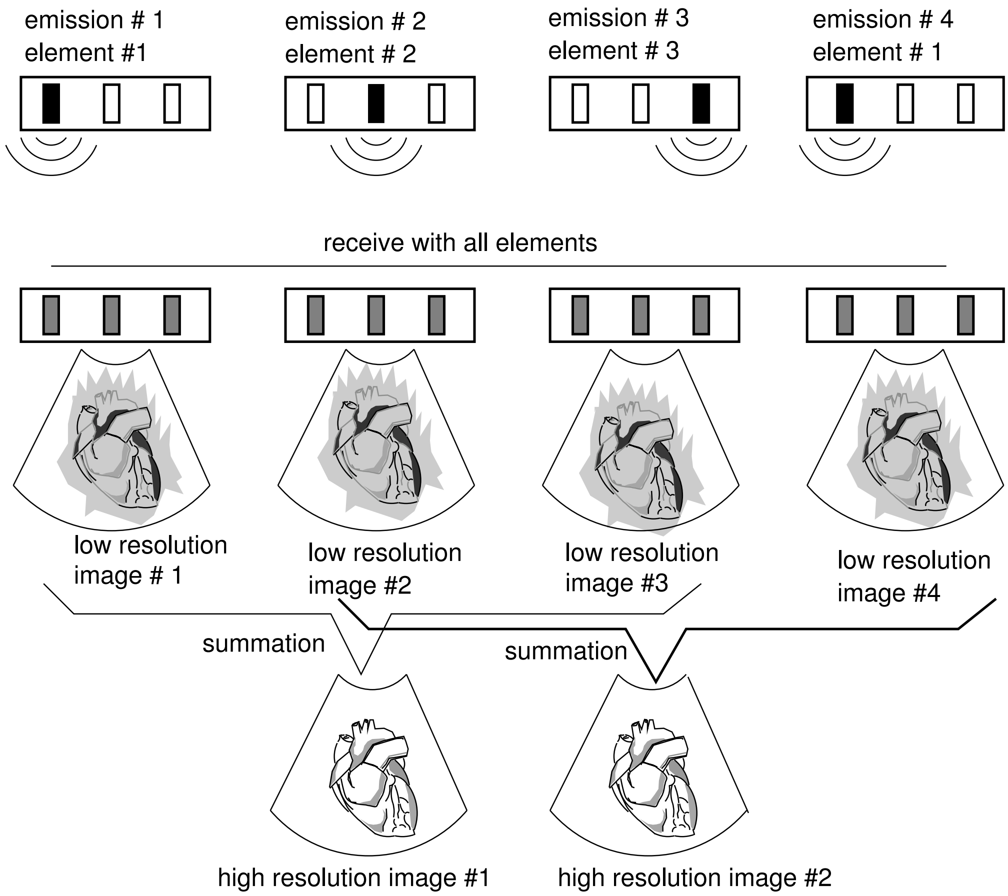


Figure 6.1: Synthetic transmit aperture imaging. The array consists of 3 elements. The figure shows 4 consecutive emissions. Two high resolution images are created: one at emission #3 and one at emission # 4.

by transmitting again with element # 1. Hence, combining low-resolution images from $\mathbf{L}_2(t)$ to $\mathbf{L}_4(t)$ ($\{[i=2],[i=3],[i=1]\}$), a new high-resolution one can be created at emission $n=4$.

Let's introduce an index of the emissions $n \in [1,\infty)$ counting the emissions from the moment the scanner was turned on. From Figure 6.1 it can be seen that emission # 4, is done again with element # 1. The relation between the emission number n and the index of the transmitting element i is given by

$$i = ((n-1) \bmod N_{xdc}) + 1 \tag{6.2}$$

It can be easily demonstrated that the same element i is used at emissions n, $n-N_{xdc}$ and $n+N_{xdc}$:

$$\begin{aligned} i &= ((n-1) \bmod N_{xdc}) + 1 \\ &= ((n-N_{xdc}-1) \bmod N_{xdc}) + 1 \\ &= ((n+N_{xdc}-1) \bmod N_{xdc}) + 1 \end{aligned} \tag{6.3}$$

In the previous chapter the signal received by element j after transmitting with element i was denoted as $r_{ij}(t)$. Thus the low resolution scan line $L_i(t)$ obtained after emitting with element i is expressed as:

$$L_i(t) = \sum_{j=1}^{N_{xdc}} a_{ij}(t) r_{ij}(t-\tau_{ij}(t)), \tag{6.4}$$

where $a_{ij}(t)$ is a dynamic apodization coefficient and $\tau_{ij}(t)$ is a delay dependent on the transmitting element i, the receiving element j and the current focal point. For simplicity the time dependence of the delay and the apodization coefficients will be omitted further in the chapter.

Let the signal at emission n received by element j be $r_j^{(n)}(t)$. This is an alternative notation for $r_{ij}(t)$, which exploits the relation between n, and i. In other words:

$$\begin{gathered} r_{ij}(t) \equiv r_j^{(n)}(t) \\ i = ((n-1) \bmod N_{xdc}) + 1 \end{gathered}$$

The same substitution will be made for the scan lines, i.e., $L_i(t)$, will be denoted with $L^{(n)}(t)$, meaning "the low-resolution scan-line obtained at emission n". The beamformation of a single low-resolution scan-line $L^{(n)}(t)$ can then be expressed as:

$$L^{(n)}(t) = \sum_{j=1}^{N_{xdc}} a_{ij} r_j^{(n)}(t - \tau_{ij}). \tag{6.5}$$

The indexing for a and τ is preserved, in order to emphasize that the transmission is done by element i. Assuming a stationary tissue, the signal received by element j after transmitting with element i will be the same, regardless of the emission number. The following equation is therefore valid:

$$r_j^{(n)}(t) = r_j^{(n+kN_{xdc})}(t), \qquad \text{where } k = 0, \pm 1, \pm 2 \cdots \tag{6.6}$$

Therefore the low-resolution images obtained at emissions n and $n + kN_{xdc}$ are the same, provided that the scanned tissue is stationary:

$$L^{(n)}(t) = \sum_{j=1}^{N_{xdc}} a_{ij} r_j^{(n)}(t - \tau_{ij}) \tag{6.7}$$

$$L^{(n+kN_{xdc})}(t) = \sum_{j=1}^{N_{xdc}} a_{ij} r_j^{(n+kN_{xdc})}(t - \tau_{ij}) \tag{6.8}$$

$$r_j^{(n+kN_{xdc})}(t) = r_j^{(n)}(t) \tag{6.9}$$

$$\Downarrow$$

$$\mathbf{L}^{(n)}(t) = \mathbf{L}^{(n+kN_{xdc})}(t) \tag{6.10}$$

As we saw the order in which the transducer elements emit does not matter. Two high resolution images at two consecutive emissions can be beamformed as follows:

$$\mathbf{H}^{(n)}(t) = \sum_{k=n-N_{xdc}+1}^{n} \mathbf{L}^{(k)}(t) \tag{6.11}$$

$$\mathbf{H}^{(n-1)}(t) = \sum_{k=n-N_{xdc}}^{n-1} \mathbf{L}^{(k)}(t) \tag{6.12}$$

Subtracting $\mathbf{H}^{(n-1)}(t)$ from $\mathbf{H}^{(n)}(t)$ gives:

$$\boxed{\mathbf{H}^{(n)}(t) = \mathbf{H}^{(n-1)}(t) + \mathbf{L}^{(n)}(t) - \mathbf{L}^{(n-N_{xdc})}(t)} \tag{6.13}$$

$L^{(n-N_{xdc})}(t)$

A high resolution line $H^{(n)}(t)$ at emission n can be obtained from the high resolution line $H^{(n-1)}(t)$ created at emission $n-1$ by adding the newly formed low-resolution line $L^{(n)}(t)$ and subtracting the low resolution line $L^{(n-N_{xdc})}(t)$ formed N_{xdc} emissions ago

emission
$(n-N_{xdc})$

$H^{(n-1)}(t)$

emission
$(n-1)$

$L^{(n)}(t)$ $H^{(n)}(t)$

current
emission # (n)

RF data

Figure 6.2: Recursive ultrasound imaging.

This gives a recursive formula for creating a high-resolution image from the previous high resolution one. The process is illustrated in Figure 6.2. The figure shows several consecutive emissions: from emission # $n-N_{xdc}$ to emission n.

Now one may ask the question about the advantages of the "*Recursive Ultrasound Imaging*". As a first thing, the calculation procedure is much simpler. If the number of elements N_{xdc} was, say 64, for every sample in the high resolution line $H^{(n)}(t)$ there would be 64 summations of the samples from the low-resolution lines $L^{(n)}(t)$, $n \in [1, N_{xdc}]$. In the new procedure the amount of summations is reduced to only 3, independent of N_{xdc}. However, the "real" advantage is that a new high-resolution image is created at every emission giving a high correlation between two consecutive high-resolution images. This makes it possible to estimate the motion of the tissue and the blood flow, something which is normally considered impossible to obtain using synthetic aperture imaging techniques.

Among the drawbacks is that the high resolution image still contains information obtained in the course of multiple transmit events. Therefore the method is susceptible to motion artifacts, as are all synthetic aperture methods. The decrease of these artifacts will be considered in Chapter 12.

Another of the draw-backs is the low signal-to-noise ratio, which is due to the use of only a single element in transmit. Solutions to this problem is given in Chapters 7 and 9.

In the paper [80] presented at 1999 IEEE INTERNATIONAL ULTRASONICS SYMPOSIUM we have given a generalized equation for Recursive Ultrasound Imaging:

$$\mathbf{H}^{(n)}(t) = \sum_{k=1}^{B} c_k \cdot \mathbf{H}^{(n-k)}(t) + \sum_{q=0}^{Q} b_q \cdot \mathbf{L}^{(n-q)}(t), \tag{6.14}$$

where B and Q are the numbers of high- and low-resolution images on which $\mathbf{H}^{(n)}(t)$ depends, and c_k and b_q are weighting coefficients. This formula was, however, never really explored in its full potential.

As previously discussed the RUI suffers from motion artifacts. They can be decreased by decreasing the time on which the high resolution image depends, i.e. by decreasing B and Q. In its extreme $B = 1$ and $Q = 0$, which leads to the Add-only Recursive Imaging.

6.2 Add-only recursive imaging

As mentioned the Add-only Recursive Ultrasound Imaging can be derived from the Generalized Recursive Ultrasound Imaging by letting $Q = 0$ and $B = 0$ in (6.14). The calculation procedure becomes:

$$\mathbf{H}^{(n)}(t) = c_1\mathbf{H}^{(n-1)}(t) + b_0\mathbf{L}^{(n)}(t) \tag{6.15}$$

The difference between equations (6.13) and (6.15) is that instead of being subtracted, the information obtained after the emission with element i decays exponentially with time. The system gradually "forgets" it. This approach has the advantage of having a much simpler implementation as shown in Figure 6.3. One can adjust the forgetting abilities of the system by changing the value of c_1.

An obvious drawback is that the so created high resolution image is not equal to the high-resolution image created using the "classical" recursive algorithm.

In the following we will try to show that these two images are similar.

Because of the recursive nature of the beamforming procedure, there will be some transient effects when the scanner is started, and when the scan position is changed. These transient effects will disappear shortly afterwards and are dependent on our "forgetting" constant c_1.

A normal high-resolution line created using synthetic transmit aperture is expressed as:

$$H(t) = \sum_{i=1}^{N_{xdc}} L_i(t), \tag{6.16}$$

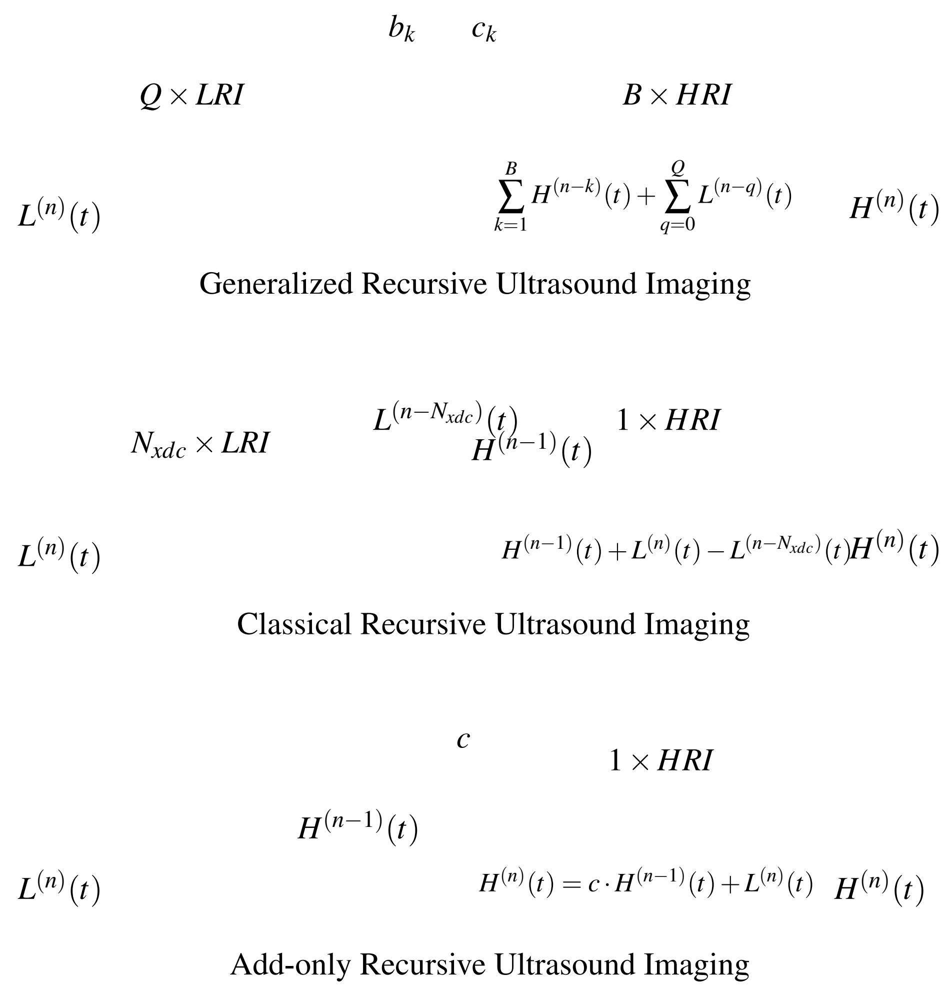


Figure 6.3: Block-scheme representation of the different variations of the Recursive Ultrasound Imaging.

where $L_i(t)$ are the low resolution lines obtained after emitting with element i.

In the recursive ultrasound imaging, a low resolution line $L_i(t)$ is obtained after every N_{xdc} emissions. In the classical approach, the old line $L_i(t)$ is replaced with the newly acquired one, thus keeping the sum constant. In the add-only algorithm, these lines are summed. It can easily be shown that this sum tends to $L_i(t)$, in other words $\sum_k c_1^{kN_{xdc}} L^{(n)}(t) \to L_i(t)$.

Assuming stationarity, ($L^{(n)}(t) = L^{(n-N_{xdc})}(t) \equiv L_i(t)$), the sum of the low-resolution lines $L_i(t)$ is:

$$L_{\Sigma;i}(t) = L_i(t) + c_1^{N_{xdc}} L_i(t) + c_1^{2N_{xdc}} L_i(t) + \cdots \tag{6.17}$$

This is a simple geometric series the sum of which is:

$$L_{\Sigma;i}(t) = (1 + c_1^{N_{xdc}} + c_1^{2N_{xdc}} + \cdots) L_i(t) \tag{6.18}$$

$$\Downarrow$$

$$L_{\Sigma;i}(t) = L_i(t) \cdot \frac{1}{1 - c_1^{N_{xdc}}} \tag{6.19}$$

If $c_1 = 0.9$ and $N_{xdc} = 64$, then $1/(1 - c^{N_{xdc}}) \approx 1$, and $L_{\Sigma;i}(t) \approx L_i(t)$.

Thus it was shown that for the add-only recursive imaging, the currently emitting channel contributes to the high resolution image almost the same information as for the case of normal synthetic transmit aperture.

The difference is that the contributions to the total sum from the neighboring channels is suppressed, or in other words “apodized”:

$$H(t) = \sum_{i=1}^{N_{xdc}} c_1^i \cdot L_{\Sigma;i}(t) \tag{6.20}$$

One should not immediately discard the ARUI as a non-viable imaging modality. For one thing, there is a certain amount of flexibility in the choice of the constant c_1. It can be a function of the transducer element number, i.e. $c_1 = c_1(i)$. An improvement in the image quality can be achieved by rearranging the order of transmissions[1], the frequency with which the individual elements transmit[2] and the weighting coefficients. As it will be seen in Chapter 11 the motion artifacts are dependent not only on the velocity of the tissue, but also on the distance between two consecutive transmitting elements. The decreased resolution, which is a result of the weighting is therefore to some extent compensated by the effect of motion artifacts.

6.3 Performance of the methods

In this section the performance of the CRUI and ARUI is discussed in terms of resolution and vulnerability to motion artifacts. The resolution is usually determined by the point spread function of the system and will be examined in the next section.

6.3.1 The point spread function

From Equation 6.13 it can be seen that the high resolution images in CRUI are formed from the exactly the same information as the high resolution images in standard synthetic transmit aperture focusing. Therefore the resolution is the same as in the standard B-mode imaging.

The point spread function of the ARUI can be deduced by the following simple considerations:

1. The high-resolution image is constantly multiplied with the forgetting factor c_1. This is equivalent to applying an additional apodization factor in transmit. Therefore the point-spread function will be broader in lateral direction.

2. The last acquired low resolution image will currently have the biggest weight on the high resolution sum. Seen from a fixed point, this results in a constant change in the angle from which the point is seen. Unless special care is taken to use emissions from the outermost element towards the innermost, the point spread function will be tilted.

Figure 6.4 shows a comparison between three point spread functions. They are a result of a measurement performed by the XTRA system (see Appendix J for description). The phantom was a wire lying in the elevation plane and parallel to the transducer surface. Only 13

[1] Nobody says that the relation between i and n is given *only* by Equation 6.2. The transmissions can be done from the outermost elements inwards to the center of the transducer

[2] For example the innermost elements can transmit twice more often than the outermost elements.

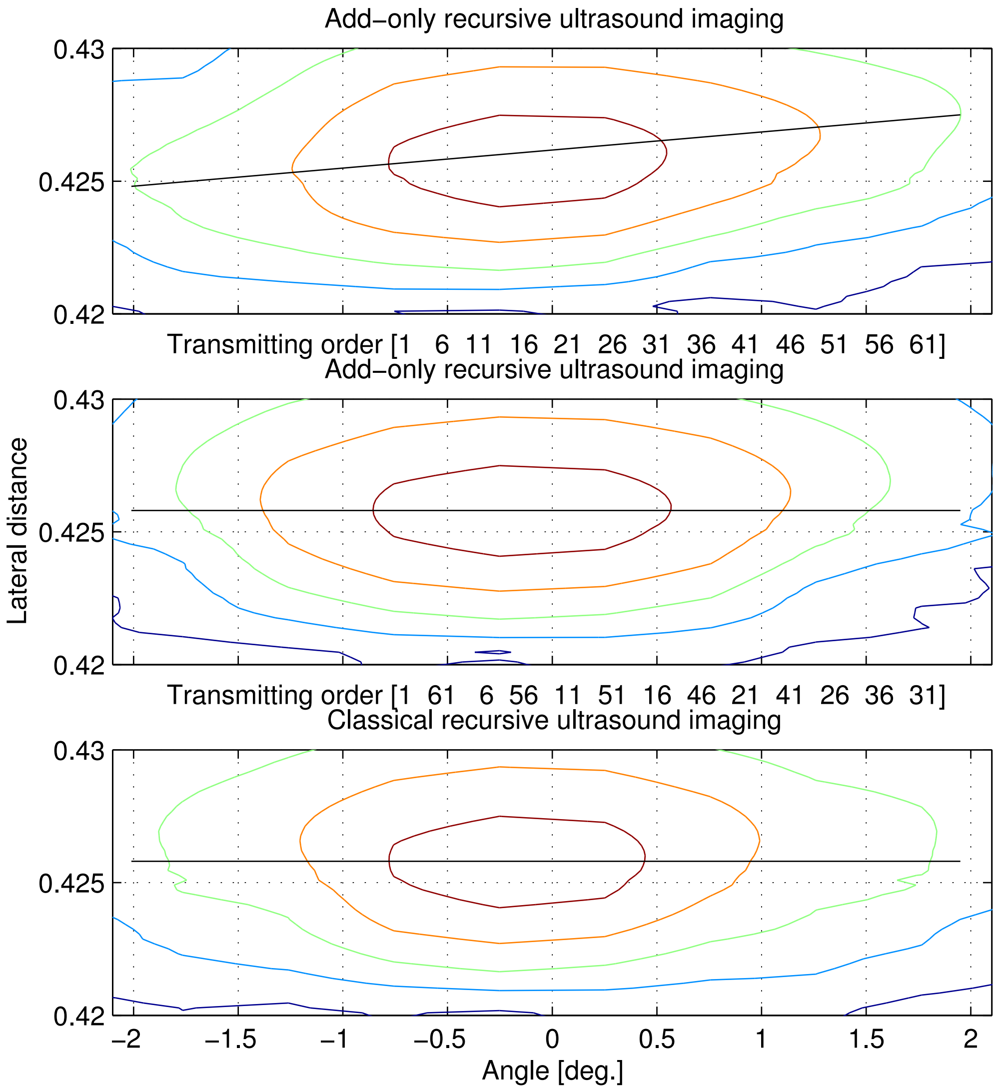
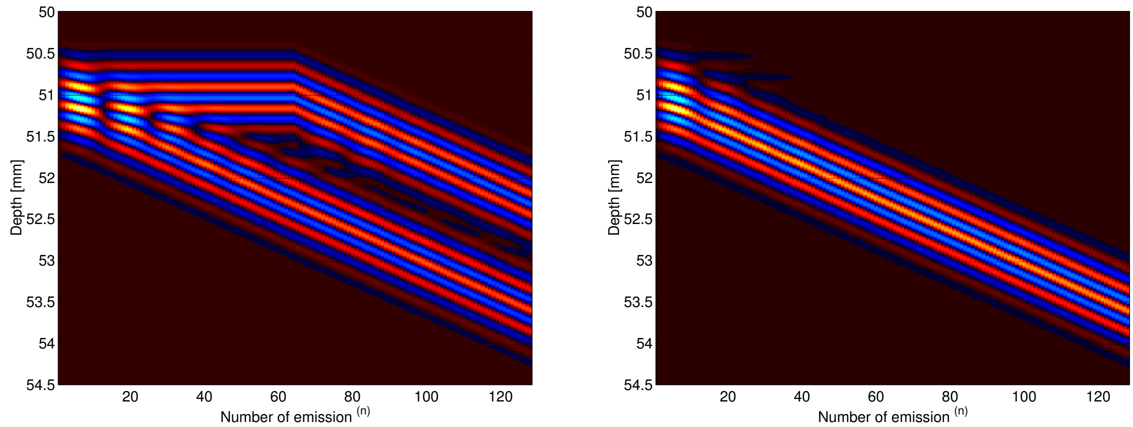


Figure 6.4: Comparison between the PSF of the CRUI and ARUI. The ARUI is given for two transmit sequences. The innermost contours are at level -6 dB. The neighboring contours are at levels 6 dB apart.

out of the 64 transducer elements were used in transmit (see Chapter 8 for details on sparse synthetic transmit aperture imaging). From the top-most plot it can be seen that the point-spread-function is tilted as discussed above. The middle plot shows that this can be compensated for by changing the order of the transmitting elements from the outermost inwards. The point-spread-function still remains wider, but is not skewed any more.

6.3.2 Motion artifacts

Figure 6.5 illustrates the source of motion artifacts in the case of classical recursive imaging and add-only recursive imaging. The figure was created by simulating a single scatterer moving away from the transducer. Initially the scatterer was positioned at coordinates $(x = 0, y = 0, z = 51 \text{ mm})$. After simulating the response of a single emission, the position of the scatterer was moved axially with a fixed step. Figure 6.5 shows a single beamformed high-resolution line as a function of emission number. The number of transmitting elements is $N_{xdc} = 64$.



(a) CRUI (b) ARUI

Figure 6.5: Illustration of the source of motion artifacts in recursive ultrasound imaging for (a) classical and (b) add-only recursive ultrasound imaging.

The classical recursive imaging "*remembers*" all of the last N_{xdc} emissions, and therefore up to emission number $n = 64$, the information from the initial position is still present in the image.

In the case of ARUI, the information is gradually forgotten, and therefore the final result is less affected by it. Figure 6.5 also shows the transient effects present in the beginning of the scanning procedure.

CHAPTER SEVEN

Virtual ultrasound sources

One of the drawbacks of the generic synthetic aperture, the synthetic transmit aperture, and recursive ultrasound imaging is the low signal-to-noise ratio (SNR) of the received signals. This is due to the low power output by a single transducer element of an array. One way to increase the transmitted energy is to use several elements in transmit. This was done by O'Donnell and colleagues in their intra-vascular imaging system [59]. Because of the array geometry the emitted waveform was diverging in a manner similar to the diverging wave generated by a single element. Multiple elements were used also by Ylitalo [81], who studied the signal-to-noise ratio in synthetic aperture ultrasound [64]. An extensive study of the phase front due to the use of multiple elements in transmit was carried out by M. Karaman and colleagues [65].

A completely separate thread of work was initiated by Passmann and Ermert in their high-frequency ultrasound system [82, 83]. In their work they treat the focal point of a single element transducer as a "*virtual source*" of ultrasound energy. This approach was studied by Frazier for different transducers with different F-numbers [84]. The virtual source elements were used by Bae and colleagues to increase the resolution and the SNR of a B-mode imaging system [76].

In this chapter, an attempt will be made to put all of the above approaches in a common frame, and an answer to the question "where is the virtual source?" will be sought.

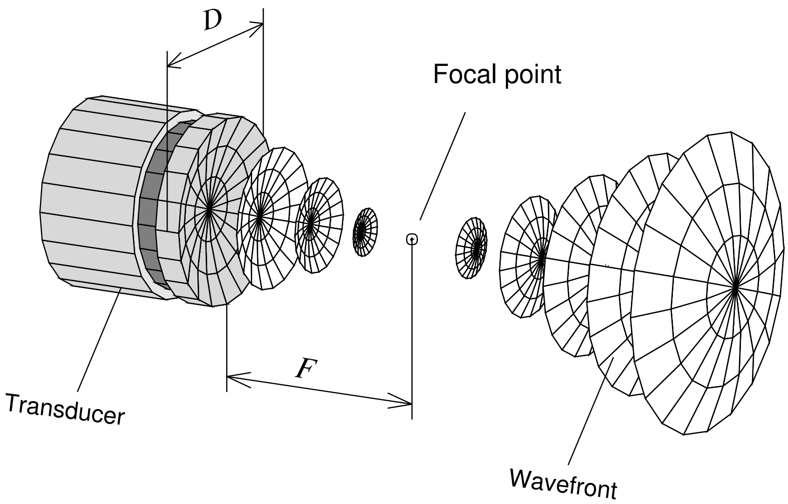


Figure 7.1: Idealized wavefront of a concave transducer.

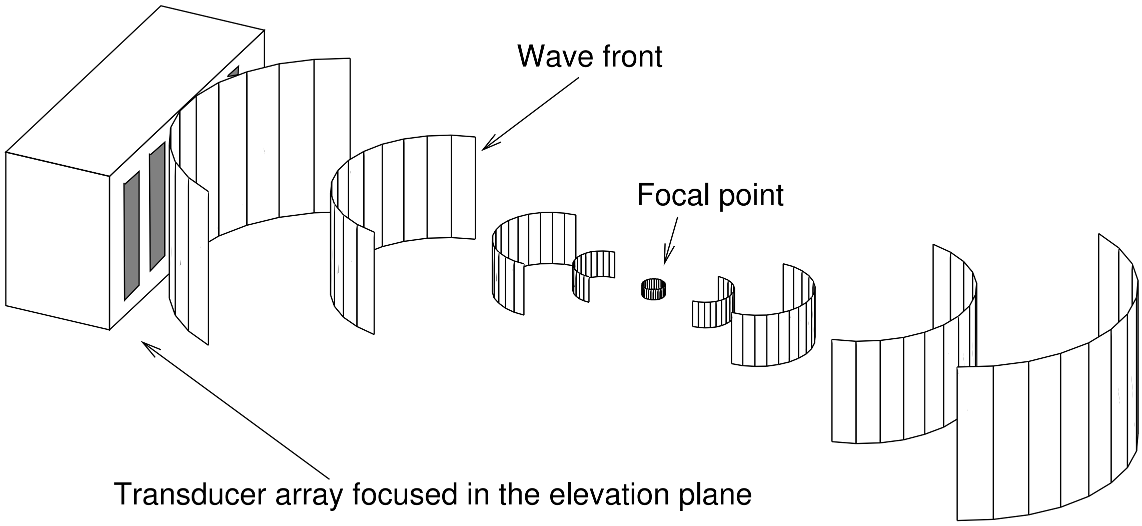


Figure 7.2: Idealized wavefront of a focused array transducer.

7.1 Model of virtual source element

Using geometrical relations one can assume that the wavefront of a transducer is a point at the focus as shown for a single element concave transducer in Figure 7.1. The wavefront before the focal point is a converging spherical wave, and beyond the focal point is a diverging spherical wave. Because of the shape of its wavefront the focal point can be considered as a "*virtual source*" of ultrasound energy. This type of arrays was used by Ermert and Passman in [82, 83] and by Frazier and O'Brien [84]. The focal point was considered as a virtual source and synthetic aperture focusing was applied on the signals recorded by the transducer.

Figure 7.2 shows the idealized wavefront for a focused linear array transducer. The generated wavefronts are cylindrical, because the elements have a height of several millimeters, while their width is a fraction of a millimeter. The width of the beam in the elevation plane is determined by the height of the elements and by the presence/absence of an acoustical lens. The shown array has an acoustical lens and therefore the height of the cylindrical waves initially converges. After the elevation focus the height of the cylinders increases again. The transducer is focused using electronic focusing in the azimuth plane. For the case depicted in Figure 7.2 the focus in the azimuth plane coincides with the focus in the elevation plane. Generally they can be different, with the possibility of a dynamic focusing in the azimuth plane. This type of focusing was used by Bae and colleagues [76] to increase the resolution and the signal-to-nose ratio in B-mode images by recalculating the delays for the dynamic focusing.

In the situations discussed so far the virtual sources were thought to be the focal points *in front* of the transducer. In the following the idea will be extended to virtual sources that are *behind* the transducer.

In quite a number of papers multiple elements have been used in transmit. As it was shown in Chapter 4, the synthetic aperture focusing relies on spherical waves being generated by the transducer. Usually a small number of elements is used to transmit. Due to the small size of the active sub-aperture the wave starts to diverge quite close to the transducer, as for the case of a planar transducer. Such approach was used by Ylitalo [81]. This approach, however, limits the size of the active sub-aperture and hereby the number of elements used in transmit.

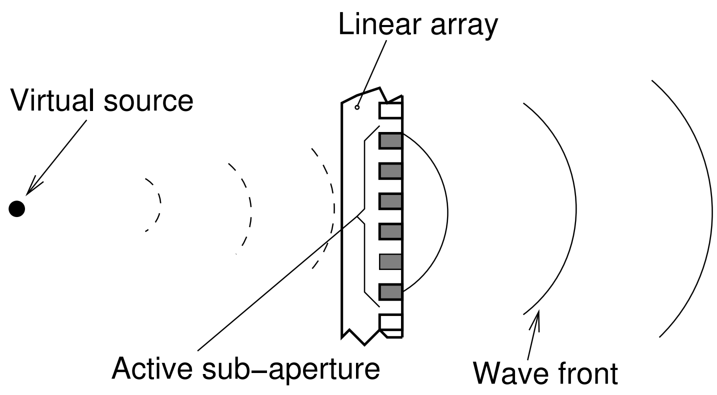


Figure 7.3: The virtual source of a ”defocused” linear array lies behind the transducer

Another approach is to ”defocus” the elements by transmitting with the central element first and with the outermost elements last. The created waveform has the shape of a diverging cylinder as shown in Figure 7.3. As it is seen from the figure, one can assume that there is a virtual source of ultrasound energy placed behind the transducer. This fact was, however, not realized in the early papers, and the delays were calculated using different approximations such as a parabolic fit [65], or as an attempt to align the wavefronts in front of the transducer [80]. The suggested approaches work, but a common framework for high-precision delay calculations is missing. Such a frame-work is provided by the use of a virtual source element, placed behind the transducer. In all of the considered papers the focusing is calculated with respect to the coordinates of the central element, which is called a ”phase center”. As we will see in the following a better estimate for the ”phase center” is provided by the coordinates of the virtual source.

7.2 Virtual source in front of the transducer

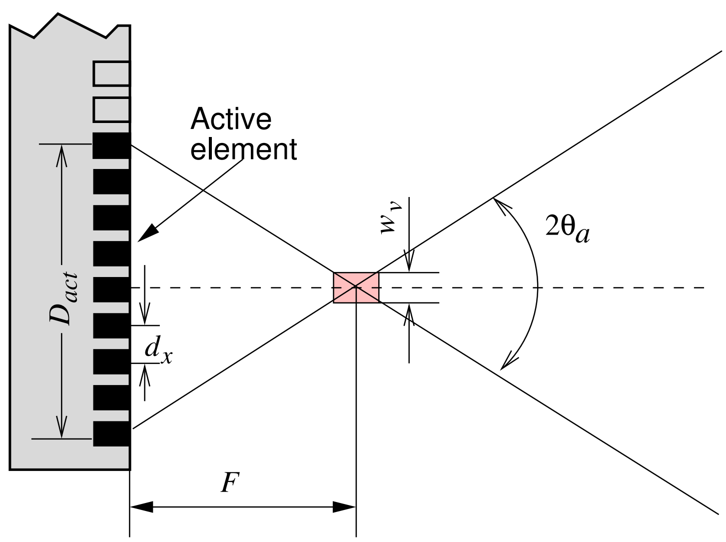


Figure 7.4: The geometry of a linear array involved in creating a virtual source.

This section starts with the introduction of some of the "parameters" of the virtual source, namely its position and directivity pattern. Figure 7.4 illustrates the geometry associated with the creation of a virtual source in front of the transducer. The delays that are applied in transmit are:

$$\tau_i = \frac{1}{c}\left(F - \sqrt{F^2 + ((i-1)d_x - D_{act}/2)^2}\right) \tag{7.1}$$
$$i \in [1, N_{act}]$$
$$D_{act} = (N_{act} - 1)d_x,$$

where i is the index of the transmitting element, F is the focal depth, and N_{act} is the number of elements in the active sub-aperture. In this case the virtual source is located at coordinates $(x_c, 0, F)$, where x_c is the coordinate of the center of the active sub-aperture.

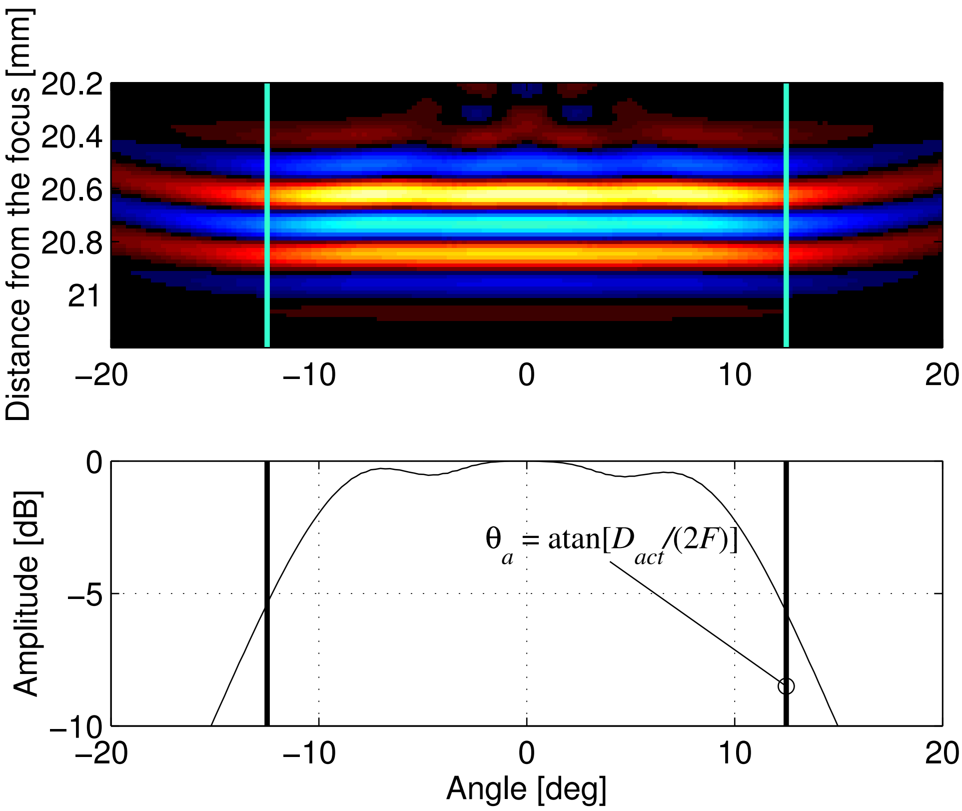


Figure 7.5: Simulated pressure field of a linear array transducer.

If a fixed focus is used in transmit, then the received signals can be processed using a synthetic aperture focusing [76], treating the received signals as echoes which are a result of a transmission with the virtual source element. The virtual ultrasound sources form a virtual synthetic transmit aperture. The main problem is to determine the useful span of the synthesized aperture L.

Consider Figure 4.4 on page 40. Let the opening angle be $2\theta_a$. The size of the virtual synthetic array used to beamform data at depth z is in this case:

$$L = 2(z - F)\tan\theta_a \tag{7.2}$$

The opening angle θ_a can be determined by simple geometric considerations as (see Figure 7.4):

$$\theta_a = \arctan \frac{D_{act}}{2F} \tag{7.3}$$

Figure 7.5 shows the simulated pressure field of a linear array. The points at which the field was simulated are placed along an arc with the focal point being at the center. The focal distance F is equal to 25 mm. The transducer pitch is $d_x = 208\ \mu$m. The number of elements is $N_{act} = 64$. The estimated angle is :

$$\begin{aligned} \theta_a &= \arctan \frac{N_{act} d_x}{2F} \\ &= \arctan \frac{11.07\ [\text{mm}]}{25\ [\text{mm}]} \\ &\approx 12.4^\circ \end{aligned} \tag{7.4}$$

and is shown in the figure as a thick line. It can be seen that θ_a almost coincides with the -6 dB drop of the amplitude. For the simulated case θ_a is also the angle at which the wavefront starts to differ from a spherical wave. The drop in the amplitude is due to misalignment of the waves transmitted by the individual elements. Thus the amplitude level can be used as an indication of the opening angle of the virtual source element.

Figure 7.6 shows the correspondence between the estimated angle using geometrical relations and the actual pressure field as a function of the focal strength. The plots are functions of the F-number, $f_\# = F/D_{act}$. It can be seen that the geometry shown in Figure 7.4 manages to delineate the -6 dB signal level. The aperture size used to create the plots is 4 mm. The fields are normalized with respect to the highest value at each depth.

7.3 Virtual source behind the transducer

As it was pointed out, in order to use synthetic aperture focusing, the transmitted wavefront should approximate a spherical or a cylindrical wave[1]. One way to create a cylindrical wave using a linear array is to fit the emitted wavefront along a circle. The geometry for this approach is shown in Figure 7.7. The idea is for the signals emitted by the elements in the active aperture to arrive at the same time along a circle with radius R. The width of the active aperture is $D_{act} = d_x N_{act}$, where d_x is the pitch of the transducer. The elements forming the sub-aperture are assigned relative indexes $k \in (-\lfloor (N_{act}+1)/2 \rfloor, \lfloor (N_{act}+1)/2) \rfloor)$. In order to create a spherical wave, the elements closest to the center of the sub-aperture emit first. If $R \geq D_{act}$ is a constant depending on the size of the aperture, then the transmit delays are determined by:

$$\tau_k = \frac{R - \sqrt{R^2 - x_k^2}}{c}, \tag{7.5}$$

where c is the speed of sound and x_k is the distance of the element from the center of the sub-aperture. If $R \to \infty$ then a planar wave is created.

[1]In fact the only two requirement for the radiated field is to be as large as possible and the shape of its wavefront to be known. The broadness of the wave field will result in a bigger phase difference in the received signals. This phase difference carries the spatial information. The knowledge of the wavefront is necessary for the beamformation.

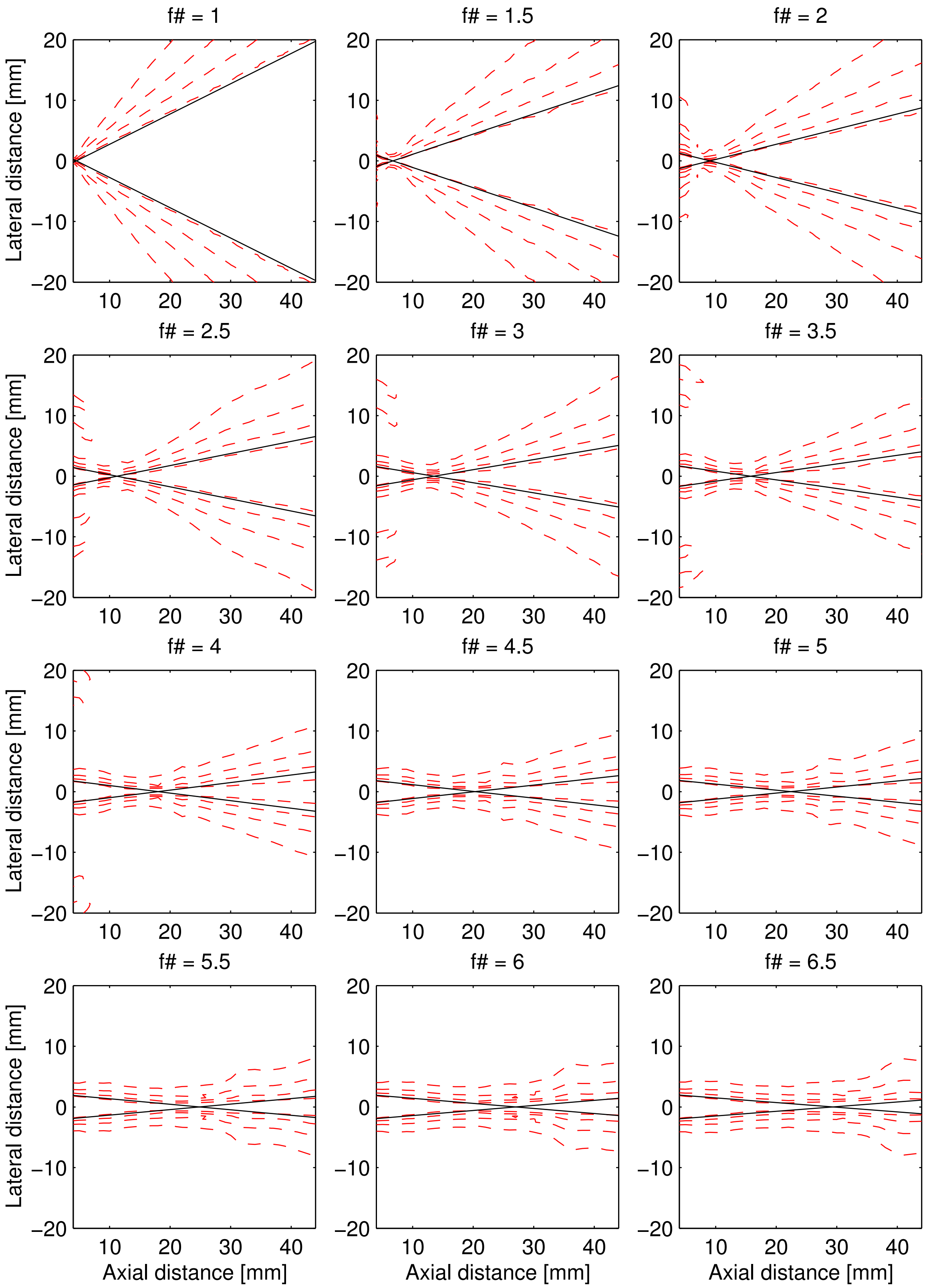


Figure 7.6: Simulated ultrasound fields. The contours given with red dashed lines start at a level of -6 dB and are drawn at levels -6 dB apart. The black lines connect the edges of the transducer with the focal point. Not that the axial distance starts from the 4th mm.

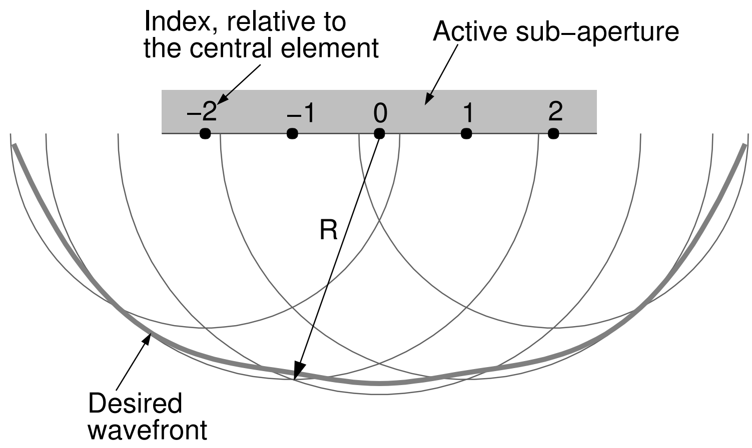


Figure 7.7: Defocusing by attempting to approximate a desired wavefront.

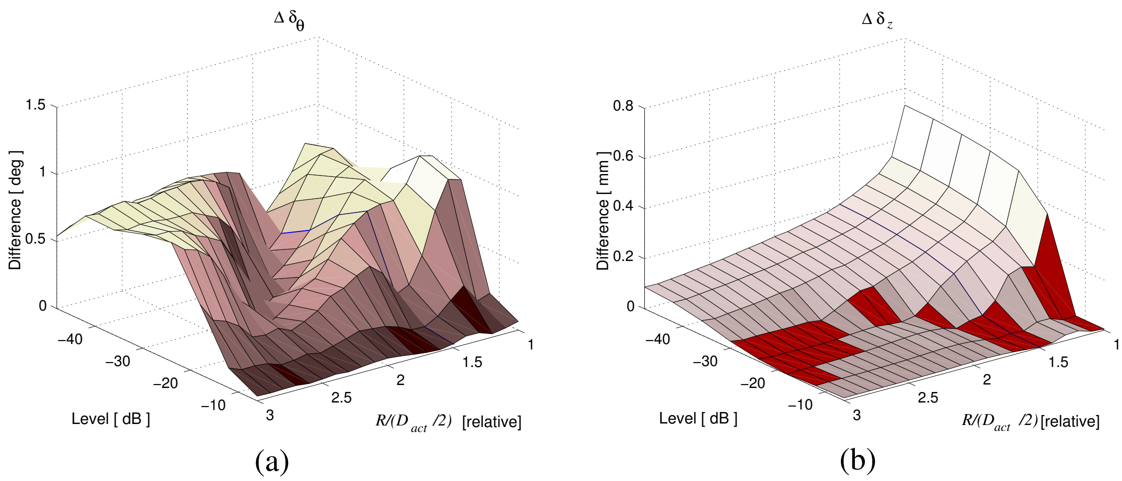


Figure 7.8: The difference between the PSF using 11 elements and a reference PSF using a single element. Plot (a) shows the difference in lateral direction, and plot (b) shows the difference in the axial direction.

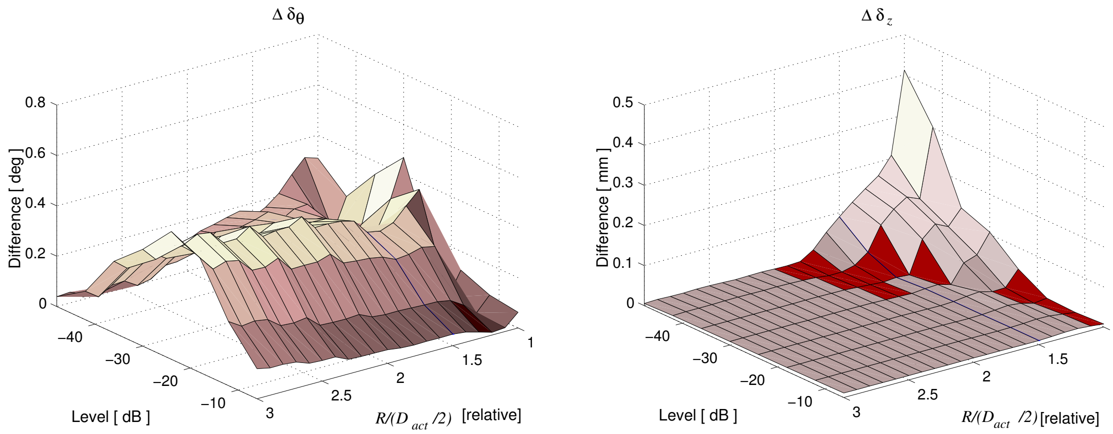
This approach was investigated using the XTRA system and the optimum for 11 transmit elements was found to be at $R = 1.5D_{act}/2$. This minimum was found by comparing the lateral dimension of point-spread functions obtained by emitting with a single element and by emitting by a defocused transmit using $N_{act} = 11$ active elements.

Figure 7.8 shows the difference in the lateral and axial size of the point spread functions for different values of R. In the figure, δ_θ is the beamwidth in the azimuth plane and δ_z is the pulse length in axial direction. $\Delta\delta_\theta$ and $\Delta\delta_z$ are the differences between these parameters obtained by using multiple elements and a single element:

$$\Delta\delta_{\theta\mathrm{XdB}}(R) = \delta_{\theta\mathrm{XdB}\ \mathrm{N_{act}=11}}(R) - \delta_{\theta\mathrm{XdB}\ \mathrm{N_{act}=1}}$$
$$\Delta\delta_{z\mathrm{XdB}}(R) = \delta_{z\mathrm{XdB}\ \mathrm{N_{act}=11}}(R) - \delta_{z\mathrm{XdB}\ \mathrm{N_{act}=1}}$$

This difference is given for different levels, from -6 to -48 dB, as a function of R. The values of R vary in the range $1 \le \frac{R}{D_{act}/2} \le 3$ at a step of 0.1. It must be noted that the position of the virtual source was not taken into account, and the delays were set as if the spherical wave emanates from the center of the active sub-aperture. The blue line shows the minimum mean squared error over all the levels for $\Delta\delta_\theta$.

The finite size of the active sub-aperture gives a rise to edge waves which generally add coherently outside the intended focal point. Apodization should generally decrease them as seen from Figure 7.9. It can be seen that the difference in the angular and lateral sizes of the point spread function are smaller when apodization is applied.



(a) (b)

Figure 7.9: The difference between the PSF using 11 elements with apodization applied on the active elements, and a reference PSF using a single element. Plot (a) shows the difference in lateral direction, and plot (b) shows the difference in the axial direction

Because both approaches show a minimum in the lateral size of the PSF for $R = 1.5D_{act}/2$, their lateral size is compared for this value of the radius in Figure 7.10. The difference is most pronounced in the level region between -24 and -40 dB. Keeping in mind that the apodization decreases the overall transmitted energy, this difference is not sufficient to justify the applied apodization. Additionally the attenuation in tissue will attenuate the edge waves.

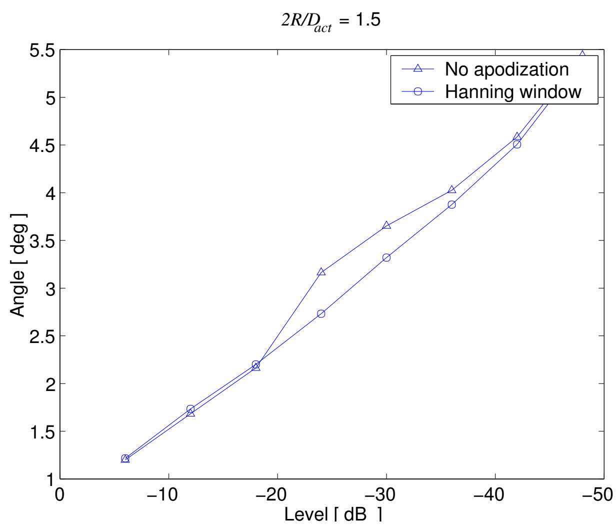


Figure 7.10: The angular size of the PSF with and without apodization in transmit. The apodization was a Hamming window.

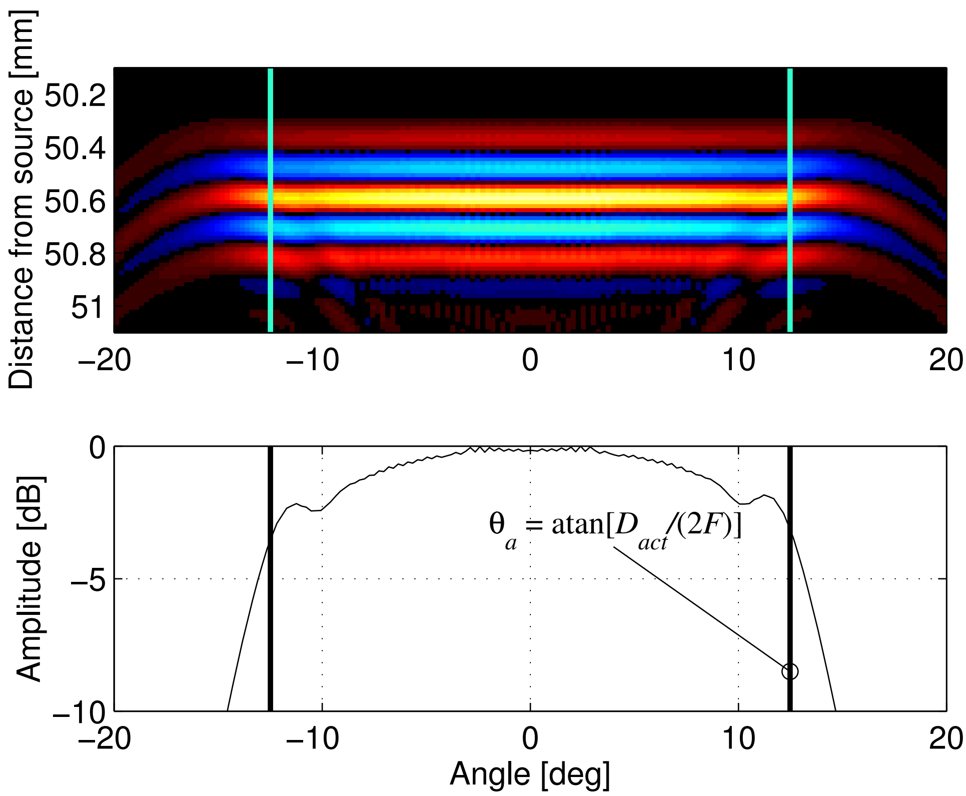


Figure 7.11: The wavefront created by a virtual source behind the transducer. The plot is at a constant distance from the virtual source.

Another way of calculating the delays was suggested by Karaman and colleagues [65]. They introduce a focal distance F, and calculate the delays as:

$$\tau_i = \frac{1}{c}\frac{x_i^2}{2F}, \tag{7.6}$$

where x_i is the coordinate of the transmitting element relative to the center of the active sub-aperture. This formula is a parabolic approximation to the real spherical profile that must be applied:

$$\tau_i = \frac{1}{c}\left(-F + \sqrt{F^2 + ((i-1)d_x - D_{act}/2)^2}\right) \tag{7.7}$$

$$i \in [1, N_{act}]$$

$$D_{act} = (N_{act} - 1)d_x,$$

The wavefront created by applying this delay profile is shown in Figure 7.11

The virtual element can be created at any spatial position $\vec{x}_v = (x_v, y_v, z_v)$. The equation for the delays becomes:

$$\tau_i = \mathrm{sgn}(z_v - z_i)\frac{1}{c}\left(|\vec{x}_i - \vec{x}_v|\right) \tag{7.8}$$

where $\vec{x}_i$ and $\vec{x}_v$ are the coordinates of the of the transmitting element and the virtual source, respectively. The function $\mathrm{sgn}(a)$ is defined as:

$$\mathrm{sgn}(a) = \begin{cases} 1 & a \geq 0 \\ -1 & a < 0 \end{cases}$$

The virtual sources are used to increase the signal-to-noise ratio. For the virtual sources behind the transducer it increases with $10\log_{10}N_{act}$ [65]. In this case the use of virtual source in front of the transducer has the advantage of being "closer" to the imaging region. Because the pulses from the elements *do* physically sum in the tissue at the focal point, it is possible to achieve high pressure levels comparable to those obtained by the conventional transmit focusing. This gives the possibility to exploit new fields of synthetic aperture imaging such as non-linear synthetic aperture imaging.

Another use of the virtual source elements is to achieve spacing between the transmissions less than $\lambda/4$ as required for the mono-static synthetic aperture imaging.

In the 2D B-mode imaging only a cross-section of the body is displayed. The beamwidth in the elevation plane is, however, neither narrow, nor uniform as shown in Figure 7.2 on page 63. It can be seen that the wavefront in the elevation plane first converges towards the elevation focus, and then, after the focus, it diverges. The thickness of the cross-section 100 mm away from a transducer, which is 4 mm high and is focused in the elevation plane at 20 mm, can reach 16 mm. This is a significant problem when scanning a volume plane by plane. It is possible, however, to extend the concept of virtual sources to the elevation plane and use it in order to increase the resolution after the elevation focus. The next section outlines a post-focusing procedure based on virtual source in the elevation plane.

7.4 Virtual source in the elevation plane

In the last years the 3D ultrasound scanners have firmly taken their place in routine medical exams [9]. Only a single system using 2D matrix arrays and capable of acquiring more than 10 volumes per second has been developed and commercially available up to this moment [6]. Most of the systems use a combination of electronic beam steering in the azimuth plane and some kind of mechanical motion in the elevation plane to scan the whole volume [9]. The motion of the transducer can either be controlled mechanically [9, 85, 86] or be a free hand scan. In the latter case the position of the transducer must be determined using some tracking device [87, 88, 8, 10].

Figure 7.12 shows such a mechanically scanned volume. The planes are parallel to each other in the elevation direction (along the y axis). In each of the planes the beam is electronically focused. Due to the dynamic focusing upon reception, the beam is narrow. In order to obtain good focusing 1.5 and 1.75 dimensional[2] are used which improve the imaging abilities of the system [89]. The complexity of the design of the array increases significantly. For example if the array has 8 rows in the elevation direction and 128 elements in the azimuth direction, then the total number of elements is 1024. A more widespread approach is to use acoustic lens. The beam profile after the focal point of the transducer diverges[3] as shown in Figure 7.13. The figure shows the simulated beam profile of a model of a B-K 8804 transducer. If this transducer is used for 3D scanning, then the generated images will have different resolutions in the azimuth and elevation planes. Using up to 128 elements and dynamic apodization the resolution in the azimuth plane can be sustained at say 1.5 - 2 mm, while in the elevation plane it can become as large as 7-10 mm for the same depth. Wang and colleagues [90] suggested the

[2]The terms were coined in order to distinguish these transducers from the "true" 2D matrix arrays which can steer the beam in the elevation direction. The 1.5 and 1.75-D transducers have limited capability to control the beam profile [48].

[3]This divergence depends on the F-number of the elevation focus.

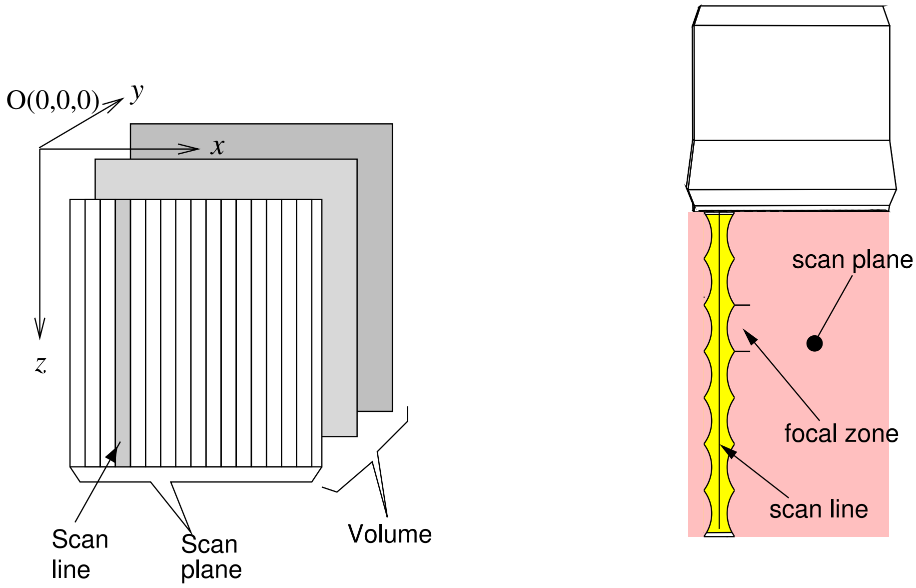


Figure 7.12: Volumetric scan using a linear array. The volume is scanned plane by plane. The beam is electronically focused in the plane. Due to multiple focal zones a tight beam is formed. From plane to plane the transducer is moved mechanically.

use of deconvolution in the elevation plane. An alternative approach is to use matched filtering instead, which can be looked upon as synthetic aperture imaging in the elevation plane [91].

The approach relies on the assumption that the ultrasound field is separable in the elevation and azimuth planes and that the focusing in the azimuth plane does not influence the properties of the beam in the elevation plane. The focal point in the elevation plane can be treated as a virtual source of ultrasound energy. The wave front emanating from this point has a cylindrical shape within the angle of divergence as for the case of a virtual source element in front of the transducer. The angle of divergence to a first order approximation is:

$$\theta_a = \arctan \frac{h}{2F}, \tag{7.9}$$

where h is the height of the transducer (the size in the elevation plane) and F is the distance to the fixed focal point in the elevation plane.

Figure 7.14 shows the creation of a "virtual array". The real transducer is translated along the y axis. Let the distance between two positions successive positions be Δy. At every position m an image is formed consisting of scan lines s_{lm}, $1 < l < N_l$. For notational simplicity only a single scan line from the image will be treated and the index l will be further skipped. The line in question will be formed along the z axis and positioned in the center of the aperture. The collection of focal points lying on the scan line positioned along the y axis (see: Figure 7.14) forms the virtual array. The virtual pitch (the distance between the centers of the virtual elements) of the virtual array is equal to the step Δy. The width of the virtual elements is assumed to be equal to the -6 dB beam width of the beam at the focus in the elevation plane. It is possible for the virtual pitch to be smaller than the virtual width. Each of the beamformed scan lines (whose direction and position relative to the transducer remains constant) is considered as the RF signal obtained by transmitting and receiving with a single virtual element. This corresponds to the case of generic synthetic aperture imaging. These RF signals can be passed to a beamformer for post focusing in the elevation plane.

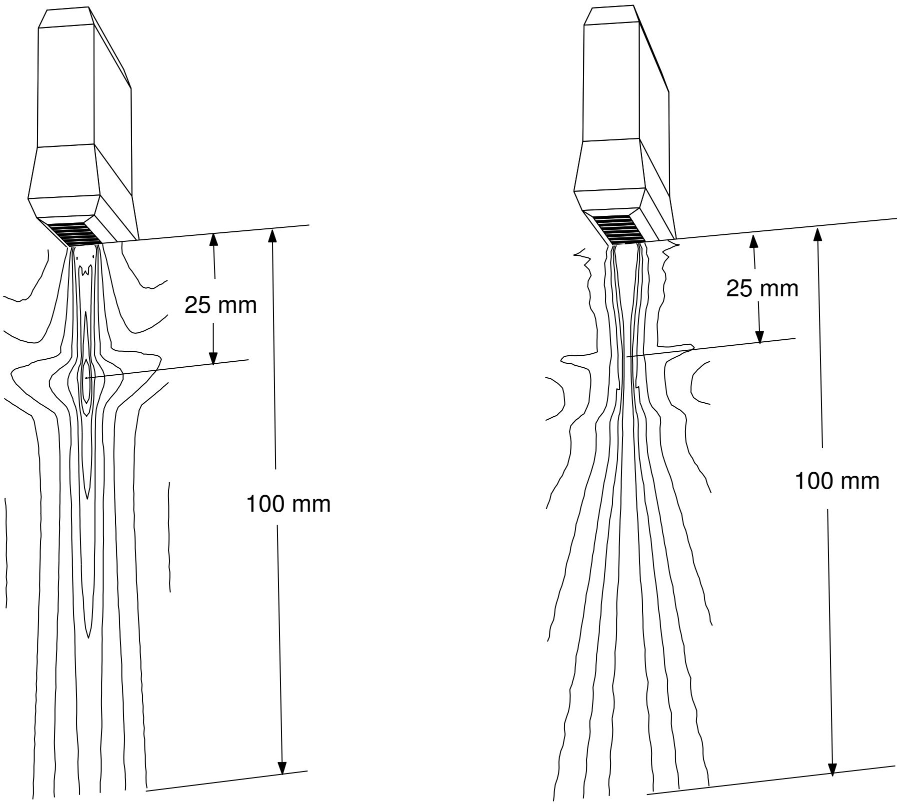


Figure 7.13: Peak pressure distribution in the elevation plane. The contours are shown at levels 6 dB apart. Sub-plot (a) shows the pressure distribution normalized to the peak at the focus, and sub-plot (b) shows the pressure distribution normalized to the maximum at each depth. The transducer has a center frequency of 7 MHz and 60 % fractional bandwidth. The excitation is a 2-cycles Hanning weighted pulse. The height of the transducer is 4.5 mm, and the elevation focus is 25 mm away from the transducer surface.

The final lines in the volume $S(t)$ are formed according to:

$$S(t) = \sum_{m=1}^{N_{pos}} a_m(t) s_m(t - \tau_m(t)), \tag{7.10}$$

where $a_m(t)$ are apodization coefficients, $\tau_m(t)$ are the delays applied, and N_{pos} is the number of positions. The coefficients a_m must be controlled in order not to exceed the size L covered by the virtual array (see Figure 4.4 on page 40). Taking into account the position of the virtual elements, the delays are:

$$\tau_m(t) = \frac{c}{2}\left[z - F - \sqrt{(z-F)^2 + \left(\left(m - 1 - \frac{N_{pos}-1}{2}\right)\Delta y\right)^2}\right] \tag{7.11}$$

$$t = \frac{2z}{c}, \tag{7.12}$$

where c is the speed of sound, z is the depth at which the line is formed[4] and F is the distance

[4]This is valid only for scan lines which are perpendicular to the surface of the transducer. If in each plane a

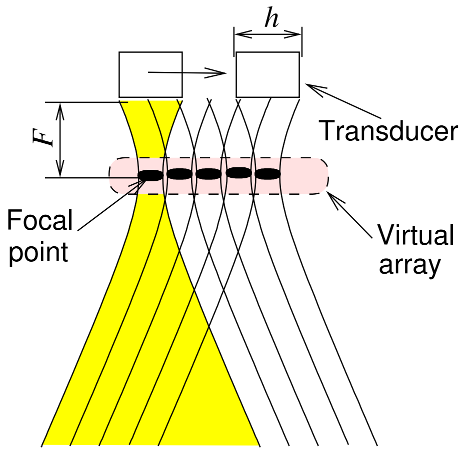


Figure 7.14: Forming a virtual array from the focal points of the scan lines.

to the elevation focus. The best obtainable resolution in the elevation plane is [83][5]:

$$\delta y_{6\mathrm{dB}} \approx m \frac{0.41}{\tan \frac{\theta_a}{2}}, \tag{7.13}$$

where λ is the wavelength. The variable m ($m \geq 1$) is a coefficient dependent on the apodization. For a rectangular apodization m equals 1 and for Hanning apodization it equals to 1.64. Substituting (7.9) in (7.13) gives:

$$\delta y_{6\mathrm{dB}} \approx 0.82 m \frac{F}{h} \tag{7.14}$$

Equation (7.14) shows that the resolution is depth independent. However, this is true only if the number of transducer positions is large enough to maintain the same F-number for the virtual array as a function of depth. For real-life applications the achievable resolution can be substantially smaller.

Lockwood et al. [1] suggested that synthetic aperture imaging should be used in the azimuth plane to acquire the data. To increase the transmitted energy multiple elements are used in transmit. The transmit delays applied on them form *virtual elements* behind the transducer. The approach can be extended by sending the lines from the high-resolution images to a second beamformer, which utilizing the concept of virtual sources in the elevation plane increases the resolution in the elevation plane. The process is summarized and illustrated in Figure 7.15.

The approach was verified using simulations and measurements.

7.4.1 Simulations

The simulations were done using Field II [19]. The parameters of the simulation are given in Table 7.1, and were chosen to model the XTRA system used for the measurements.

sector image is formed, then z and F must be substituted with the distances along the beam.

[5]This formula is not derived neither in the paper in question, nor in any of the materials referenced by it. The same equation (taken from the same paper) is, however, quoted by Frazier and O'Brien [84]. The author assumes that it has been verified by Passman and Ermert [83] and therefore it is used in the dissertation.

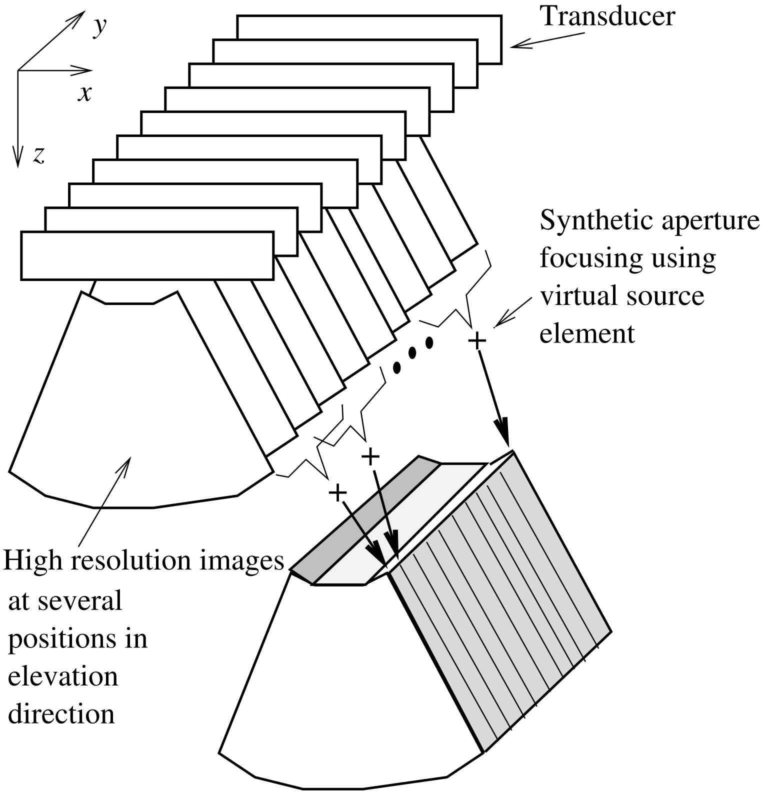


Figure 7.15: Two stage beamforming for 3D volumetric imaging. In the first stage high-resolution images are formed using synthetic transmit aperture focusing. In the second stage each of the RF scan lines from the high resolution images is post focused to increase the resolution in the elevation plane.

Parameter name	Notation	Value	Unit
Speed of sound	c	1480	m/s
Sampling freq.	f_s	40	MHz
Excitation freq.	f_0	5	MHz
Wavelength	λ	296	μm
-6 dB band-width	B	4.875 - 10.125	MHz
Transducer pitch	d_x	209	μm
Transducer kerf	$kerf$	30	μm
Number of elements	N_{xdc}	64	-
Transducer height	h	4	mm
Elevation focus	F	20	mm

Table 7.1: Simulation parameters.

Figure 7.16 shows the 3D structure of the point spread function before and after synthetic aperture post focusing in the elevation plane. The plots show the -30 dB isosurface of the data. The high resolution images in the azimuth plane are phased array sector images and the point spread function is shown before scan conversion. The lateral direction is therefore displayed in degrees. The step in the elevation direction is $\Delta y = 0.5$ mm, which is 1.7λ. The total number of positions is 95. The number of positions N_{pos} used in the post beamforming of the high

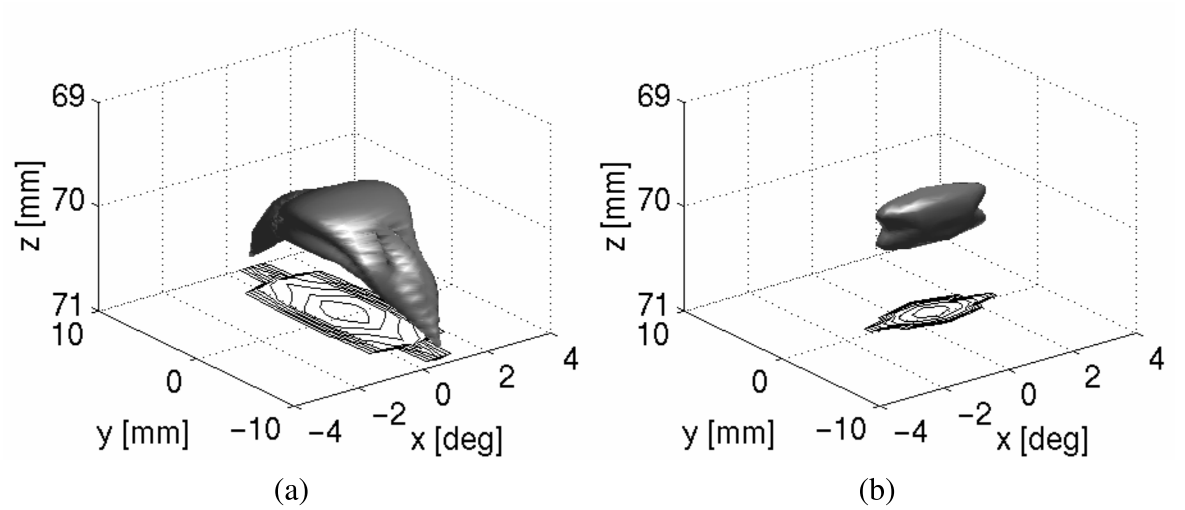


Figure 7.16: The 3D point spread function, (a) before, and (b) after synthetic aperture post focusing in the elevation direction. The surface of the point spread function is reconstructed at a level of -30 dB.

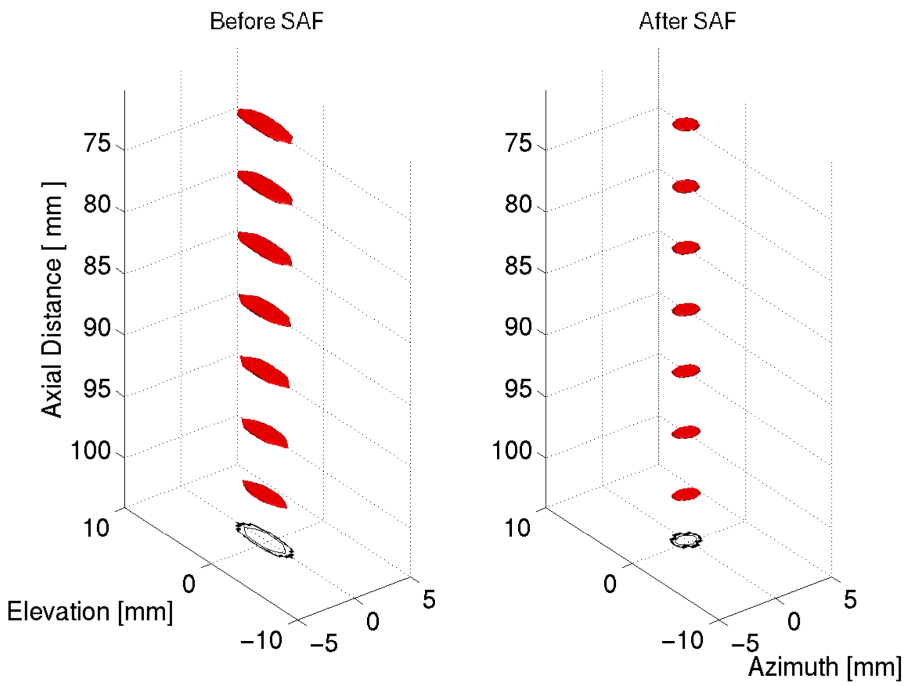


Figure 7.17: The point spread function as a function of depth. The surface is reconstructed at a level of -10 dB.

resolution lines is 40.

Figure 7.17 shows the result of a simulation of the point spread function for different depths and Table 7.2 gives the values of the -6 dB beam width. The seven point scatterers are positioned at depths ranging from 70 to 100 mm. The step in the elevation plane is larger, $\Delta y = 0.7$ mm. The number of lines N_{pos} is reduced to 30. From Table 7.2 it can be seen that the lateral size of

Depth [mm]	δx_{6dB} [mm]	Before SAF δy_{6dB} [mm]	After SAF δy_{6dB} [mm]
70	1.44	4.78	1.72
75	1.54	5.16	1.72
80	1.65	5.48	1.72
85	1.75	5.80	1.85
90	1.85	6.18	1.85
95	1.96	6.56	1.85
100	2.06	6.75	1.97

Table 7.2: The resolution at -6 dB as a function of depth.

the point spread function in the azimuth plane increases linearly with depth:

$$\delta_{x6\mathrm{dB}} = z \sin \delta_{\theta 6\mathrm{dB}}, \tag{7.15}$$

where $\delta_{\theta 6\mathrm{dB}}$ is the angular size of the point spread function at -6 dB in polar coordinates. The elevation resolution $\delta_{y6\mathrm{dB}}$ prior to the synthetic aperture focusing also increases almost linearly with depth, which is consistent with the model of the virtual source element. After applying the synthetic aperture focusing, $\delta_{y6\mathrm{dB}}$ becomes almost constant as predicted by (7.14). A Hanning window was used for the apodization coefficients, which corresponds to broadening of the point spread function at -6 dB approximately 1.6 times ($m \approx 1.6$ in (7.14)). Substituting $h = 4$ mm, $F = 20$ mm, and $\lambda = 0.296$ mm gives $\delta_{y6\mathrm{dB}} \approx 1.87$, which is in agreement with the results from the simulations.

7.4.2 Measurements

The measurements were done with the XTRA system. The parameters of the system are the same as the ones given in Table 7.2. Two experiments were conducted:

1. A point scatterer mounted in an agar block 96 mm away from the transducer was scanned. The step in the elevation direction was $\Delta y = 375$ μm. The diameter of the point scatterer was 100 μm.

2. A wire phantom was scanned at steps of Δy = 700 μm. The wires were positioned at depths from 45 to 105 mm, 20 mm apart. At every depth there were two wires perpendicular to each other. The diameter of the wires was 0.5 mm.

The purpose of the first experiment was to compare the resolution of the measurement with the resolution obtained in the simulations. The purpose of the second experiment was to show that the resolution obtained in the elevation and azimuth planes for different depths are comparable. Both experiments were done using steps Δy bigger than one wavelength to show that grating lobes, if present, have low levels due to low transducer sensitivity at large angles in the elevation plane (due to the focusing lens). In none of the experiments the level of grating lobes was above -40 dB, and did not appear outside the -6 dB beamwidth of the original point spread function (before post focusing).

Figure 7.18 shows the measured point point scatterer. The contours are drawn starting from -6 dB and are 6 dB apart. The plot was obtained using $N_{pos} = 60$ RF scan lines in the post

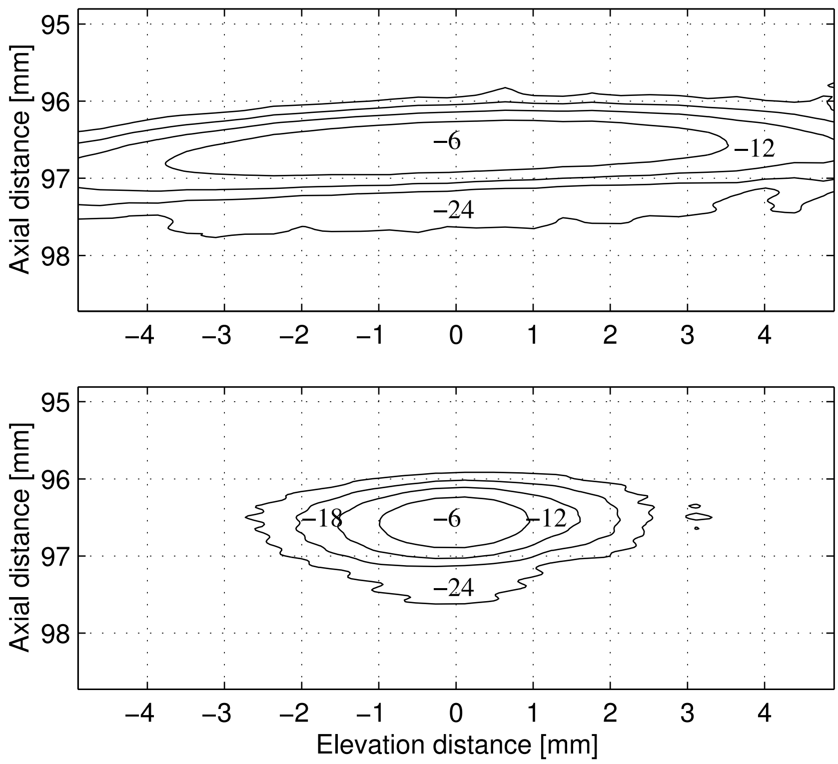
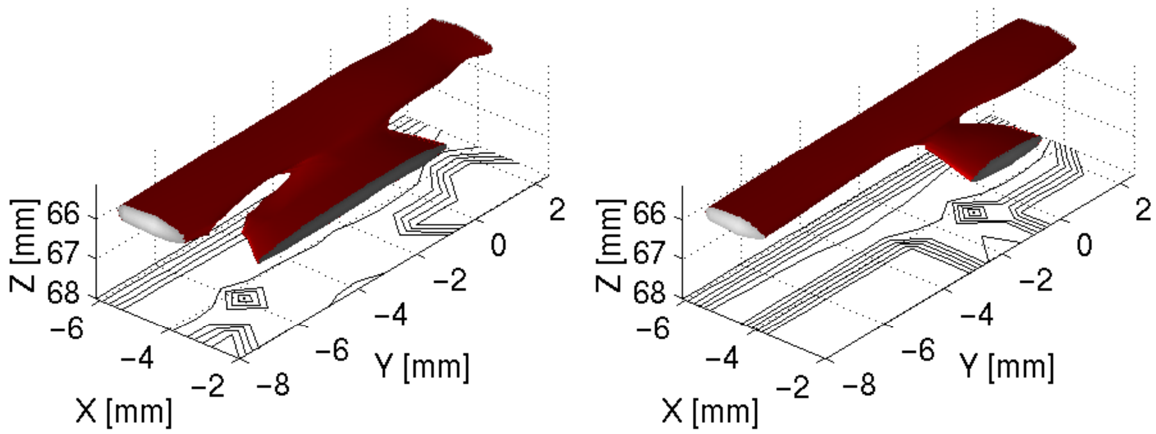


Figure 7.18: PSF in the elevation plane: (top) before and (bottom) after synthetic aperture post focusing. The numbers show the level in the respective region, i.e., the first contour line delineates the -6 dB region.



(a) (b)

Figure 7.19: The wire phantom: (top) before, and (bottom) after post focusing. The surface is reconstructed from the volumetric data at a level of -10 dB.

focusing. This gives a virtual array of the same length as in the simulations. Comparing the dimensions of the measured and simulated PSFs (see the entry for 95 mm in Table 7.2) reveals a good agreement between simulations and experiment.

Figure 7.19 shows the reconstructed surface of the scanned wire phantom range gated between 65 and 68 mm. Both wires lie off-axis. The post focusing was done using $N_{pos} = 21$ planes for each new one. The image shows (the contours at the bottom of the plots) that the resolution in elevation and azimuth planes are of comparable size as predicted by the simulations. The wires are made from nylon and their surface is smooth. Due to specular reflection the echoes from the wire along the x axis weaken and this effect can be seen at the crossing point.

One of the problems of the approach are the grating lobes. They can be caused by a large step Δy. The step must not exceed the size of the PSF of the physical transducer in the elevation plane, otherwise not only grating lobes, but also discontinuities in the volume will appear. For a fixed number of emissions per second this limits the size of the volume that can be scanned.

7.5 Conclusion

One of the major limitation to the routine use of synthetic aperture ultrasound is caused by the low energy that can be transmitted by a single element of a linear array. The transmitted power can be increased by using multiple elements in transmit. In this chapter the different approaches for using multiple elements in transmit were put in the common frame of the virtual sources of ultrasound. It was shown that apart from the increased transmit energy, the concept of virtual ultrasound sources can be used to increase the resolution in the elevation plane of the 3D scans obtained using linear arrays.

CHAPTER

EIGHT

Sparse synthetic transmit aperture

The ultrasound beamforming is based on the coherent summation of the signals. The synthetic aperture imaging uses data acquired over several emissions. If the position of the scatterer changes, the echoes from it arrive at different time instances and cannot be summed in phase. In the worst case, the phase shift can be $\lambda/2$ and the signals cancel out.

One way to reduce the motion artifacts is to reduce the acquisition time by using only few of the transducer elements in transmit. This is equivalent to synthesizing sparse transmit aperture. The receive aperture can be either dense or sparse. The chapter starts by introducing the basics sparse arrays design. The introduction is limited to linear arrays. The performance in terms of resolution and grating and side-lobe levels for synthetic transmit aperture are discussed in the next section. The chapter ends with 2D sparse arrays design and their performance.

8.1 Fundamentals of sparse array design

Using sparse arrays as the one given in Figure 8.1 for imaging is a well studied topic and is described in a number of sources such as Section 3.3 of "*Array signal processing. Concepts and techniques.*" [22]. As seen from the previous sections, the resolution of a given system is

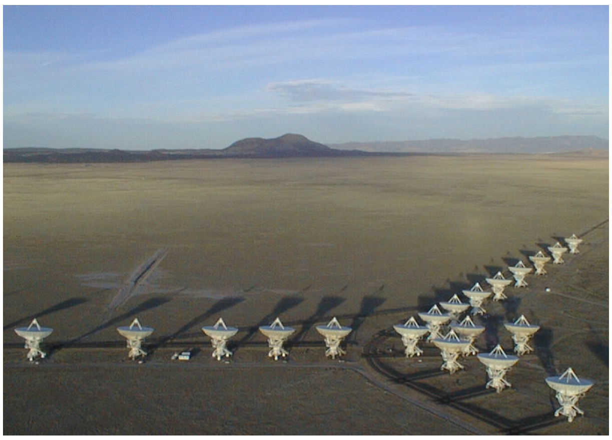


Figure 8.1: The Very Large Array radio telescope in New Mexico. It consists of 27 identical 25-m antennas arranged in a Y-shaped array. Two arms are 21 km long, and the third arm is 19 km long. The array belongs to National Radio Astronomy Observatory, and the picture is taken from `http://www.aoc.nrao.edu`.

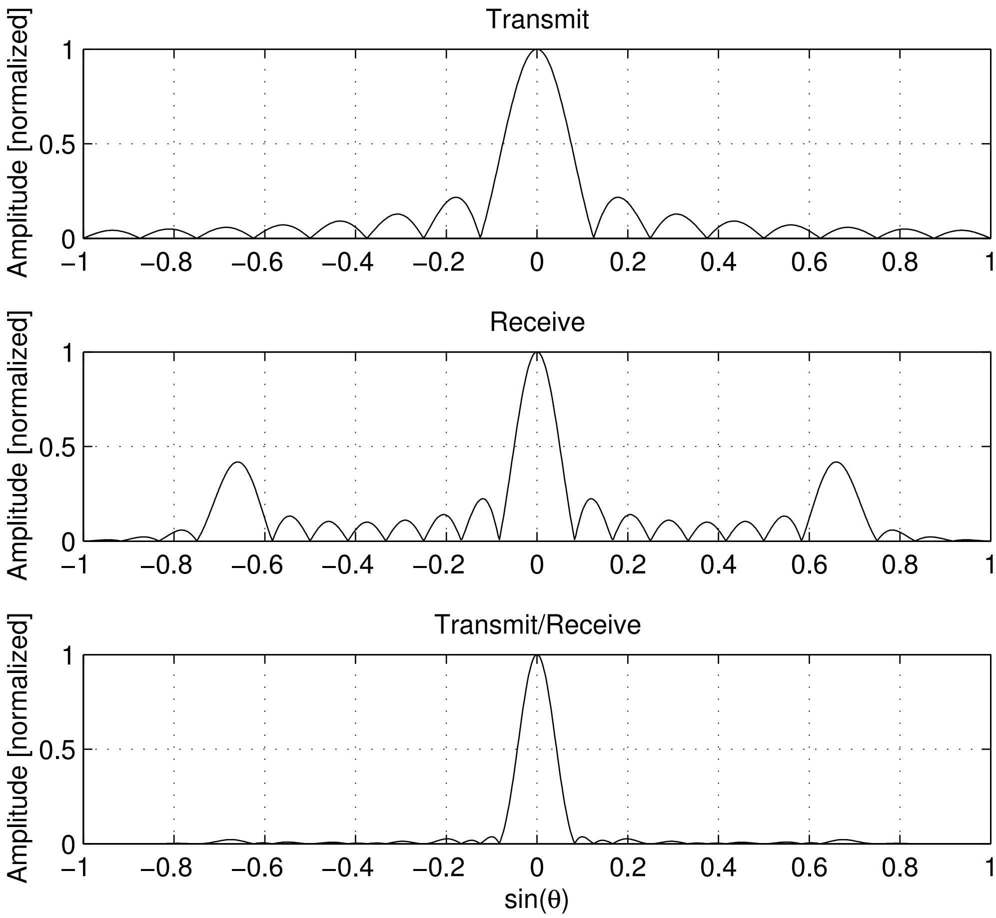


Figure 8.2: Reducing the level of the side and grating lobes. The distance between the elements in transmit is λ, and in receive is 1.5λ.

determined by the distance between the outer-most elements. If the number of transducer elements is reduced, the distance between two neighboring elements is increased and the grating lobes enter the field of view. The goal of the designer of sparse arrays is to position the sensors or to choose them from a fixed grid in such a way as to minimize the grating lobes. One way to reduce the grating lobes is by breaking the periodicity in their position. In ultrasound imaging, where the position of the elements is fixed, this can be done by picking randomly the active elements.

Ultrasound scanning is active imaging - energy is emitted first, and then received. The thing that matters is the two-way radiation pattern. This was realized by the group from Duke university in the design of the first real-time 3D scanner [6, 92]. The elements were chosen at a distance λ in transmit and 1.5λ in receive. This choice of elements ensures that the nulls in the radiation pattern in transmit cancel the grating lobes in receive and vice-verse as shown in Figure 8.2.

As it was shown in Section 3.7 the imaging abilities of a system depend on how the system fills in the k-space . Most of the sparse array designs are considered in the light of the "*co-array*" [52], or the "*effective aperture*" [50]. In the following we will use the term "*effective aperture*". The effective aperture $a_{t/r}(x)$ is defined as the spatial convolution between the transmitting $a_t(x)$ and receiving $a_r(x)$ aperture functions:

$$a_{t/r}(\vec{x}) = a_t(\vec{x}) \underset{x}{*} a_r(\vec{x}), \tag{8.1}$$

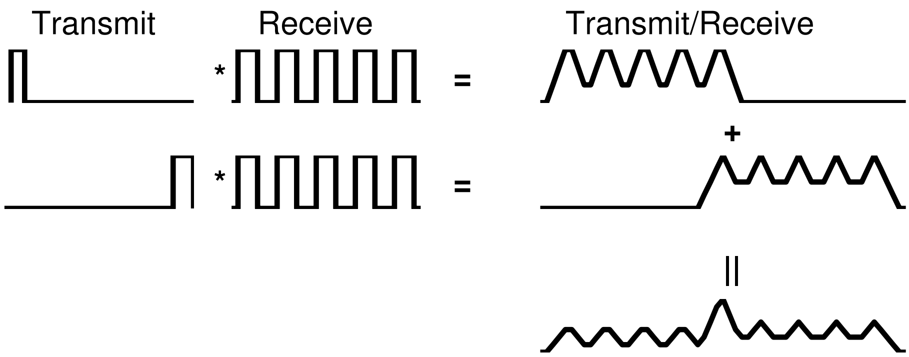


Figure 8.3: Synthetic transmit aperture imaging using only two emissions.

This is equivalent to the k-space representation of the system. Using the Fourier relations between an aperture and its radiation pattern, the point spread function of the system is the Fourier transform of the effective aperture.

Most considerations on sparse arrays were made for ”conventional” imaging, where the imaging is done line by line. The restriction is imposed by the complexity of the system, and the number of available channels. For the first time sparse synthetic transmit aperture was used by Cooley and Robinson [93]. They have taken the technique to the limit, by using only two emissions with the outermost elements. Because the imaging is considered a linear process, the effective aperture of the whole imaging process is the sum of the effective apertures at every emission. The result for two emissions is shown in Figure 8.3. Comparing this figure with Figure 4.5 on page 43 shows that using only two emissions and all elements in receive gives equivalent image as for the GSAU.

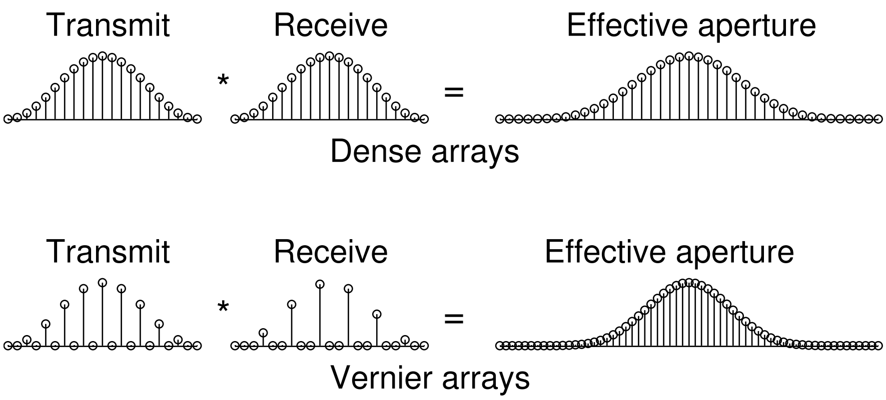


Figure 8.4: Design of Vernier arrays. The spacing between the elements in the transmit aperture is $2\frac{\lambda}{2}$ and in the receive one is $3\frac{\lambda}{2}$. The smoothness of the effective aperture is achieved by applying cosine^2 apodization. The effective aperture of the dense array is given for comparison.

Big impact had a series of articles by Lockwood and Foster [51, 50, 94], in which the use of the so-called *Vernier* arrays is suggested. The purpose of designing sparse arrays is to create an effective aperture with elements that are $\lambda/2$ spaced and has a “smooth” appearance. Such a design is shown in Figure 8.4. If the spacing in the transmit and receive apertures is $(p)\frac{\lambda}{2}$ and $(p+1)\frac{\lambda}{2}$, respectively, then the spacing between the elements in the effective aperture is $\lambda/2$.

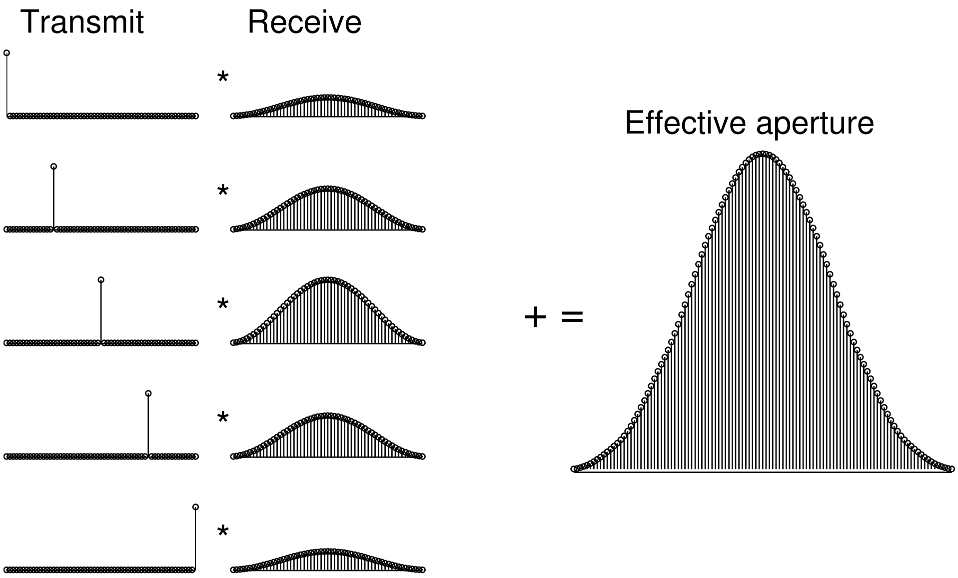


Figure 8.5: Transmit and receive aperture functions for a sparse synthetic aperture imaging system. The effective aperture is formed using 57 receive elements and 5 emissions.

The smooth appearance of the effective aperture is achieved by applying apodization, which for the case in Figure 8.4 is cosine^2. It must be noted, however, that in my experience this design fails when the arrays are "really sparse", the number of elements must be more than 20 for it to hold. This design was extended by Lockwood and colleagues to sparse synthetic transmit aperture imaging in [1]. The principle is illustrated in Figure 8.5, where the imaging is done using only 5 emissions and appropriate apodization in receive.

The conclusion is that in the case of more than two emissions, redundancy in the effective aperture is introduced allowing for the elimination of grating lobes.

The use of sparse synthetic transmit aperture imaging was also investigated by Chiao and Thomas in [3], and [95]. The design of the sparse arrays is considered in the light of the so-called transmit/receive matrix, which is just another tool to make the above considerations. Chiao and colleagues have also suggested the use of spatially encoded transmits to increase the emitted energy which will be discussed in Section 9.2. There exist, however methods that allow to overcome some of the effects of larger spacing in the effective aperture, and some of them are discussed in [33], and in [96].

8.2 Resolution and side lobe level

The performance of the method has been evaluated by scanning a wire phantom in a water bath using the XTRA system. The parameters chosen for the purpose are (1) the angular resolution δ_θ, and (b) the Integrated Main Lobe to Side Lobe Ratio (IMLSLR). The former characterizes the ability of the system to resolve bright scatterers, while the latter characterizes its ability to image dark regions, such as cysts. The resolution δ_θ is taken at several levels: -4 dB, -6 dB and -12 dB. The main lobe is defined as the energy contained in the signal samples having level

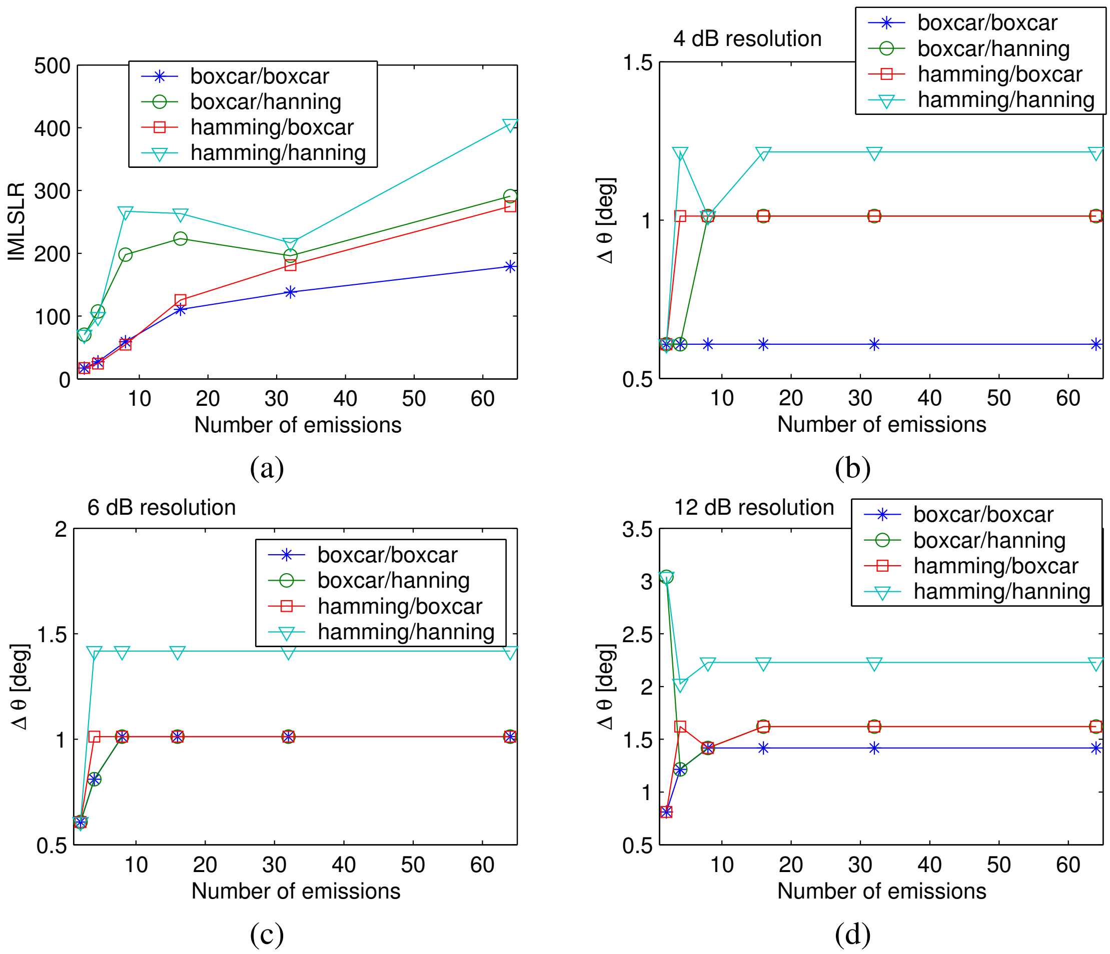


Figure 8.6: Some parameters of the point spread function of a sparse synthetic transmit aperture imaging system: (a) integrated side lobe to main lobe ratio (ISLMLR), (b) -4 dB angular width, (c) -6 dB angular width, and (d) -12 dB angular width. The PSF was obtained measuring a wire in a water bath using XTRA system. The main lobe was defined as the signal above -30 dB.

bigger than -30 dB:

$$ML(\theta, r) = \begin{cases} 1, & 20\log_{10}\left(\frac{\mathrm{psf}(\theta,r)}{\mathrm{psf}(0,0)}\right) > -30 \text{ dB} \\ 0 & \text{otherwise} \end{cases} \tag{8.2}$$

The IMLSLR is calculated as:

$$IMLSLR = \frac{\sum\limits_r \sum\limits_\theta |\mathrm{psf}(\theta, r)|^2 ML(\theta, r)}{\sum\limits_r \sum\limits_\theta |\mathrm{psf}(\theta, r)|^2 (1 - ML(\theta, r))} \tag{8.3}$$

The angular resolution is found from the projection of the maximum of the point-spread-function in axial direction:

$$\mathrm{A}_{t/r}(\theta) = \max_r \left(\mathrm{psf}(\theta, r)\right) \tag{8.4}$$

$$\delta_{\theta X \mathrm{dB}} = \max_\theta \left(\theta\big|_{(\mathrm{A}_{t/r}(\theta) > -X \text{ dB})}\right) - \min_\theta \left(\theta\big|_{(\mathrm{A}_{t/r}(\theta) > -X \text{ dB})}\right) \tag{8.5}$$

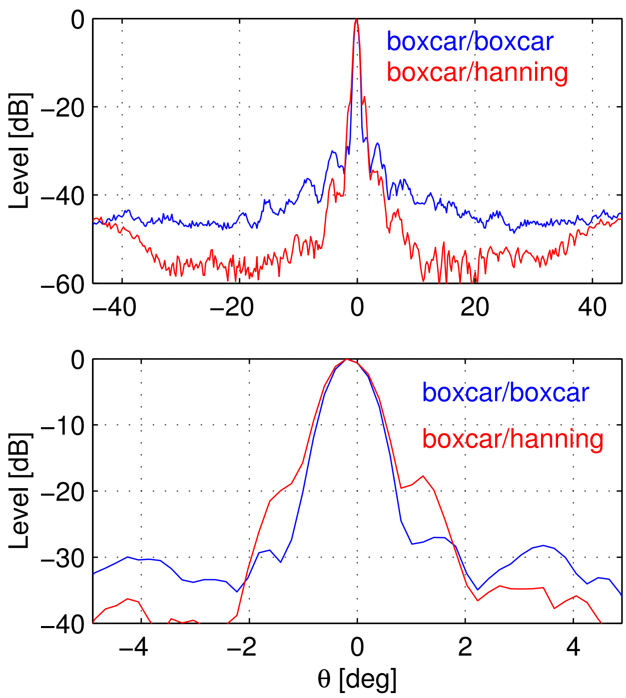


Figure 8.7: Comparison of the PSF for the apodization combinations boxcar/boxcar and boxcar/Hanning. The number of emissions is 4.

Experiments with 2, 4, 8, 16, 32 and 64 emissions were conducted and the results from them are plotted in Figure 8.6. Two types of apodization curves were tried in transmit[1] and in receive, giving a total of 4 combinations. In transmit the windowing functions were: boxcar and Hamming, and in receive: boxcar and Hanning. The reason for using Hanning window in receive is that it reaches 0 at the edges, thus eliminating to a great extent the edge effect. In transmit Hamming window was used because it does not reach zero at the outermost elements and can be applied with only a few emissions.

Consider IMLSLR shown in Figure 8.6(a). As a general rule, an increase in the number of emissions leads to reduction of the side lobe energy. Another observation is that a bigger influence has the apodization in receive, especially for a small number of emissions. Having Hanning apodization leads to a smaller level of the side lobes. The initial increase of IMLSLR for the case of a Hanning receive apodization is due to side lobes “entering” the main lobe as shown in Figure 8.7. The shape of the point spread function for the case with apodization is characterized also with a “clean” region around the main lobe.

The plots of the angular resolution comply with the general theory of windowing. In all of the depicted cases the highest resolution is achieved by the combination of apodization windows boxcar/boxcar, and the lowest with the combination Hamming/Hanning. The use of only 2 emissions has a resolution close to the one of GSAU, in which only a single element is used in transmit and receive. Considering Figure 8.6(d) an ”anomaly” can be observed: the 12 dB width of $A_{t/r}(\theta)$ is much wider for the case of 2 emissions and then decreases. This can be explained by the plots in Figure 8.8. The first two side lobes have a level above -12 dB and

[1] The transmit apodization is actually not applied in transmit. It is implemented as coefficients in the sum of the low resolution images.

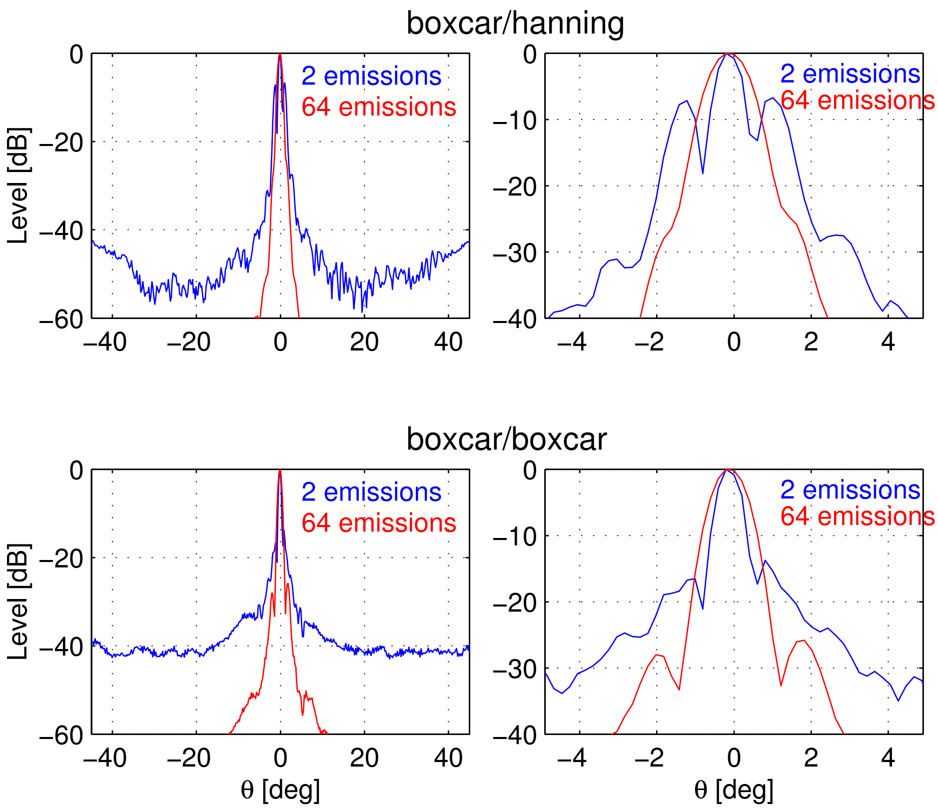


Figure 8.8: Comparison of the resolution between when imaging with 2 and 64 emissions. Two cases for a receive apodization are depicted: Hanning (top), and boxcar (bottom).

are considered as a part of the main lobe. This effect is absent for the case of boxcar receive apodization, in which most of the energy is dispersed at lower levels. The measured levels of the side and grating lobes are comparable to the theoretical finding of Hazard and Lockwood [75]. The number of bits used to quantize the values is essential, and if the signal does not fill in the input range of the A/D converter, the levels of the side lobes increase significantly.

8.3 Frame rate and deterioration of the point spread function

If sparse transmit aperture imaging is used, then the frame rate is determined by :

$$f_{fr\ \mathrm{STAU}} = \frac{f_{prf} N_{parallel}}{N_{xmt} N_l}, \tag{8.6}$$

where N_{xmt} is the number of the elements used in transmit, $N_{parallel}$ is the number of lines beamformed in parallel, and N_l is the number of lines in the image. Assuming these two parameters to be equal, the frame rate is dependent only on N_{xmt} and f_{prf}. The "sparser" the transmit aperture, the higher the frame rate is. This is for the case of synthetic transmit aperture imaging. The consequences for the recursive imaging is that less memory is necessary since only N_{xmt} low resolution images are involved in the process. The most important consequence for both types of imaging is the reduction of the effect of motion. The tolerance to motion was investigated experimentally by Hazard and Lockwood in [97].

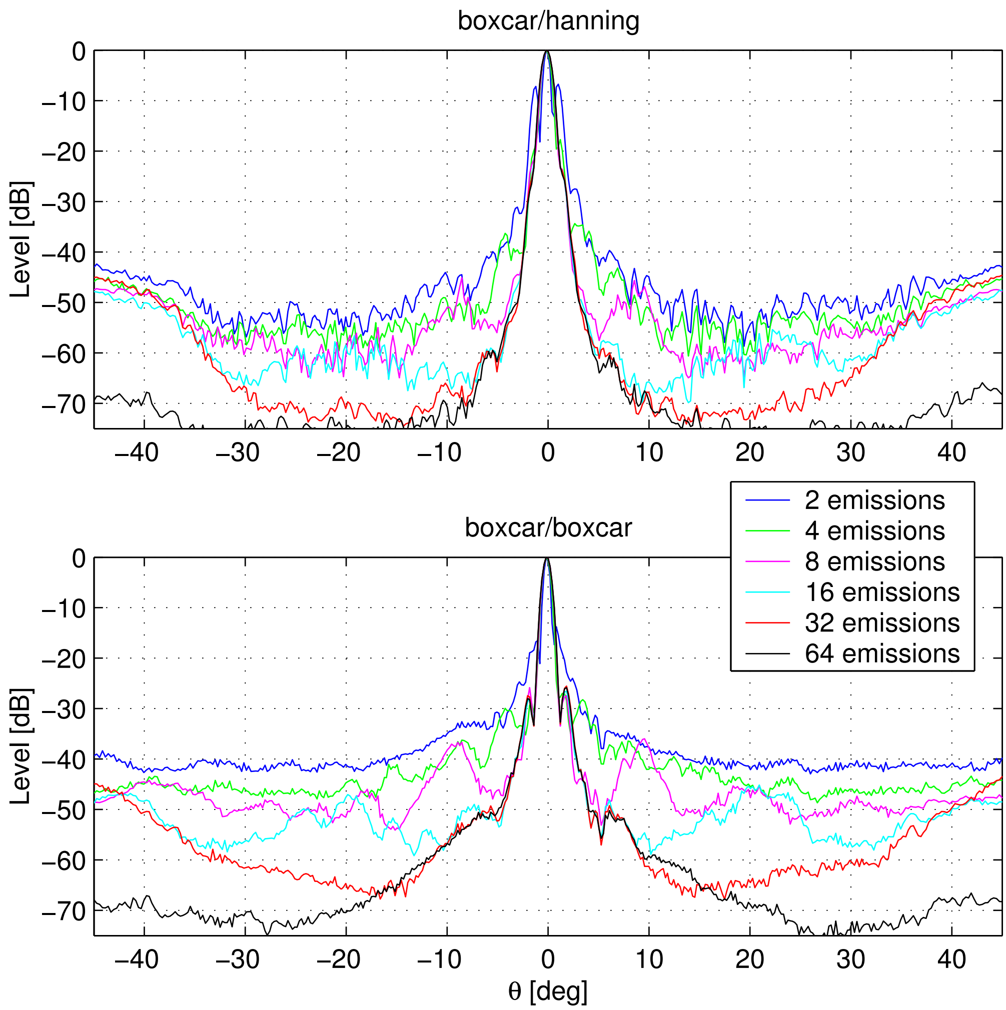


Figure 8.9: The point spread function of the XTRA system obtained with 2,4,8,16, 32, and 64 emissions. The transmit apodization is boxcar. The receive apodization is Hanning, top, and boxcar, bottom.

All their simulations and experiments were conducted for the case of a sparse transmit aperture with only 3 emissions, $N_{xmt} = 3$. Their findings are:

> The results of the simulations show that lateral motion does not significantly distort the radiation pattern for the velocities simulated ($\leq$ 2 m/s). There is no rise in the level of the secondary lobes and only a minor shift and widening of the main beam.
>
> While axial motion is a more serious concern, simulations reveal that at speeds typical for cardiac motion (0.08 m/s) there is little distortion of the main beam and virtually no increase in the level of the secondary lobes. At higher speeds (0.4 - 0.8 m/s), there is about a 4 dB increase in the level of the secondary lobes. Axial motion also produces fine structure in the main lobe and widens the beam slightly. These results were seen in both experimental data and simulation.

The simulations and the experiments conducted at the Center for Fast Ultrasound Imaging generally confirm these results. There is, however, one small detail which is not discussed in [97], the shape of the point spread function and its position. Most of their work was carried out on a point centered in front of the transducer, which is a special, and often favorable case. The motion artifacts are considered in greater detail in Chapter 11.

8.4 Extension to 2D arrays

The use of the principles of synthetic aperture imaging can be used to speed up the simulation of ultrasound images [43], especially when applied to 3D ultrasound imaging. The approach used in the cited work is instead of simulating the scan process line by line, to calculate the received signals $r_{ij}(t)$ for all permutations of transmit i and receive j channels. The signals are beamformed afterwards using the principles of synthetic transmit aperture imaging.

Assuming that the desired 2D aperture is a 64×64 elements matrix gives 4096 transmit events to acquire a single volume, provided that there are 4096 receive transducer elements connected 4096 channels in the system. Some day systems with 4096 channels will exist, but using 4096 transmit events per acquired volume will still result in 1 acquired volume per second. Thus, for 3D imaging sparse arrays are necessary both in transmit and receive. The 2D arrays have been investigated by several research groups for years. Their efforts were aimed at the choice of active elements and the apodization applied on them. In the following we will set aside the synthetic aperture, and concentrate the attention on the design of the 2D arrays.

As seen from Chapter 3 the increased distance between the centers of the elements causes the appearance of grating lobes in the region under investigation. The grating lobes are a result from the regular nature of the element position. Thus breaking the periodic structure of the array removes the grating lobes. The use of random 2D arrays have been investigated in a number of works [98, 99, 100]. The problem with the sparse random arrays is that a "pedestal" of side lobe energy is created (typically at -20, -30 dB), thus decreasing the contrast range in the images. This makes them unusable for imaging small, echo poor anatomical structures such as cysts. A number of works have been targeted towards the optimization of the position of the transducer elements [101, 102, 103]. The optimization can involve not only the element position, but also the apodization coefficients applied [104, 105, 106]. All of the optimization procedures presented in the cited papers *do* improve the performance of the arrays, but the results obtained are not adequate for ultrasound imaging. Another way of breaking the periodicity in the elevation and azimuth directions is to use arrays whose elements are placed as a function of angle, such as the spiral arrays [107, 108].

All of these designs were, however, not capable of surpassing the performance of regular arrays designed with the *two-way* radiation pattern in mind. A substantial amount of work has been carried out at Duke university by the group(s) led by Stephen Smith [11, 92, 109] and Olaf von Ramm [6, 110]. In the past decade this group was the only one able to use their transducers in a clinical environment.

Another significant contribution to the layout and apodization design of 2D arrays was the work by Geoffrey Lockwood [51, 94, 111]. He proposed the extension of the Vernier arrays to the two dimensional case. The layouts are considered to be independent in the azimuth and axial directions. The arrays showed very good performance. The Vernier relation is present only along the x and y axes. The design was further elaborated by the author [112]. The new design keeps the Vernier relation also along the diagonals of the array, thus reducing the grating lobes with 10 dB compared to the original suggestion. In the work presented at Ultrasonics International'99 Copenhagen two more layouts were suggested. These are based on a few dense clusters of elements in the receive aperture and a regular sparse transmit aperture. Approximately at the same time Austeng and Holm suggested the introduction of redundancy in the apertures [113, 114]. In their considerations the apertures are treated as spatial filters. They use the principle of cascade filters. The idea is to find a simple convolution kernel which convolved with the apodization function of the transmit aperture introduces zeros at the position of the grating lobes of the receive aperture. This design is applied onto the Lockwood's Vernier arrays and in [114] it is called *enahanced Vernier arrays*. In the next sections different approaches of designing sparse 2D arrays based on the Vernier design will be presented, and their performance will be compared based on the simulated point spread functions.

8.4.1 Design of 2-D arrays.

The first approach to design sparse 2D arrays is to assume that the 2D apodization function is a separable function. It can be separable either in Cartesian (x,y) or in polar (r,θ) coordinate systems. This assumption gives:

$$a_t(x,y) = a_t(x)\cdot a_t(y) \tag{8.7}$$
$$a_r(x,y) = a_r(x)\cdot a_r(y), \tag{8.8}$$

where a_t and a_r are the transmit and receive apodization functions. They include also the extent of the aperture.

For the radially separable functions one gets:

$$a_t(r,\theta) = a_t(r)\cdot a_t(\theta) \tag{8.9}$$
$$a_r(r,\theta) = a_r(r)\cdot a_r(\theta). \tag{8.10}$$

Usually the apertures that are separable in polar coordinates are circularly symmetric, and therefore the transmit and receive apertures are defined only as a function of the radial distance [2].

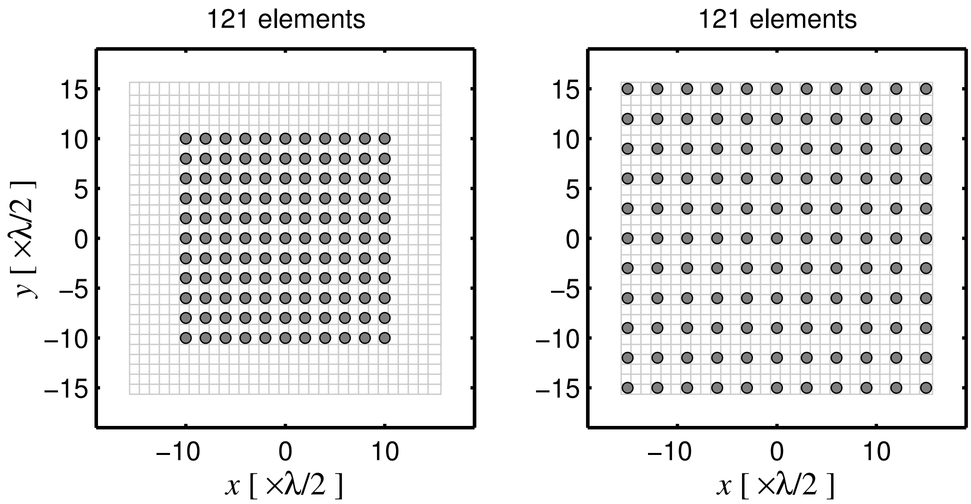
The Vernier arrays are defined as :

$$\begin{aligned} a_t(x) &= \begin{cases} 1 & x = p\cdot d_x, \text{ and } |x| \leq \frac{N_{xdc}}{2} d_x \\ 0 & \text{otherwise} \end{cases} \\ a_r(x) &= \begin{cases} 1 & x = (p-1)\cdot d_x, \text{ and } |x| \leq \frac{N_{xdc}}{2} d_x \\ 0 & \text{otherwise} \end{cases} \end{aligned} \tag{8.11}$$

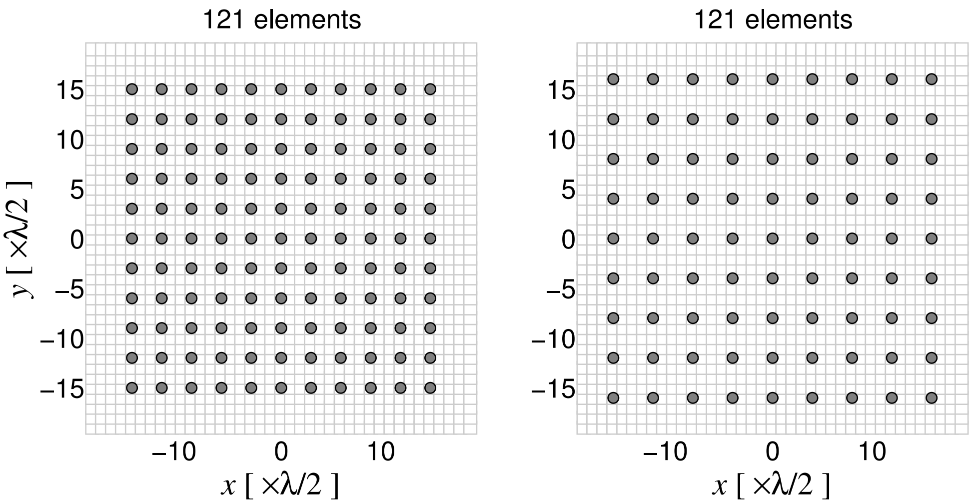
where p is an integer number, d_x is the pitch of the transducer and N_{xdc} is the total number of transducer elements.

Figures 8.10 and 8.11 show sparse arrays whose apodization functions are separable in Cartesian and polar coordinates, respectively. Sub-figures (a), (b), and (c) show the cases for $p = 2$, $p = 3$, and $p = 4$, respectively. All of the sparse layouts presented in this chapter are designed with the restriction that the sum of the active elements in transmit and receive should be between 240 and 270 elements. For the same number of active elements, the arrays separable in Cartesian coordinates cover a larger area, and are expected to have higher resolution than the circularly symmetric ones. Equation (8.11) gives only the position of the active elements (as

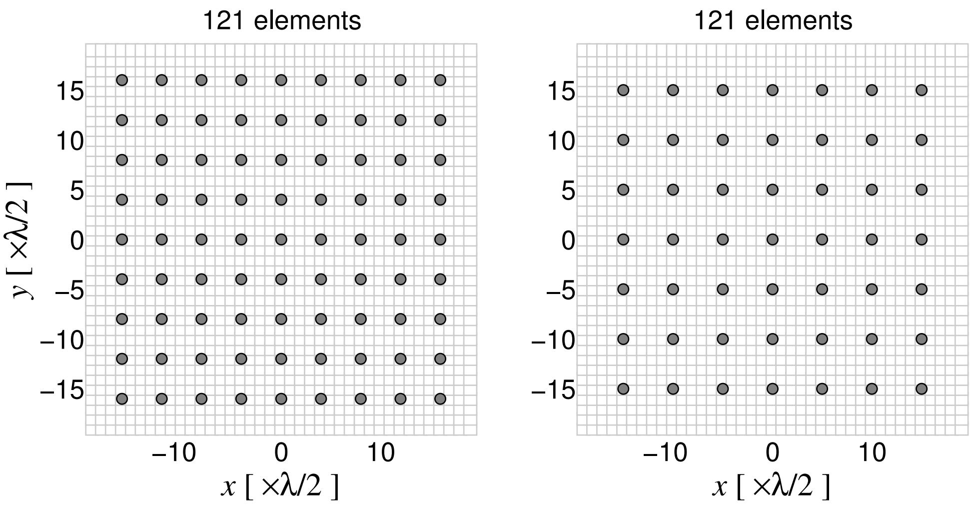
[2] Remember that the origin of the coordinate system lies in the geometrical center of the aperture.



(a) vern_2

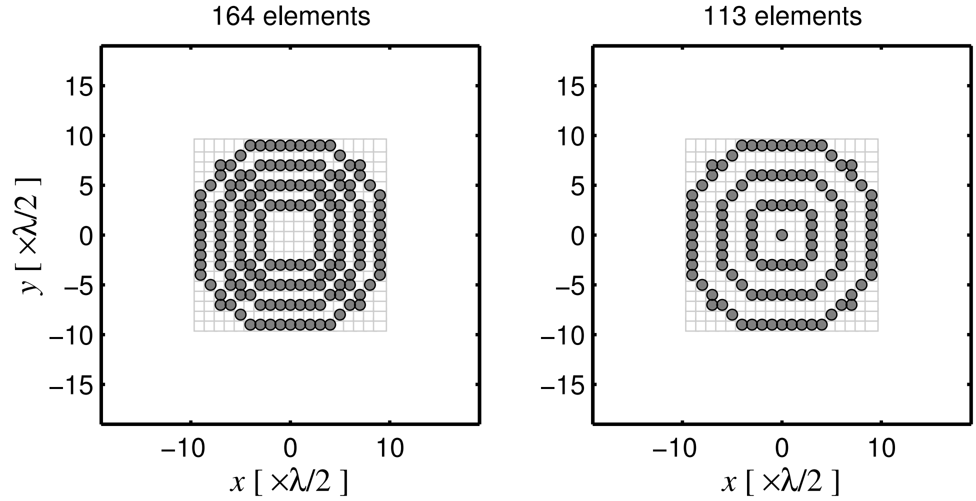


(b) vern_3

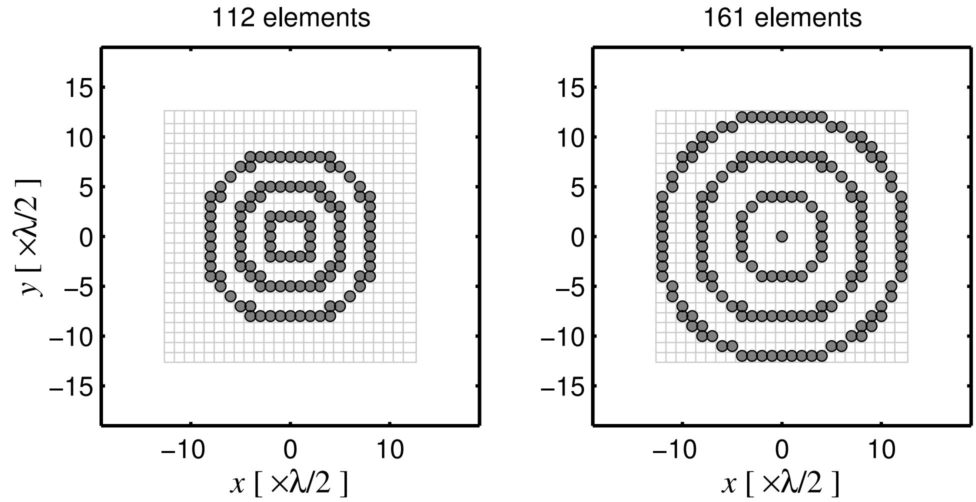


(c) vern_4

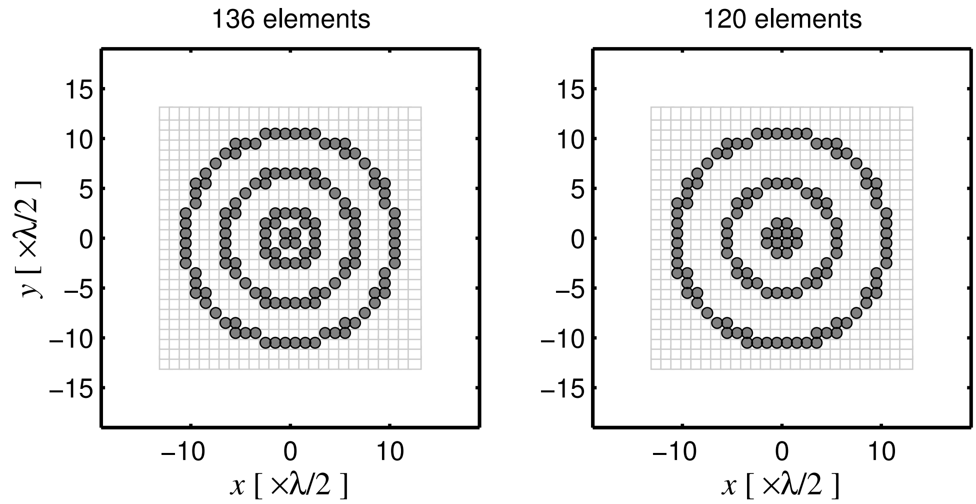
Figure 8.10: Classical Vernier arrays as suggested by Lockwood and Foster [50].



(a) vern_2_rad

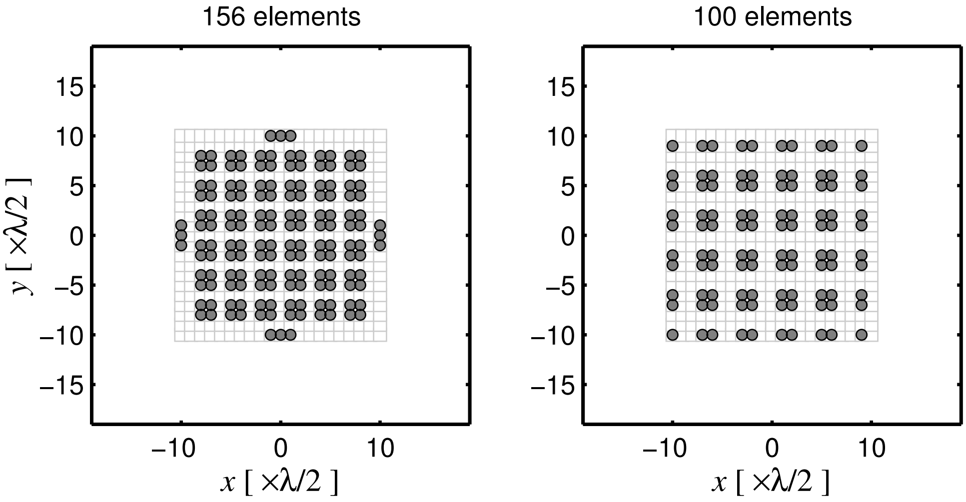


(b) vern_3_rad

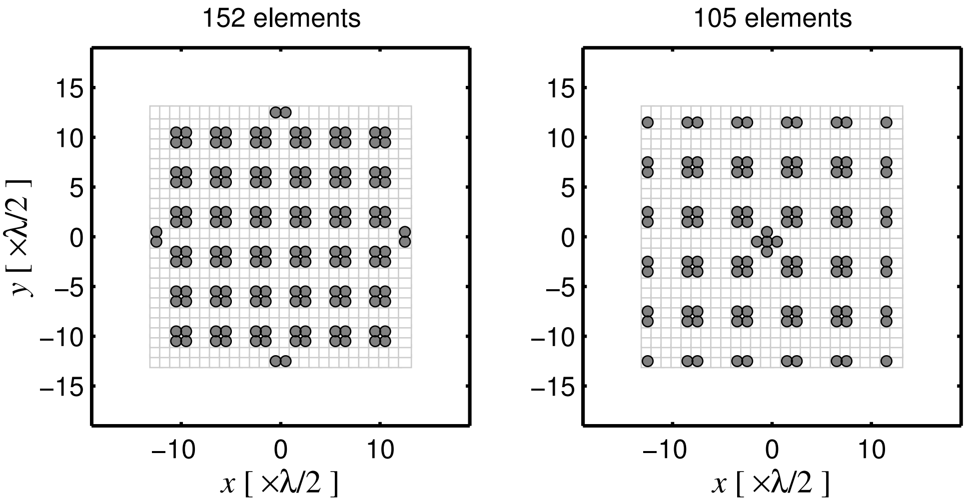


(c) vern_4_rad

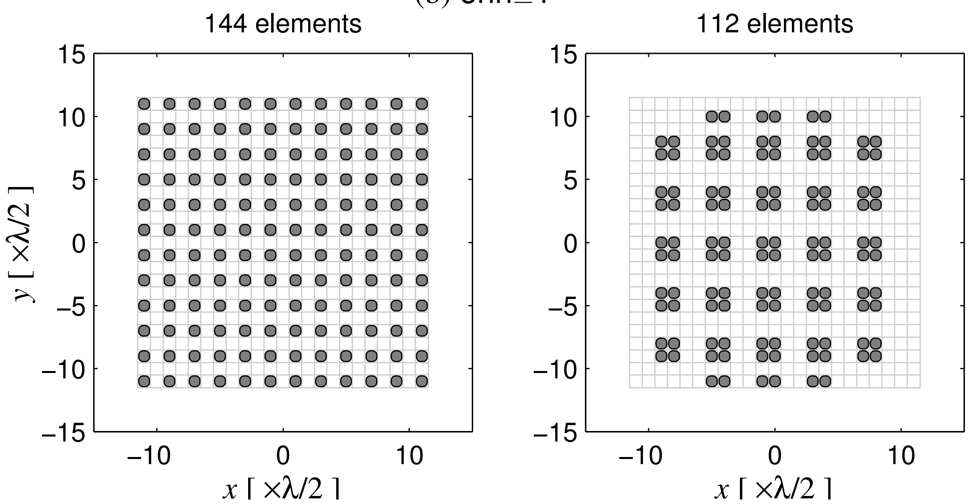
Figure 8.11: Circularly symmetric Vernier arrays.



(a) enh_3

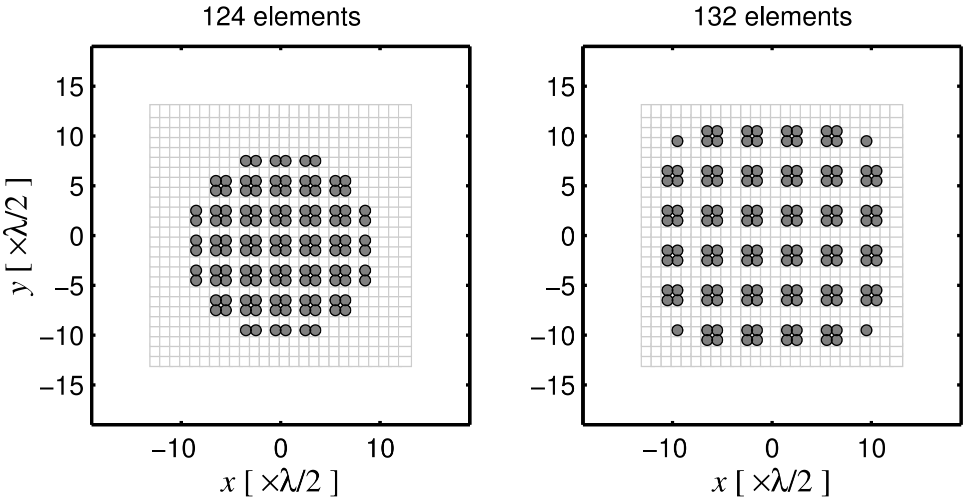


(b) enh_4

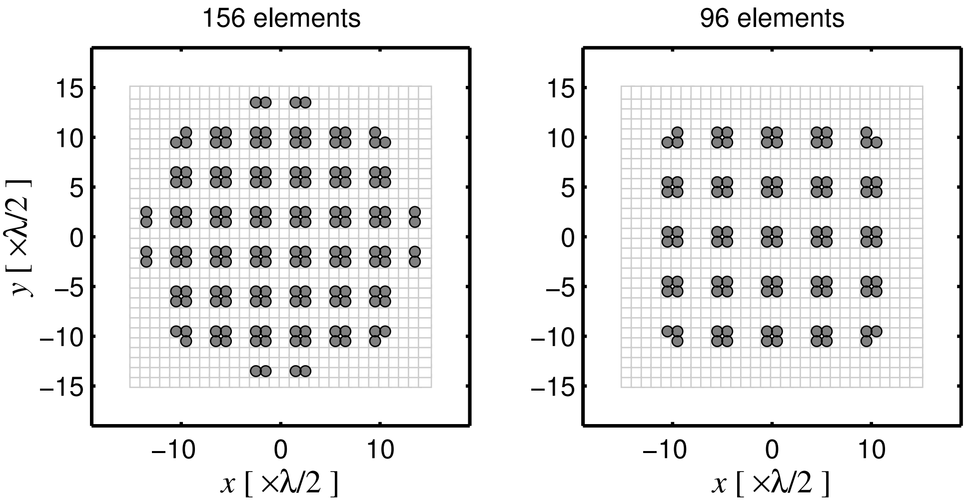


(c) enh_vern

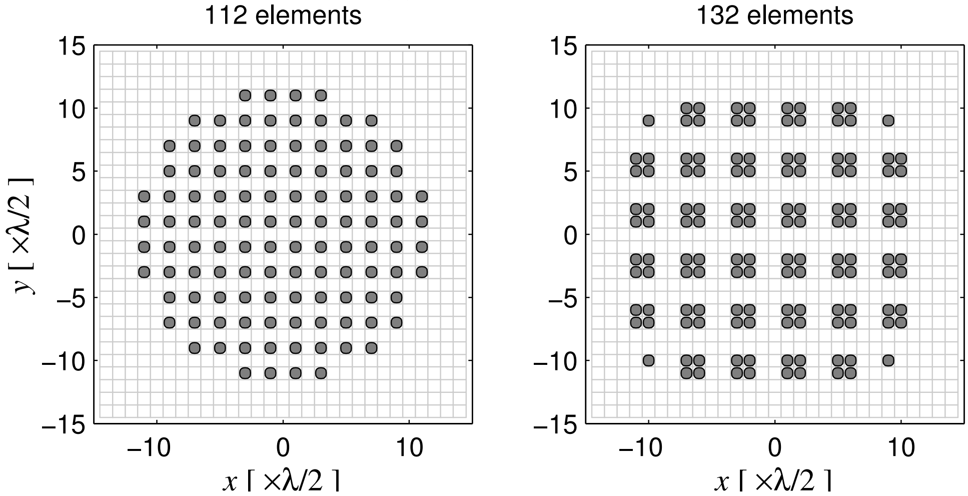
Figure 8.12: Enhanced Vernier arrays as suggested by Austeng and Holm [114, 113].



(a) enh_3_circ



(b) enh_4_circ



(c) enh_vern_circ

Figure 8.13: Enhanced Vernier arrays with circular footprint.

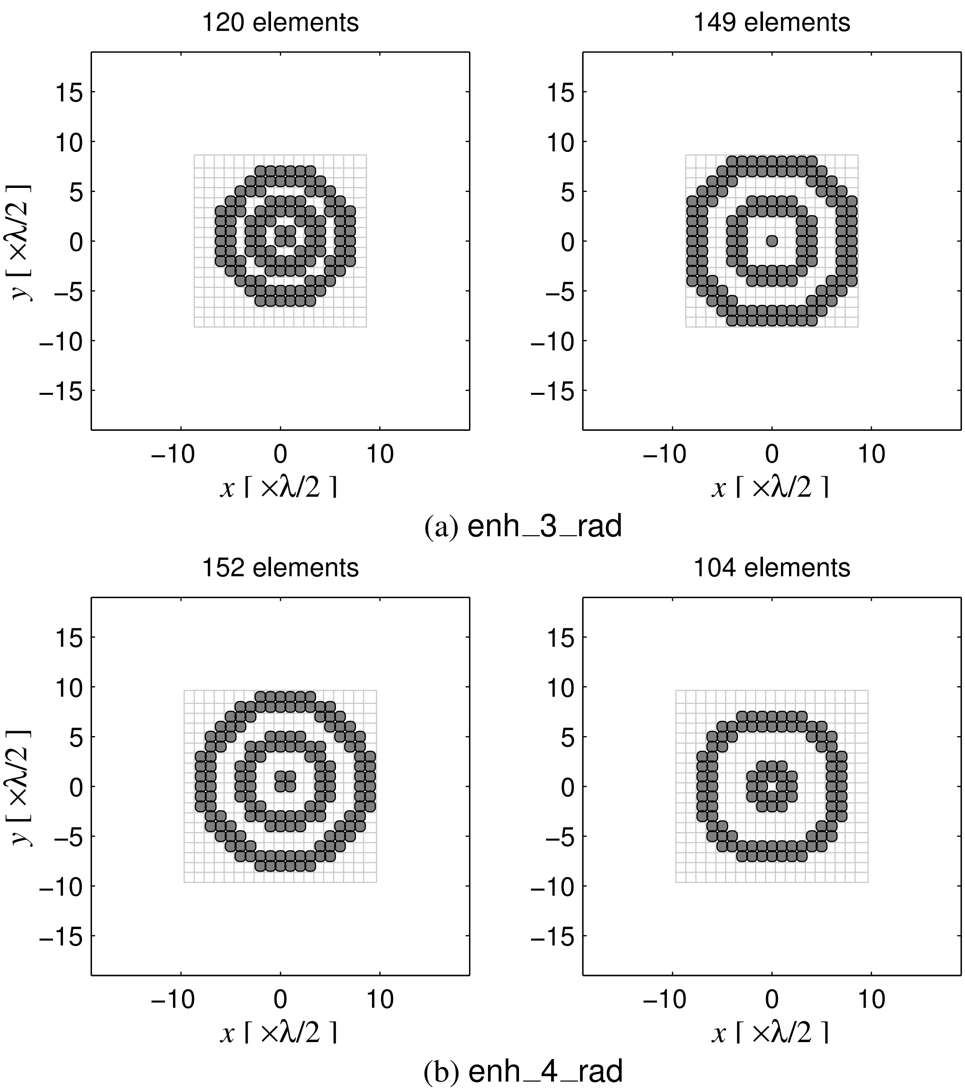


Figure 8.14: Circularly symmetric enhanced Vernier arrays.

a boolean matrix). Additional weighting is applied on the elements to smoothen the effective aperture and further reduce the grating lobes. While Lockwood and colleagues suggests the use of apodization functions such as $\cos^2 x$ [51], Austeng and Holm suggest to weight only the outermost elements [113].

The choice of active elements for circularly symmetric apertures like the ones in Figure 8.11 depends on the way a given element is treated as a being part of the circle with radius r. Different 2D layouts may be created from the same 1D function.

It must be noted that while the Vernier arrays reduce the grating lobes, they do not eliminate them. Austeng and Holm [113, 114] suggest to introduce redundancy in the array to further decrease the grating lobes. To do so one must have one or more zeros in the transmit radiation pattern located at the angles of the receive grating lobes and vice versa. One way of generating zeros at desired angles is by convolving the sparse array with short kernel. The kernel used in [114] is simply $k = [1, 1]$. The resulting arrays are named "*enhanced*" Vernier arrays. Figures 8.12 and 8.13 show enhanced vernier arrays separable in Cartesian coordinates with a rectangular and circular footprints respectively. Figure 8.14 shows circularly symmetric enhanced

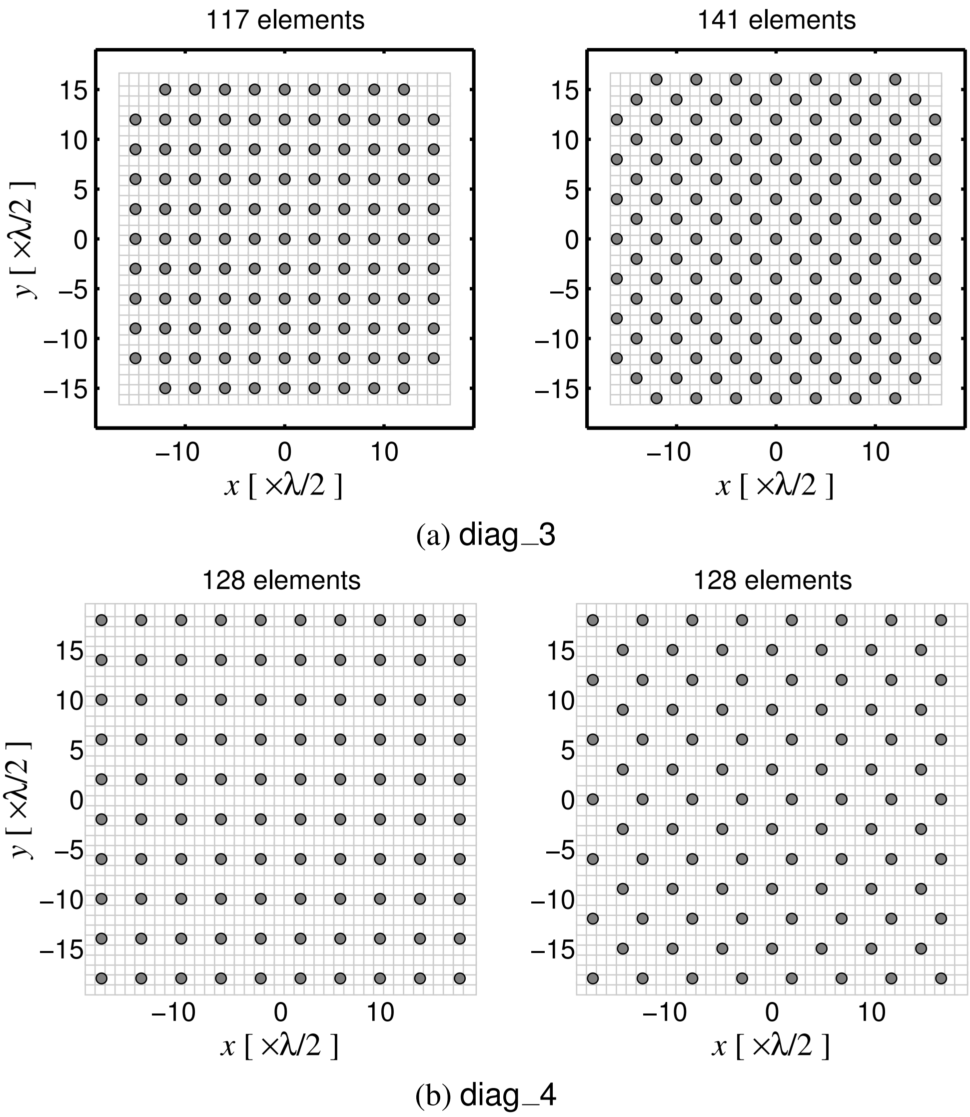


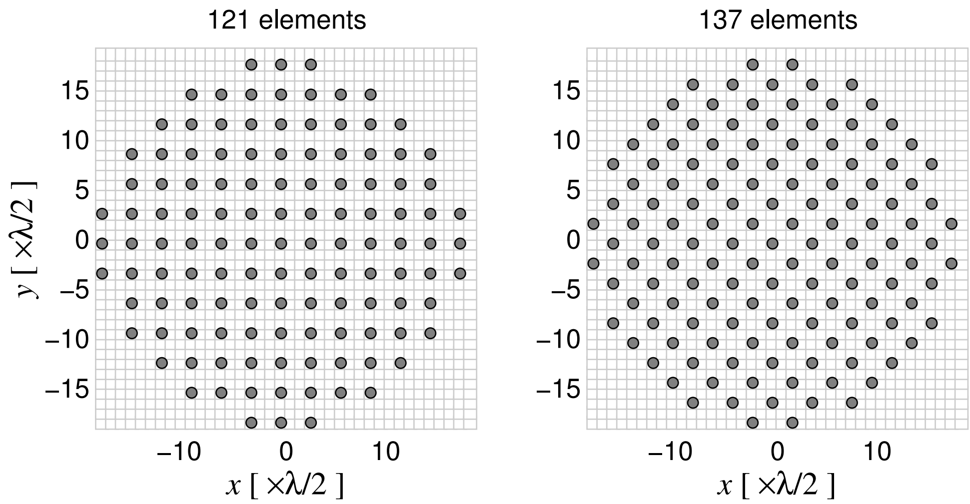
Figure 8.15: Diagonally optimized Vernier arrays as suggested by the author [112].

Vernier arrays.

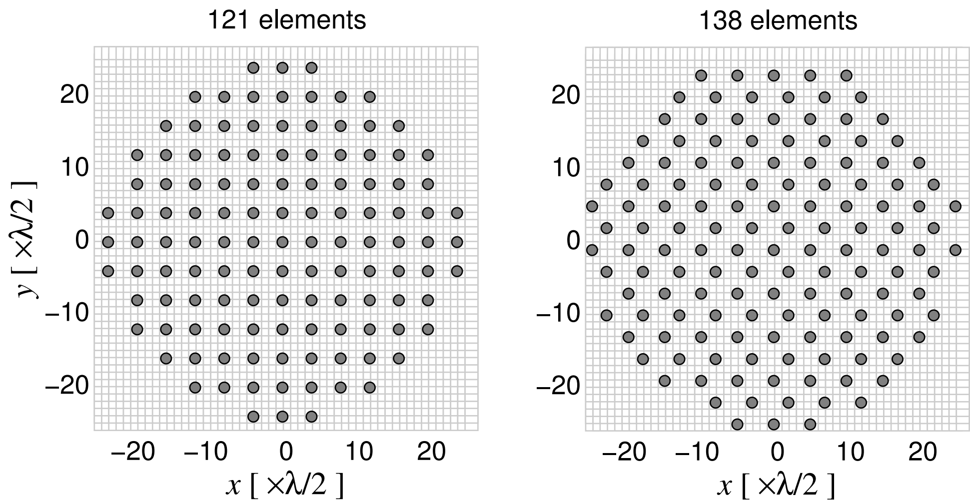
One of the problems of the Vernier designs is that along the diagonal the distances between the elements is too large. Consider Figure 8.10. Let the pitch in azimuth and elevation directions be the same, and equal to $\lambda/2$. Along the axes x and y the distance between the elements is $p\lambda/2$ and $(p-1)\lambda/2$. Along the diagonals, however, the distances are $p\sqrt{2}\lambda/2$ and $(p-1)\sqrt{2}\lambda/2$. The suggestion in [112] is to position the elements of one of the apertures along the diagonals as shown in Figure 8.15. The distance along the axes for the transmit and receive apertures is $p\lambda/2$ and $(p-1)\lambda/2$, respectively. Along the diagonals these values are $(p-1)\sqrt{2}\lambda/2$ and $(p-2)\sqrt{2}\lambda/2$, respectively. The reduced distance along the diagonals results in reduced grating lobes.

This type of design will be called "*diagonally optimized Vernier arrays*". Figures 8.15 and 8.16 show diagonally optimized Vernier arrays with rectangular and circular footprints respectively.

In [112] yet another type of arrays were suggested - clustered arrays. Such arrays are shown in Figure 8.17.



(a) diag_3_circ



(b) diag_4_circ

Figure 8.16: Diagonally optimized Vernier arrays with circular footprint.

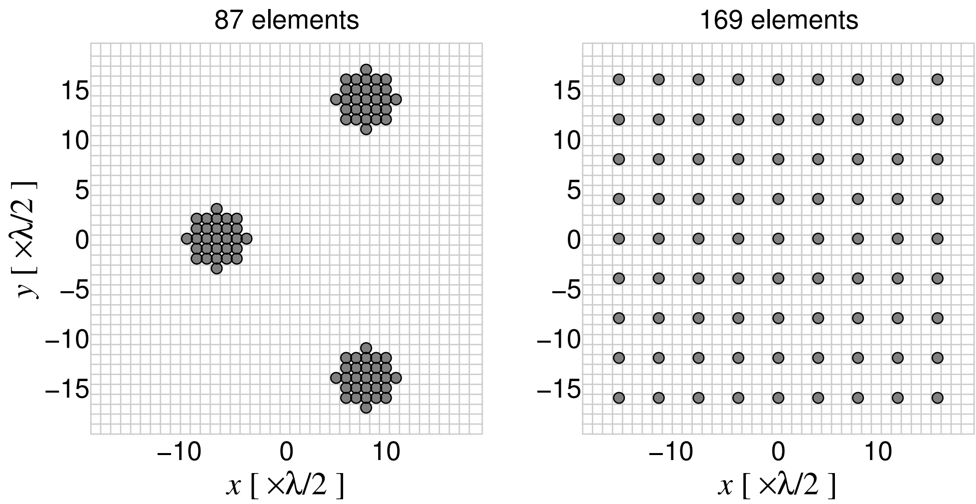
One of the apertures consists of several clusters of elements. The number of elements in a cluster and the number of clusters depends on the constraints imposed by the total number of active elements allowed. The clusters are identical in shape, number of elements and apodization coefficients. The clusters are circularly symmetric. Let the radial size of a cluster be D_{cluster} and the number of elements in azimuth and elevation directions be:

$$N_{x\ \text{cluster}} = N_{y\ \text{cluster}} = \left\lfloor \frac{D_{\text{cluster}}}{d_x} \right\rfloor, \tag{8.12}$$

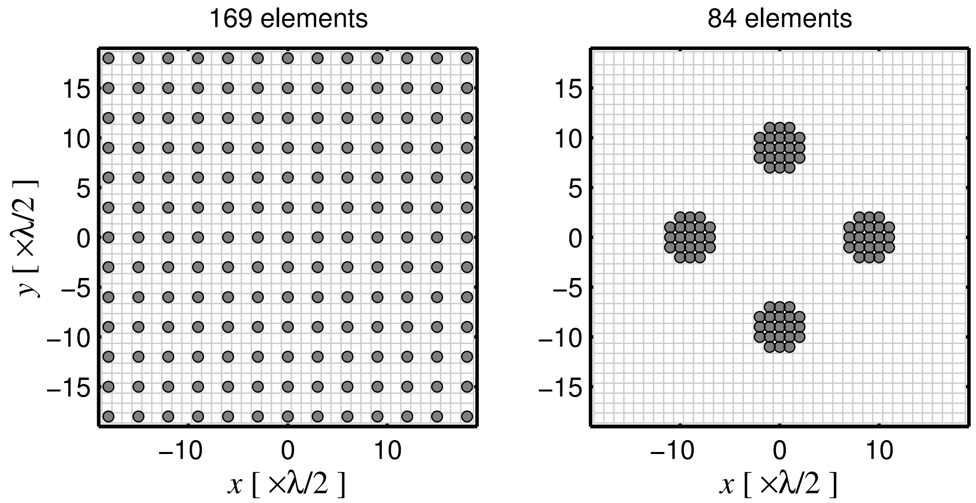
where d_x is the pitch of the transducer. Let the receive aperture be the one consisting of clusters. The transmit aperture is, in this case, a sparse array with regular spacing between the elements. The number of skipped elements between two active elements must be:

$$N_{x,\text{skip}} = \left\lfloor \frac{N_{x\ \text{cluster}} - 1}{2} \right\rfloor \tag{8.13}$$

If the area covered by the elements of the transmit aperture is Nd_x, then the number of elements



(a) clust_3



(b) clust_4

Figure 8.17: Clustered arrays [112].

between the centers of two clusters is:

$$N_{\text{clusters skip}} = \left\lfloor \frac{N-2}{2} \right\rfloor. \tag{8.14}$$

The results of these designs is given in Figure 8.17. The designs in Figures 8.17(a) and 8.17(b) consist of 3 and 4 clusters, respectively. The size of a single cluster is determined by the spacing between the elements of the sparse aperture. The number of clusters is limited by the total number of elements allowed.

8.4.2 Performance of the sparse arrays

The performance of the arrays shown in Figures 8.10 – 8.17 is compared based on the simulated point spread functions of the arrays. All of the arrays are designed by limiting the total number of active elements:

$$240 \leq N_{xmt} + N_{rcv} \leq 270, \tag{8.15}$$

Design	Figure	N_{xmt}	N_{rcv}	N_{over}	N_{tot}	$L_{x\ xmt}$	$L_{y\ xmt}$	$L_{x\ rcv}$	$L_{y\ rcv}$
vern_2	8.10(a)	121	121	9	242	21	21	31	31
vern_3	8.10(b)	121	121	9	242	31	31	41	41
vern_4	8.10(c)	121	121	9	242	41	41	51	51
vern_2_rad	8.11(a)	164	113	76	277	19	19	19	19
vern_3_rad	8.11(b)	112	161	56	273	17	17	25	25
vern_4_rad	8.11(c)	136	120	72	256	22	22	22	22
enh_3	8.12(a)	156	100	27	256	21	21	20	20
enh_4	8.12(b)	152	105	16	257	26	26	25	25
enh_vern	8.12(c)	144	112	28	256	23	23	18	22
enh_3_circ	8.13(a)	124	132	32	256	18	18	22	22
enh_4_circ	8.13(b)	156	96	32	252	28	28	22	22
enh_vern_circ	8.13(c)	112	132	28	244	23	23	22	22
enh_3_rad	8.14(a)	120	149	45	269	14	14	17	17
enh_4_rad	8.14(b)	152	104	40	256	18	18	15	15
diag_3	8.15(a)	117	141	13	258	31	31	33	33
diag_4	8.15(b)	128	128	8	256	45	45	41	43
diag_3_circ	8.16(a)	121	137	13	258	37	37	37	37
diag_4_circ	8.16(b)	121	138	8	259	49	49	51	49
clust_3	8.17(a)	87	169	6	256	22	35	49	49
clust_4	8.17(b)	169	84	4	253	37	37	23	23

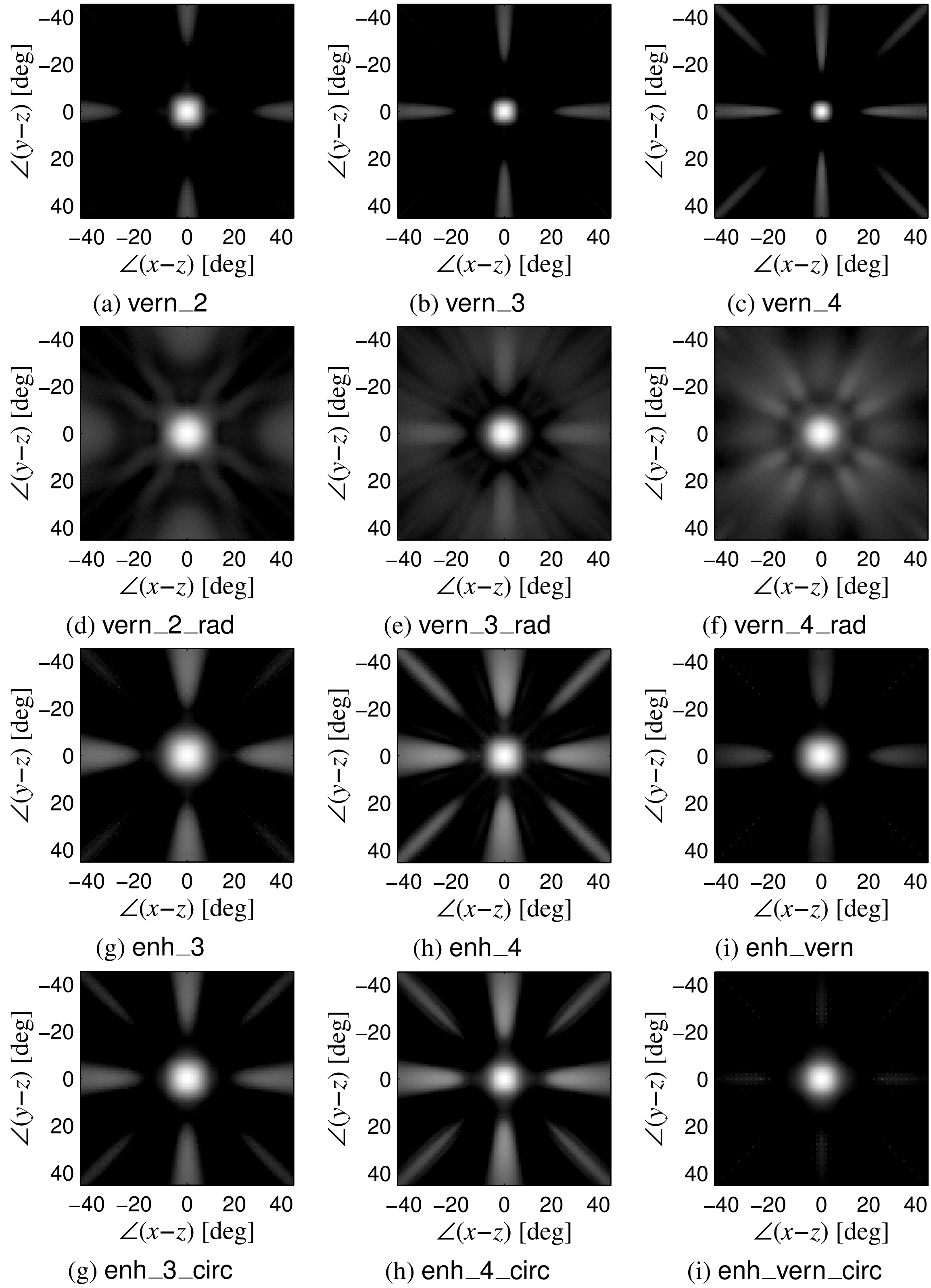
Table 8.1: Properties of the designs. N_{xmt}, N_{rcv}, N_{tot}, and N_{over} are the number of transmit, receive, total and overlapping elements, respectively. $L_{x\ xmt}$, $L_{x\ rcv}$, $L_{y\ xmt}$, and $L_{y\ rcv}$ are the distances between the outermost elements in x and y directions for the transmit and receive apertures given in number of elements.

where N_{xmt} and N_{rcv} is the number of transmit and receive elements, respectively. More precisely the targeted sum is 256, with small deviations (± 15 elements) allowed in order not to handicap some of the layouts.

Table 8.1 summarizes the parameters of the designs. Each of the designs is assigned a name, with vern, enh, diag and clust standing for “Vernier”, “Enhanced Vernier”, “Diagonally Optimized Vernier” and “Clustered”, respectively. The numbers show the value of p, and circ and rad mean ”with circular footprint” and ”radially symmetric” respectively. The table also gives references to the figures, where the respective designs are shown. The number of transmit, receive, and the number of total and overlapping elements are also given. The radially symmetric designs have the highest total number of transmit and receive elements. The design vern_2_rad exceeds the limit of total number of elements, but if reduced its performance would be “artificially” worsened.

The distance between the outermost elements along the x and y axes are denoted with L_x and L_y. They are given in number of elements. The letters rcv and xmt, that are appended to the subscripts, show whether the number is for the receiving or the transmitting apertures, respectively.

B-mode images of the pulse echo response of the various designs are given in Figures 8.18 and 8.19. Below each of the images the abbreviation of the design used to acquire it is given. The images were simulated using the program Field II. The relevant parameters of the simulation



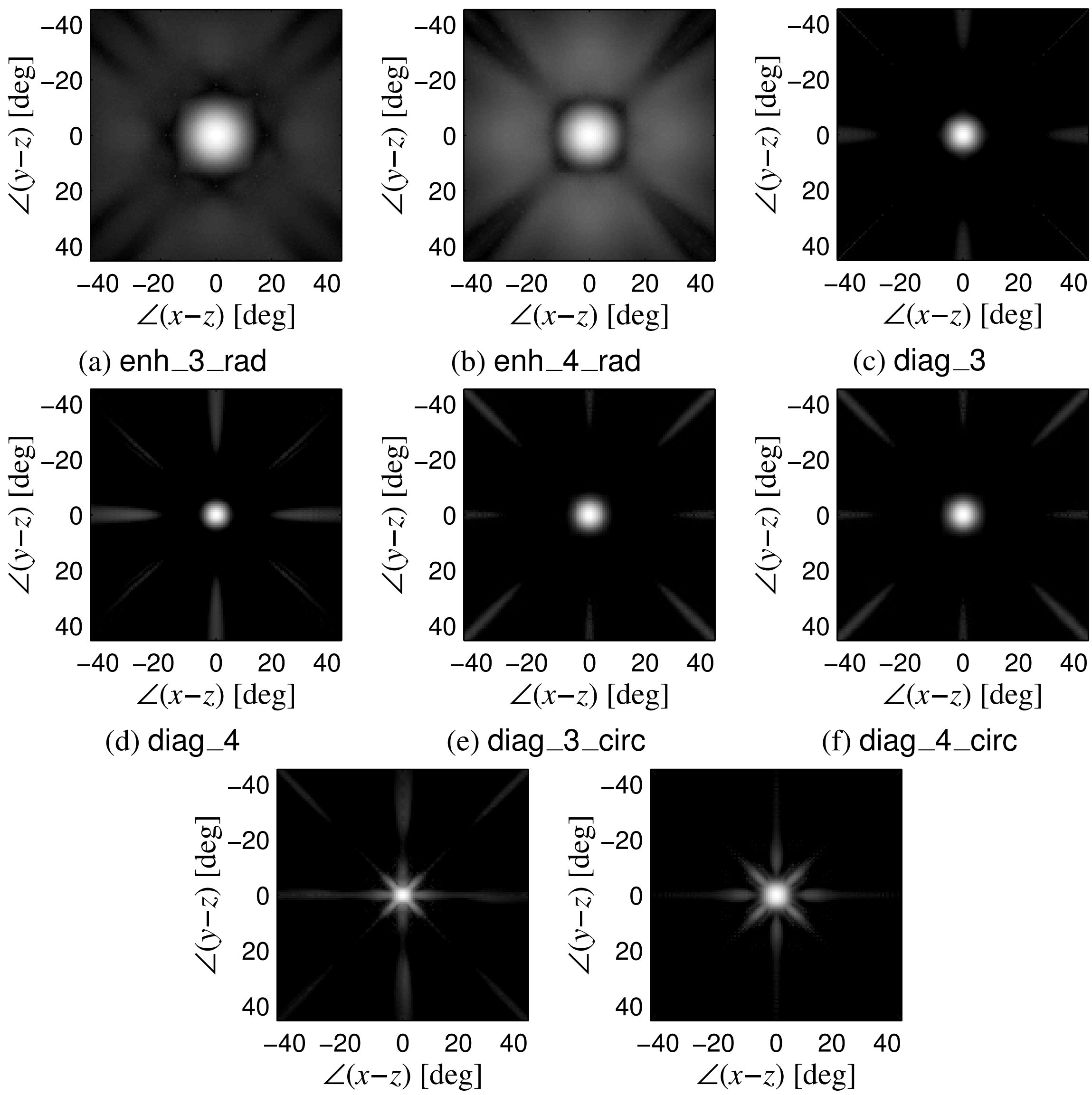
(a) vern_2 (b) vern_3 (c) vern_4

(d) vern_2_rad (e) vern_3_rad (f) vern_4_rad

(g) enh_3 (h) enh_4 (i) enh_vern

(g) enh_3_circ (h) enh_4_circ (i) enh_vern_circ

Figure 8.18: B-mode images of the point spread function of the array pairs shown in Figures 8.10, 8.11, 8.12, and 8.13. The dynamic range of images is 60 dB.



(a) enh_3_rad (b) enh_4_rad (c) diag_3

(d) diag_4 (e) diag_3_circ (f) diag_4_circ

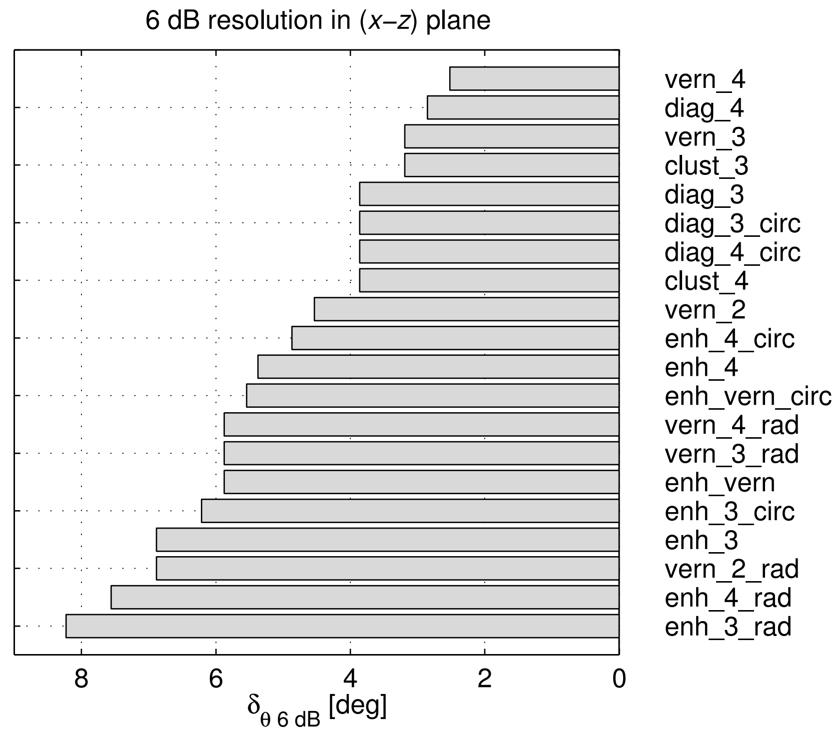
(d) clust_3 (e) clust_4

Figure 8.19: B-mode images of the point spread function of the array pairs shown in Figures 8.14, 8.15, 8.16, and 8.17. The dynamic range of images is 60 dB.

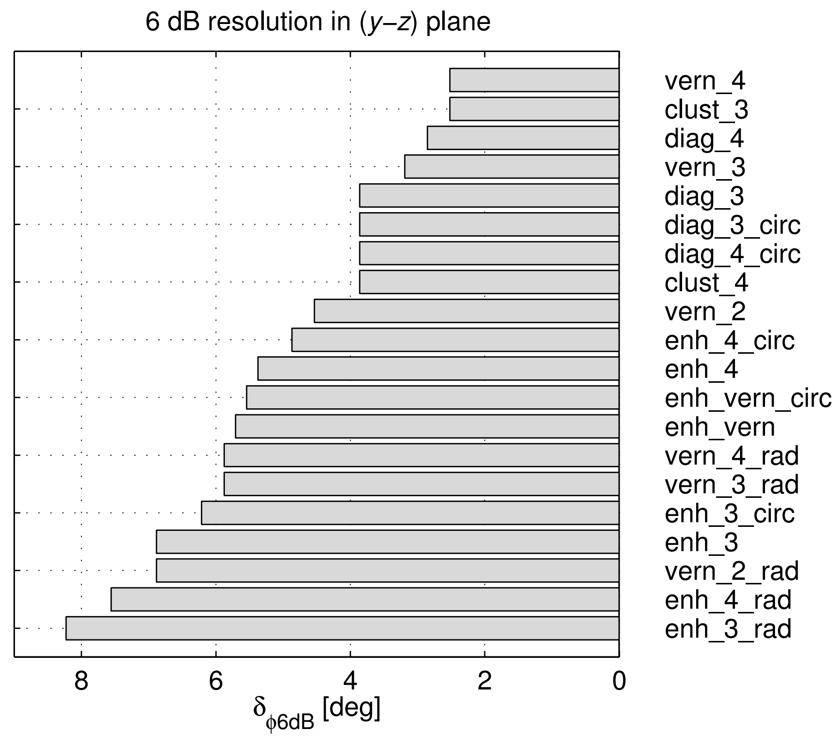
are given in Table 8.2.

The scanned point scatterer is centered 50 mm away from the transducer surface. The angles in the elevation and azimuth directions range from -45 to 45 degrees. The number of scan lines is 135×135, or in other words 1.5 lines per degree. The images shown in Figures 8.18 and 8.19 are obtained by taking the maximum of the response along the lines at each angle. The data was logarithmically compressed prior to display. Levels below -60 dB are cut from the plots.

The performance of the arrays is assessed on the basis of the resolution, maximum grating lobe level, and integrated main lobe to side lobe ratio, whose values are shown in Table 8.3. Some of the designs such as clust_3 are asymmetric and the parameters are therefore given for the $(x-z)$ and $(y-z)$ planes separately. This is done by projecting the maximum of the point spread function onto the planes, and picking the values for the table from the projected values.

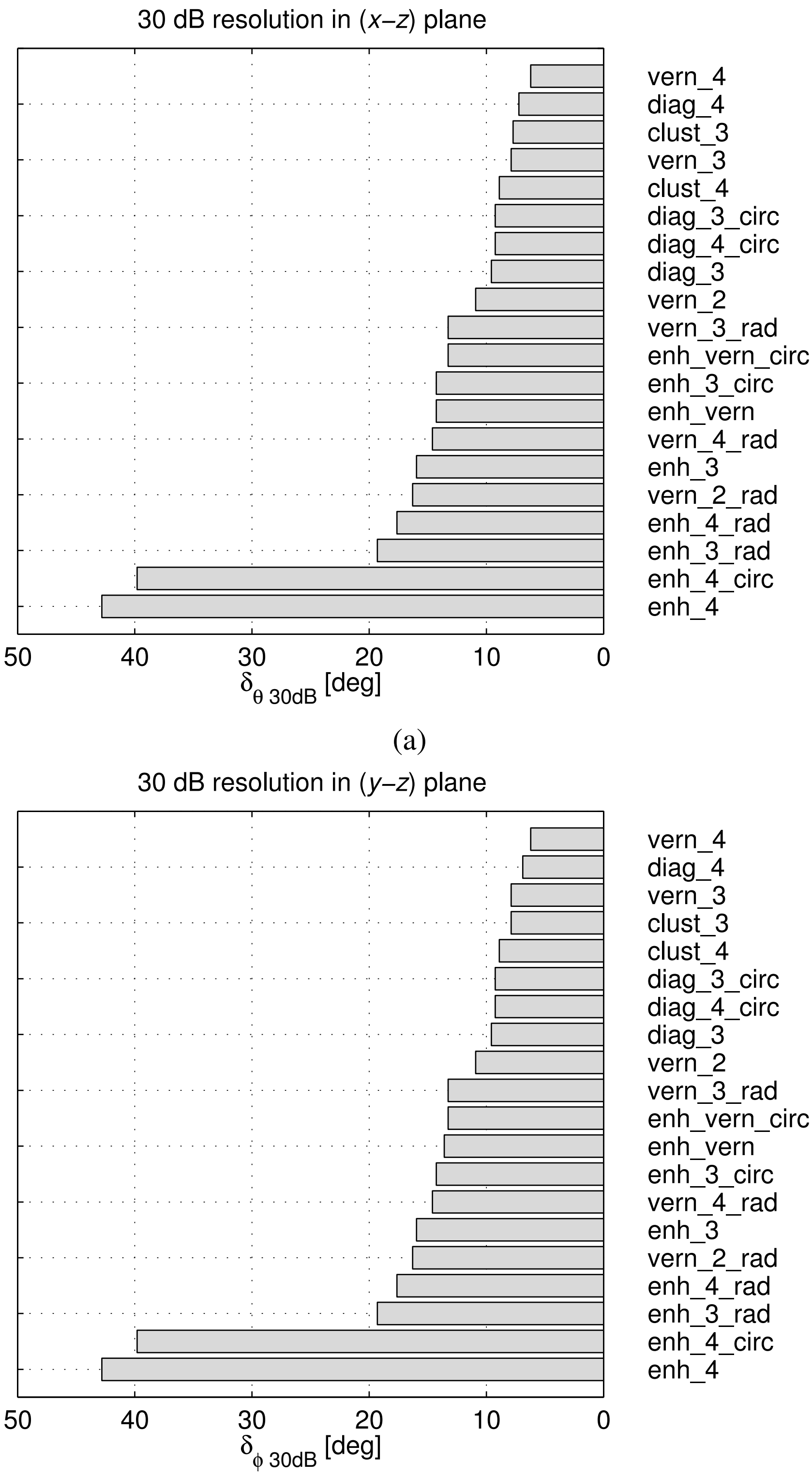


(a)



(b)

Figure 8.20: The angular resolution at -6 dB in the $(x-z)$ and $(y-z)$ planes.



(a)

(b)

Figure 8.21: The angular resolution at -30 dB in the $(x-z)$ and $(y-z)$ planes.

Parameter	Notation	Value	Unit
Speed of sound	c	1540	m/s
Center frequency	f_0	3	MHz
Sampling Frequency	f_s	105	MHz
Pitch	d_x	0.5	mm
Bandwidth	B	70	%

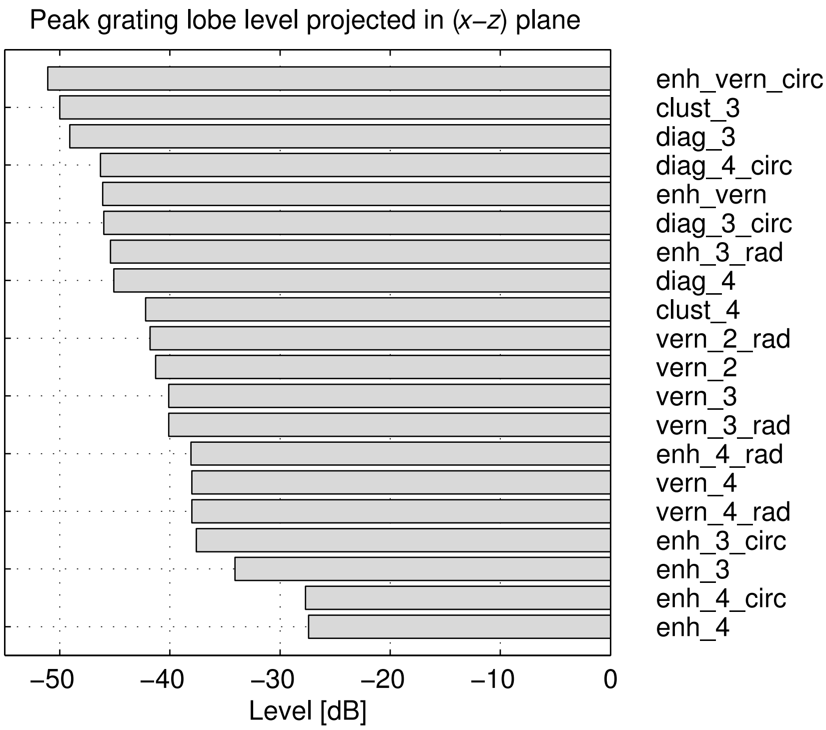
Table 8.2: Simulation parameters.

Design	$\delta_{\theta 6\mathrm{dB}}$ [deg]	$\delta_{\theta 30\mathrm{dB}}$ [deg]	$\delta_{\phi 6\mathrm{dB}}$ [deg]	$\delta_{\phi 30\mathrm{dB}}$ [deg]	IMLSLR [dB]	peak_xz [dB]	peak_yz [dB]
vern_2	4.53	10.91	4.53	10.91	25.69	-41.30	-41.30
vern_3	3.19	7.89	3.19	7.89	22.81	-40.10	-40.10
vern_4	2.52	6.21	2.52	6.21	19.85	-38.00	-38.00
vern_2_rad	6.88	16.29	6.88	16.29	24.11	-41.80	-41.80
vern_3_rad	5.88	13.27	5.88	13.27	21.90	-40.10	-40.10
vern_4_rad	5.88	14.61	5.88	14.61	18.68	-38.00	-37.00
enh_3	6.88	15.95	6.88	15.95	21.31	-34.10	-34.10
enh_4	5.37	42.82	5.37	42.82	16.44	-27.40	-27.40
enh_vern	5.88	14.27	5.71	13.60	26.49	-46.10	-46.10
enh_3_circ	6.21	14.27	6.21	14.27	22.78	-37.60	-37.60
enh_4_circ	4.87	39.79	4.87	39.79	16.81	-27.70	-27.70
enh_vern_circ	5.54	13.27	5.54	13.27	26.69	-51.10	-51.10
enh_3_rad	8.23	19.31	8.23	19.31	24.86	-45.40	-45.40
enh_4_rad	7.56	17.63	7.56	17.63	18.61	-38.10	-38.10
diag_3	3.86	9.57	3.86	9.57	27.33	-49.10	-49.10
diag_4	2.85	7.22	2.85	6.88	24.82	-45.10	-48.20
diag_3_circ	3.86	9.23	3.86	9.23	26.31	-46.00	-46.00
diag_4_circ	3.86	9.23	3.86	9.23	26.31	-46.30	-46.30
clust_3	3.19	7.72	2.52	7.89	21.60	-50.00	-48.50
clust_4	3.86	8.90	3.86	8.90	24.44	-42.20	-42.20

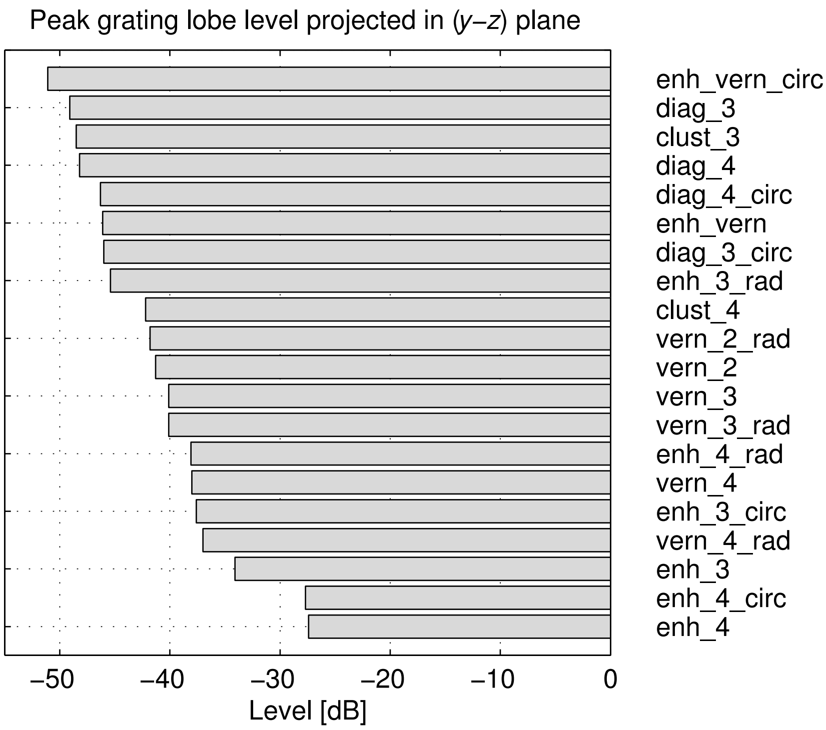
Table 8.3: Summary of some indicators for the performance of the different designs. The table shows the resolution in the azimuth δ_θ and elevation δ_ϕ planes at levels of -6 and -30 dB. The integrated main lobe to side lobe ratio and the peak grating lobe levels are also given.

In order to make the comparison between the arrays, their performance is plotted in separate figures (Figures 8.20 - 8.23), each depicting one of the indicators. The plots are given as bars, and are sorted with respect to the performance. The top bar corresponds to the best performance in the given category and the bottom bar to the worst performance.

Figure 8.20 gives the size of the point spread function at a level of -6 dB. Here the classical Vernier design vern_4 is unsurpassed in its performance, which is to be expected since the area covered by the elements in the array is the largest (see Table 8.1). It is closely followed by the diagonally optimized Vernier arrays. The array formed from 3 clusters also performs well, and equals the resolution of vern_4 in the $(y-z)$ plane. The size of the point spread function of the enhanced Vernier arrays is almost twice as large. This is to be expected, since the restrictions



(a)



(b)

Figure 8.22: The peak levels of the grating lobes.

are imposed on the *total* number of active elements. In order to have the same resolution the enhanced arrays must have at least four times as many elements [114]. The radially symmetric designs have the lowest performance

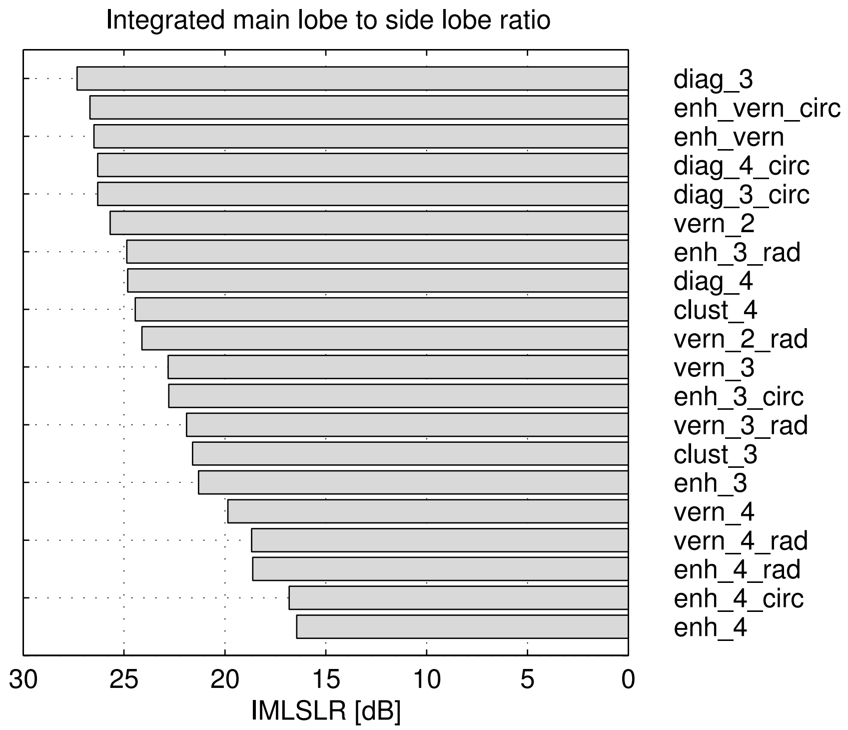


Figure 8.23: The integrated main lobe to side lobe ratio.

The performance order at the top is preserved also for the -30 dB beam width. The only significant difference is that both array pairs using clustered elements are in the top five designs.

While having best resolution, the classical Vernier arrays fall back when considering the peak side lobe level as seen from Figure 8.22. The top performing pair of arrays is enh_vern_circ as suggested by Austeng and Holm [114]. The diagonally optimized pair diag_3 and clust_3 have peak side lobe levels of 1 - 2 dB higher than enh_vern_circ. Most of top performing pairs are formed using the diagonally optimized Vernier arrays.

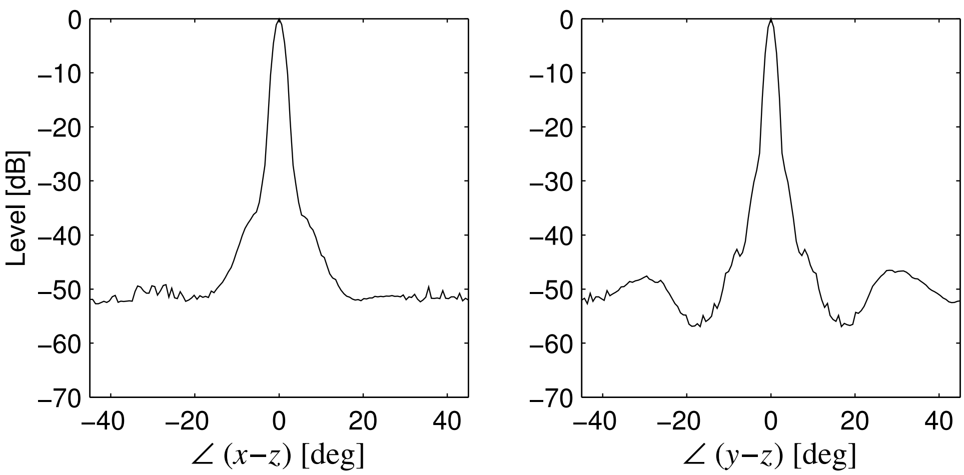
The diagonally optimized pair diag_3 has the highest ratio of the main lobe to side lobe energy as shown in Figure 8.23. Its performance is closely followed by enh_vern_circ, enh_vern, diag_4_circ and diag_3_circ.

As a conclusion, the classical Vernier arrays exhibit the highest resolution closely followed by the diagonally optimized pairs. One of the enhanced Vernier pairs has the lowest level of grating lobes, followed closely by the diagonally optimized arrays. One of the diagonally optimized arrays has the highest main to side lobe energy ratio. This makes the diagonally optimized arrays a good compromise between obtainable resolution and side and grating lobe energy leakage [3,4].

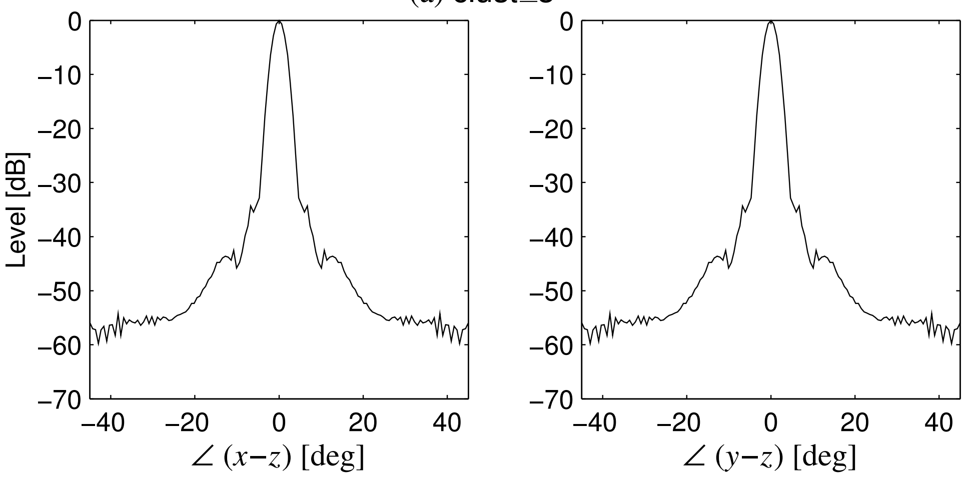
In spite of their non traditional for 2D ultrasound transducers appearance, the pairs clust_3

[3]The enhanced Vernier arrays perform best when constraints are not imposed on the number of active element, but on the total number of available elements in a matrix.

[4]Two crows had their chicks going to the same kindergarten. One day one of the crows was busy and asked the other one if she could take her chick back from the kindergarten. "How am I going to recognize it ?", asked the second crow. "It is very easy. Just look around. The best looking chick is mine.". In the evening, when the first crow came back from work she asked, "Did you bring back my chick ?". "I could not recognize it.", replied the other one. "I was looking around, but the best looking chick was mine." (Bulgarian fairy tale)



(a) clust_3



(b) clust_4

Figure 8.24: The projections of the point spread functions in $(x-z)$ and $(y-z)$ planes.

and clust_4 perform well, showing some of the best resolutions and lowest peak grating lobe levels. Another merit of these pairs is that the grating lobes appear close to the main lobe and monotonically decrease with the increasing angle. The projections of their point spread functions in the $(x-z)$ and $(y-z)$ planes are given in Figure 8.24.

8.5 Conclusion

In this chapter it was shown that the number of transmit events N_{xmt} can be significantly reduced using sparse transmit aperture. The choice of the active elements must be based on the transmit/receive radiation pattern. A very powerful approach is to consider the effective aperture of the array pairs. This is especially useful when designing 2D sparse arrays. The most suitable designs are based on the principle of Vernier arrays suggested by Lockwood. Using them as a basis the performance of the array pair can be improved by using what is called "diagonally optimized" pairs as suggested by the author, or enhanced arrays as suggested by Austeng and Holm. The latter have superior performance when the restrictions are on the available array matrix, while the former perform best when the restrictions are imposed on the available number of active elements.

CHAPTER

NINE

Coded excitations

One of the major problems of all synthetic aperture imaging techniques is the signal-to-noise ratio. The signal level decreases not only due to the tissue attenuation but also because of the very nature of the emissions - sending a spherical wave. This problem can be addressed in several ways: (1) using multiple elements to create virtual sources, (2) generating long pulses with temporal encoding, and (3) transmitting from several virtual sources at same time using spatial encoding.

The use of a long waveform is a well known and used technique in the radar systems. It involves either sending a long phase/frequency modulated pulse (PM or FM), or a series of short pulses with binary encoding.

Another way of encoding is the spatial encoding. Instead of sending only from a single virtual source, the transmission is done from all of the virtual sources at the same time and a binary encoding/decoding scheme is used to resolve the signals.

This chapter will start by discussing the properties of the matched filter, which is normally applied on the recorded data. The gain in signal-to-noise ratio and the obtained resolution will be given in terms of the signal duration and bandwidth. Then the pulse compression of linear frequency modulated pulses by using matched filter will be presented. The discussion on pulse compression will conclude with the binary Golay codes. Since the design of coded in time excitations goes beyond the scope of this work, the FM pulses used further on in the dissertation will be based on results obtained by my colleague Thanassis Misaridis.

Then the focus will be shifted towards the spatial encoding. The considerations will be limited to the use of Hadamard matrices for coding and decoding. The gain in signal-to-noise ratio will be briefly considered.

At the end of the chapter B-mode images from measurements on a phantom and *in-vivo* are presented. The scans are carried out with spatial, temporal and spatio-temporal encoding.

9.1 Temporal coding

9.1.1 The matched filter receiver

One of the steps in the reconstruction of images using synthetic aperture is the matched filtration (see Chapter 4). The matched filter also increases the signal-to-noise ratio which in an ultrasound system is limited by the peak acoustic power, rather than average power [115]. This condition can be expressed as:

$$PSNR = \frac{\max\left(|r_m(t)|^2\right)}{P_N}, \tag{9.1}$$

where $r_m(t)$ is the output signal from the filter and P_N is the average noise power. The filter which maximizes this ratio is the *matched filter* [47]. Consider a passive filter with a transfer function $\dot{H}_m(f)$ The absolute magnitude of the signal after the filter can be expressed in terms of the complex spectrum of the input signal $\dot{R}_i(f)$ as:

$$|r_m(t)| = \left| \int_{-\infty}^{\infty} \dot{R}_i(f)\dot{H}_m(f)e^{j2\pi ft}\, df \right|. \tag{9.2}$$

The noise power after the filter is given by:

$$P_N = P_0 \int_{-\infty}^{\infty} |H_m(f)|^2\, df, \tag{9.3}$$

where P_0 is the power density of the noise. By using Schwartz's inequality:

$$\left| \int_{-\infty}^{\infty} \dot{R}(f)\dot{H}(f)\, df \right|^2 \leq \int_{-\infty}^{\infty} \left|\dot{R}(f)\right|^2 df \int_{-\infty}^{\infty} \left|\dot{H}(f)\right|^2 df$$

one gets:

$$PSNR \leq \frac{\int_{-\infty}^{\infty} \left|\dot{R}_i(f)\right|^2 df \int_{-\infty}^{\infty} \left|\dot{H}_m(f)\right|^2 df}{P_0 \int_{-\infty}^{\infty} \left|\dot{H}_m(f)\right|^2 df}, \tag{9.4}$$

$$\left|e^{j2\pi ft}\right| = 1 \tag{9.5}$$

$$PSNR \leq \frac{\int_{-\infty}^{\infty} \left|\dot{R}_i(f)\right|^2 df}{P_0}. \tag{9.6}$$

The equality is only obtained when:

$$\dot{H}_m(f) = G_a \left[\dot{R}_i(f)e^{j2\pi ft_1}\right]^* \tag{9.7}$$

$$\dot{H}_m(f) = G_a \dot{R}_i^*(f)e^{-j2\pi ft_1} \tag{9.8}$$

where t_1 is a fixed delay, and G_a is a constant usually taken to be one. The impulse response of the filter in the time domain is then:

$$\boxed{h_m(t) = G_a r_i(t_1 - t).} \tag{9.9}$$

The filter that maximizes the SNR, thus has an impulse response equal to the input waveform with reversed time axis, except for an insignificant amplitude factor and translation in time. For this reason the filter is called *matched* filter. Aside from the amplitude and the phase constants,

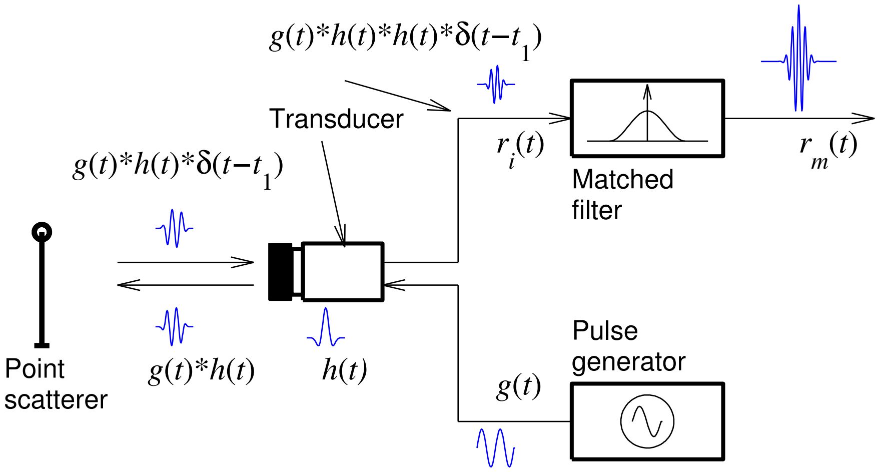


Figure 9.1: Schematic diagram of a single element transducer together with the generation of the pulse and the subsequent matched-filter processing. The picture illustrates the evolution of the time signal for a single scatterer placed at the focus or in the far field.

the transfer function of the matched filter is the conjugate of the signal spectrum. This means that maximization of the SNR is achieved by (1) removing any nonlinear phase function of the spectrum and (2) weighting the received spectrum in accordance with the strength of the spectral components in the transmitted signal.

Consider the one-dimensional measurement situation depicted on Figure 9.1. The pulse generator generates a radio frequency pulse $g(t)$. The pulse is fed into a transducer with an electromechanical impulse response $h(t)$. The generated pressure pulse $p(t) = g(t) \underset{t}{*} h(t)$ is sent into the tissue. There the pulse propagates until it reaches a point scatterer which scatters the pulse back. The propagation of the pulse in the tissue can be expressed as a convolution between the pressure pulse and the spatial impulse response (see Section 3.3.2). For the simplified case, the spatial impulse response will be assumed a delta function, which is valid for the focus and for the far field. The back-scattered signal, which reaches the transducer is $p(t) \underset{t}{*} \delta(t-t_1)$, where t_1 is the round-trip delay. The transducer converts the received pressure back to an electric signal $r_i(t)$, which is:

$$r_i(t) = \underbrace{g(t) * h(t) * h(t)}_{=g(t)} * \delta(t-t_1). \tag{9.10}$$

Assuming an ideal transducer with an impulse response $h(t) = \delta(t)$ one gets that the signal received by a single scatterer is a replica of the transmitted pulse delayed in time:

$$r_i(t) = g(t-t_1). \tag{9.11}$$

The output $r_m(t)$ of the matched filter with impulse response $h_m(t) = r_i^*(-t)$ is the autocorrelation function of the transmitted pulse:

$$r_m(t) = R_{gg}(t), \tag{9.12}$$

$$R_{gg}(\tau) = \int_{-\infty}^{\infty} g(t)g^*(t+\tau)\, dt, \tag{9.13}$$

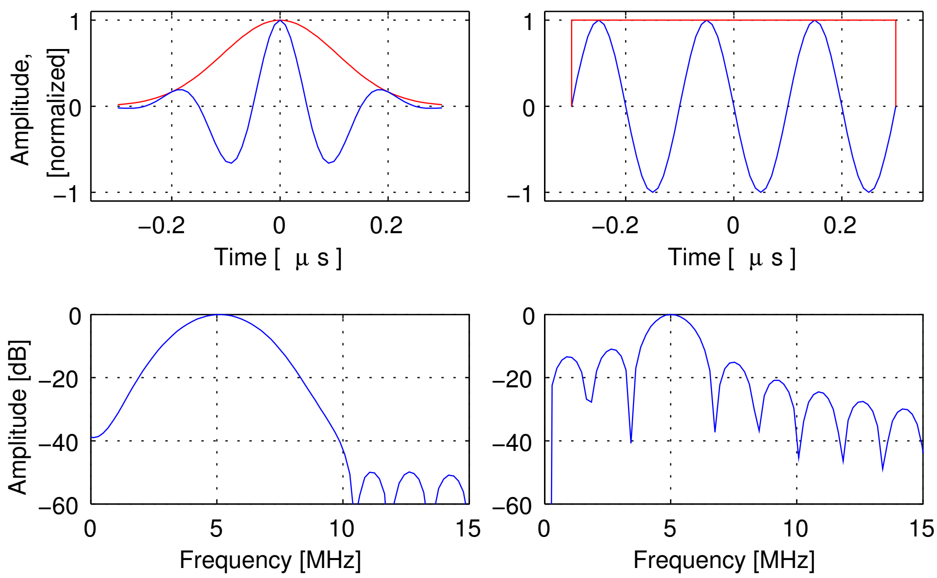


Figure 9.2: Examples of two radio pulses with center frequency f_0 = 5 MHz. The left column shows a pulse with a Gaussian window applied on it, and the right with a rectangular window applied on. The top row shows the signals in time domain, and bottom row in frequency domain. The windowing functions are drawn in red.

where the delay t_1 has been skipped for notational simplicity. The latter means that the maximum will occur at a lag corresponding to the two-way propagation time. The displayed image is the envelope of the autocorrelation function of the transmitted pulse. It can be seen that the imaging abilities of the system depend on the parameters of the autocorrelation function and therefore on the parameters of the transmitted signal.

9.1.2 Signal parameters

The signals used in ultrasound are usually RF pulses of the form:

$$g(t) = w(t)\cos[2\pi f_0 t + \phi(t)], \tag{9.14}$$

where $w(t)$ is a baseband modulation signal, f_0 is the carrier frequency, and $\phi(t)$ is the phase of the signal. For the sake of simplicity in the analysis the complex notation for the signals will be used:

$$\dot{g}(t) = w(t)e^{j\varphi(t)} \tag{9.15}$$

(9.16)

where $w(t)$ and $\varphi(t) = 2\pi f_0 t + \phi(t)$ are the amplitude and the phase, respectively.

Consider Figure 9.2 showing two RF pulses at a center frequency f_0 = 5 MHz. It can be seen that these pulses are localized in time, and therefore their power density spectra are infinite. In spite of the infinite nature of the spectra, most of the energy is concentrated around the central frequency f_0 within some band Δ_f which can be given in terms of the standard deviation [47, 116], and will be called RMS bandwidth. The energy contained in the signal is:

$$||\dot{g}(t)||^2 = \int_{-\infty}^{\infty} |\dot{g}(t)|^2\, dt = \int_{-\infty}^{\infty} |\dot{G}(f)|^2\, df = E_s \tag{9.17}$$

The mean time, and mean frequency can be found by:

$$\bar{t} = \frac{1}{E_s}\int_{-\infty}^{\infty} t|\dot{g}(t)|^2\, dt \tag{9.18}$$

$$\overline{f} = \frac{1}{E_s}\int_{-\infty}^{\infty} f|\dot{G}(f)|^2\, df. \tag{9.19}$$

Following the consideration in Section 2.4 in [116], the mean frequency $\overline{f}$ is found to be:

$$\overline{f} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{d\varphi(t)}{dt}\frac{|\dot{g}(t)|^2}{E_s}\, dt, \tag{9.20}$$

which says that the mean frequency $\overline{f}$ is the weighted average of the instantaneous quantities $\frac{d\varphi(t)}{dt}$ over the entire time domain. The value $\frac{d\varphi(t)}{dt} = f_i(t)$ is named *the mean instantaneous frequency* and is:

$$f_i(t) = \frac{1}{2\pi}\frac{d\varphi(t)}{dt} = f_0 + \frac{1}{2\pi}\frac{d\varphi(t)}{dt} \tag{9.21}$$

The RMS time duration $2\Delta_t$ and the RMS frequency bandwidth $2\Delta_f$ are defined as:

$$\Delta_t^2 = \int_{-\infty}^{\infty} (t-\bar{t})^2\frac{|\dot{g}(t)|^2}{E_s}\, dt \tag{9.22}$$

$$\Delta_f^2 = \int_{-\infty}^{\infty} (f-\overline{f})^2\frac{|\dot{G}(f)|^2}{E_s}\, df \tag{9.23}$$

They are found to be:

$$\Delta_t^2 = \int_{-\infty}^{\infty} t\frac{|\dot{g}(t)|^2}{E_s}\, dt - \bar{t}^2 \tag{9.24}$$

$$\Delta_f^2 = \frac{1}{4\pi^2 E_s}\int_{-\infty}^{\infty} \frac{d\varphi(t)}{dt}w(t)^2\, dt - \frac{j}{4\pi^2 E_s}\int_{-\infty}^{\infty} \frac{dw(t)}{dt}w(t)\, dt, \tag{9.25}$$

The latter result means that the bandwidth is completely defined by the amplitude variation $\frac{dw(t)}{dt}$ as well as phase variations $\frac{d\varphi(t)}{dt}$. If both, the amplitude and the phase are constant, such as the complex sinusoidal signal $\exp(j2\pi f_0 t)$ then the frequency bandwidth reduces to zero.

It can be shown that :

$$\Delta_t \cdot \Delta_f \geq \frac{1}{4\pi}. \tag{9.26}$$

This result[1] is known as the uncertainty principle and comes to show that the *time-bandwidth* (TB) product of a signal has a lower limit.

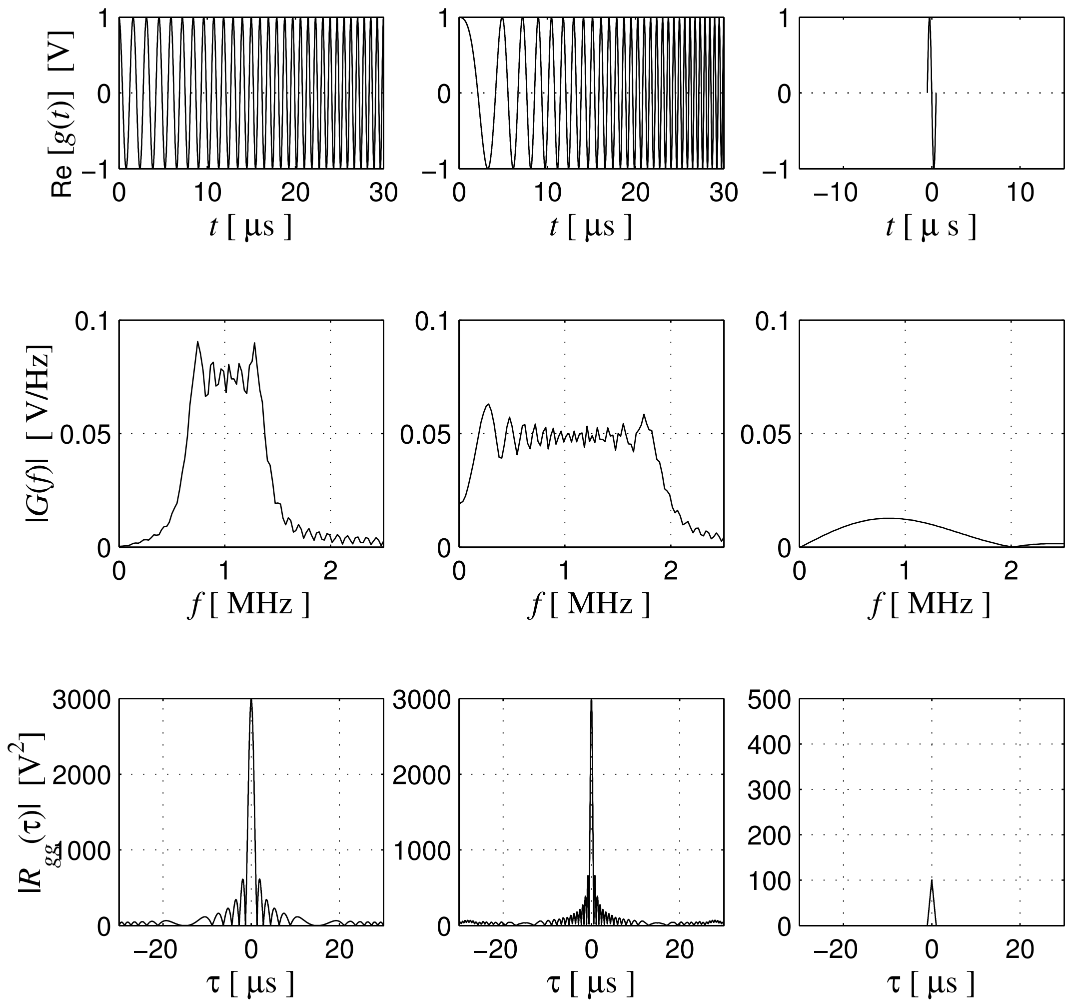


Figure 9.3: Illustration of the pulse compression and the influence of the pulse duration and bandwidth on the auto correlation function. The top row shows 3 real signals - two frequency modulated pulses and a one-cycle sine pulse. The duration of the FM pulses and the sine pulse is 30 μs and 1 μs respectively. The mean frequency $\overline{f}$ of the signals is 1 MHz. The middle row shows the frequency density spectra. The bottom row shows the magnitude (envelope) of their autocorrelation functions.

9.1.3 Pulse compression

As seen in Section 9.1.1 the signal that determines the imaging abilities of the system is $r_m(t)$, which is the output of the matched filter, and at the same time is the autocorrelation function of the transmitted pulse $g(t)$. The resolution of the system is determined by the width of the main lobe of $r_m(t)$. Rihaczek [47] reported an approximate equation for the half-power width of the main lobe:

$$\Delta\tau = \frac{1}{4\Delta_f} \tag{9.27}$$

[1] There are different definitions for the duration of the pulse and its bandwidth. Qian and Chen [116] use the same definition for Δ_t, but because they work with angular frequency ω, their product is $\Delta_t \cdot \Delta_\omega \geq \frac{1}{2}$. Rihaczek [47] on the other hand defines the RMS (root mean square) duration and bandwidth as $\delta = 2\pi\Delta_t$ and $\beta = 2\pi\Delta_f$, respectively. The uncertainty principle is then $\beta \cdot \delta \geq \pi$.

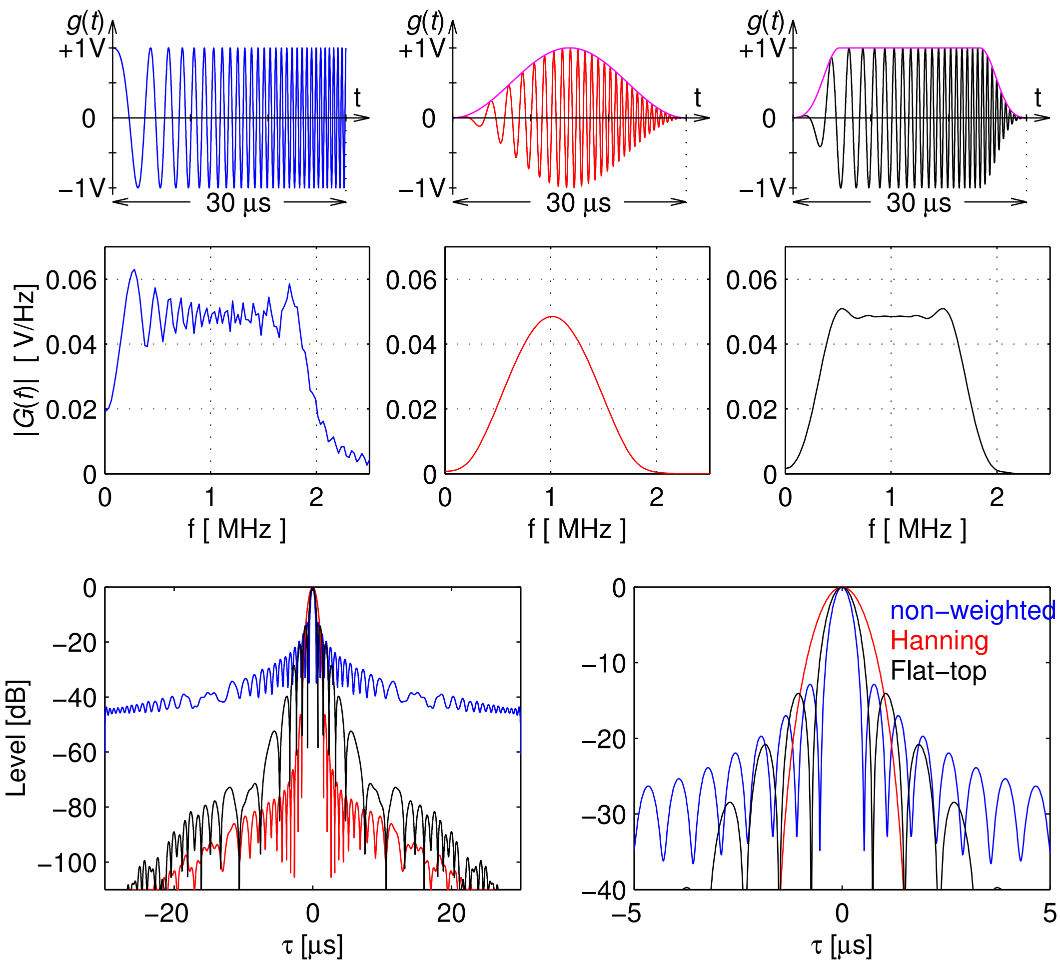


Figure 9.4: Reducing the side lobes by temporal weighting.

On the other, hand the peak value of $r_m(t)$ being the value of the autocorrelation function at lag 0, is equal to the energy contained in the signal. The gain in the signal-to-noise ratio is given by:

$$GSNR = \frac{SNR_{out}}{SNR_{in}} = TB, \tag{9.28}$$

where TB is the time-bandwidth product.

From these two results it is obvious that the duration of the pulse must be chosen according to the desired gain in signal-to-noise ratio. The bandwidth of the signal can be increased (See Eq. (9.25)) by applying either amplitude or phase modulation. Changing the phase does not affect the carried energy and therefore is the preferred approach. From (9.21) one will learn that any linear phase change only will change the mean frequency f_0, and therefore a non-linear modulation function for $\varphi(t)$ must be used. One set of signals are the frequency modulated signals, an example of which is shown in Figure 9.3. The two long signals are linear frequency modulated signals (chirps), which are defined as:

$$\dot{g}(t) = w(t)e^{j2\pi(f_0 t + \beta t^2)} \tag{9.29}$$

where β is a constant showing how fast the instantaneous frequency changes and $w(t)$ is a windowing function. All of the signals have a mean frequency $\overline{f}$ of 1 MHz. The chirp in the

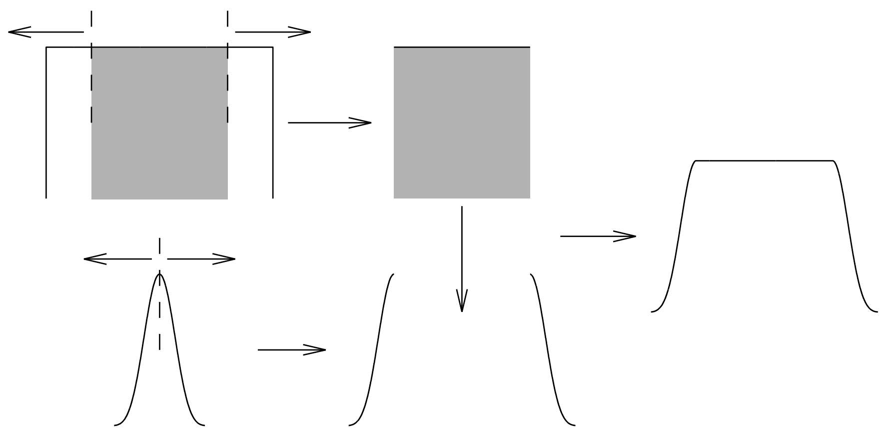


Figure 9.5: Illustration of the filter design for weighting of the chirps.

first column has RMS bandwidth $2\Delta_f$ of 807 kHz. The chirp in the middle column and the sine pulse have RMS bandwidths around 2 MHz. The figure clearly shows that the energy in the signals is proportional to their length, and that the width of main lobe of the autocorrelation function is inversely proportional to the bandwidth. Using either a chirp or a single frequency RF pulse with the same RMS bandwidth results in the same spatial/temporal resolution, but the energy contained in the chirp is bigger. From (9.28) the GSNR of the matched filter for the sine pulse and the middle chirp in Figure 9.3 are:

$$GSNR_{\text{chirp}} = T_{\text{chirp}} \times B_{\text{chirp}} \approx 30 \cdot 10^{-6}\ \text{s} \times 2 \cdot 10^{6} \frac{1}{\text{s}} = 17.78\ \text{dB} \tag{9.30}$$

$$GSNR_{\text{sine}} = T_{\text{sine}} \times B_{\text{sine}} \approx 1 \cdot 10^{-6}\ \text{s} \times 2 \cdot 10^{6} \frac{1}{\text{s}} = 3.01\ \text{dB} \tag{9.31}$$

This result shows that using the linear chirp given in Figure 9.3 instead of the sinusoidal pulse increases the signal-to-noise ratio with 14.7 dB.

Because of the nearly rectangular shape of the spectrum, the envelope of the autocorrelation function of a chirp signal has side lobes, which in the literature are known as range side lobes. These range side lobes reduce the contrast of the image and their level must be decreased. This can be done by applying a windowing function $w(t)$ different than a rectangle [117, 118] as shown in Figure 9.4. The top row and middle rows show three different signals in time and frequency domain, respectively. The bottom row shows a comparison between the output of their correlation functions. The narrowest main-lobe of the autocorrelation function is achieved by the non-tapered chirp, and the lowest side lobe level by the Hanning weighted chirp. A good compromise is the chirp apodized with a flat-top[2] filter. A substantial work in designing such flat-top filters was done by my colleague Thanassis Misaridis [119, 120] in his Ph.D. work and the signals used further on in the dissertation are based on his design.

The window used for the signal in the right-most column of Figure 9.4 was made by combining a rectangular and a Kaiser window functions as shown in Figure 9.5. The discrete Kaiser

[2]The term flat-top here is used to denote that most of the filter coefficients are equal to 1. Normally this term is used to denote filters with flat frequency response, which is not the case here.

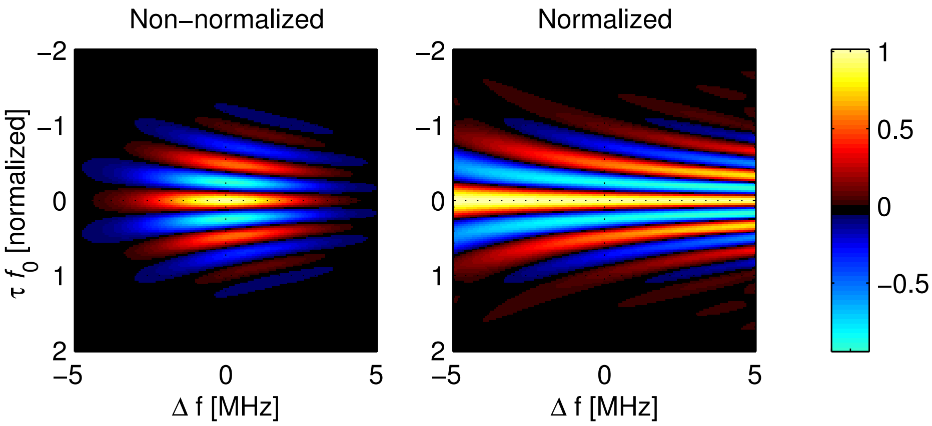


Figure 9.6: The output of a filter when there is a mismatch between the center frequency of the signal and the center frequency of the impulse response of the filter.

window $w[n] = w(n/f_s)$ is defined as [53]:

$$w[n] = \begin{cases} \dfrac{I_0\left[\zeta\sqrt{1-\frac{(n-\alpha)^2}{\alpha^2}}\right]}{I_0(\zeta)}, & 0 \le n < M, \\ 0 & \text{otherwise} \end{cases} \tag{9.32}$$

where f_s is the sampling frequency, n is the sample number, M is the number of filter coefficients, $I_0(\cdot)$ is a zeroth order modified Bessel function of the first kind, and $\alpha = \frac{M}{2}$. In the depicted case (Figures 9.4 and 9.5) the length of the Kaiser window is 40 % of the total length of the filter, the sampling frequency is $f_s = 100$ MHz, and $\zeta = 7.8$.

Up to this moment it was assumed that the received signal is a replica of the transmitted pulse. The design of the matched filter should also take into consideration the influence of the transducer bandwidth. Also, if some other optimization criteria are chosen, then the filter used for compression might deviate from the matched filter. This problem goes, however, beyond the scope of this work and will not be treated further.

Another assumption tacitly made to this moment was, that the received waveform had the same frequency content as the emitted one. The ultrasound energy, however, while propagating in the human body attenuates. The attenuation is frequency dependent and is bigger for higher frequencies. The mean frequency and bandwidth are therefore decreased. Because of the difference in the frequency content between the transmitted and received pulses, the matched filter is not any more matched. Figure 9.6 shows the real part of the output of a filter matched to a Gaussian pulse with a mean frequency $\overline{f} = f_0 = 5$ MHz and bandwidth $2\Delta_f = 3.13$ MHz. The input signal was a Gaussian pulse, whose mean frequency was varied. The left plot shows the response of the filter, when all the responses are normalized with respect to the highest output, which occurs at $\Delta f = 0$ and $\tau = 0$. The right plot shows the outputs, each of which was normalized by the highest value contained in it. It can be seen that both the amplitude and the shape of the output changes. Such an output is described by the *ambiguity function*.

9.1.4 The ambiguity function

The ambiguity function is defined [47, 116][3] as:

$$\chi(\tau, f_d) = \int_{-\infty}^{\infty} \dot{g}(t)\dot{g}^*(t-\tau)\exp(j2\pi f_d t)\, dt, \tag{9.33}$$

It can be seen that when there is no difference in the frequency, i.e. $f_d = 0$, the ambiguity function is the matched filter response. The value of the ambiguity function away from the origin shows the response of a filter which is mismatched by a delay τ and frequency shift f_d relative to the transmitted signal. Of interest is only the behavior of the function when there is a frequency difference.

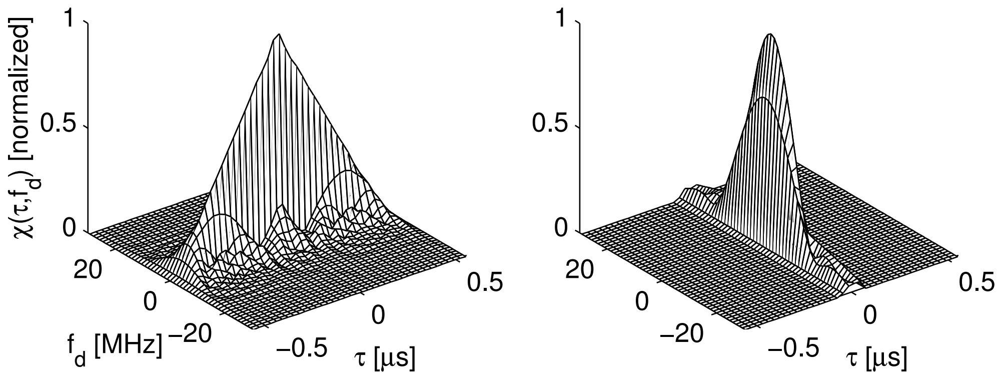


Figure 9.7: The ambiguity functions of a long (left) and short (right) pulses. The short pulse is well localized in time, while the long pulse is well localized in frequency.

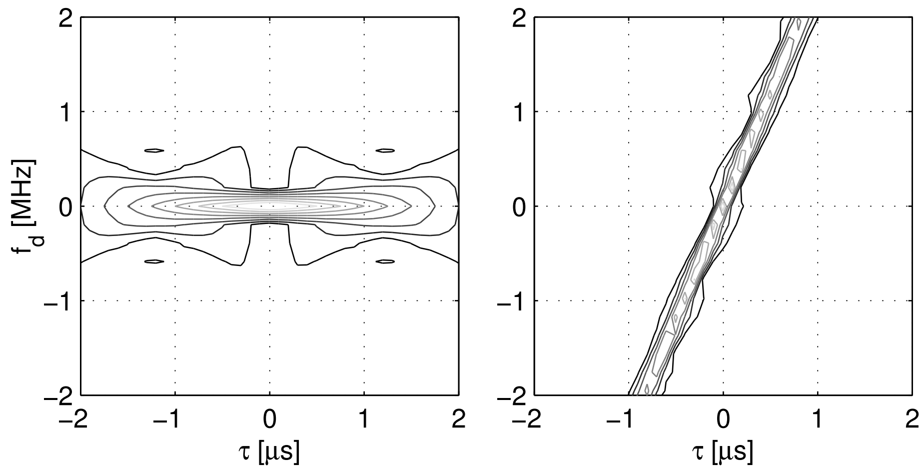


Figure 9.8: Comparison between the ambiguity function of an RF pulse (left) and a chirp (right) with the same duration.

[3]In some sources this formula is known as the *narrow-band* ambiguity function.

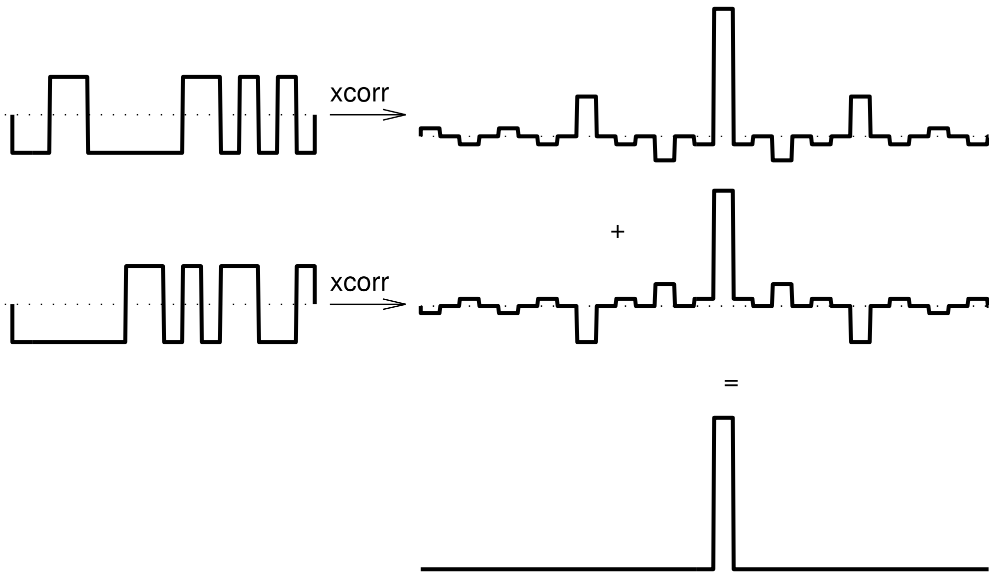


Figure 9.9: Compression using Golay codes.

Consider Figure 9.7 which shows the properties of the ambiguity function. A short pulse is well localized in time, but not in frequency domain. The output of the frequency mismatched filter will produce a sharp peak for the short pulse. The long pulse, however will be entirely filtered out.

Figure 9.8 shows the ambiguity functions of an RF pulse with a single frequency carrier with $f_0 = 5$ MHz and a linear frequency modulated RF pulse ranging from $f_1 = 2$ to $f_2 = 8$ MHz. The duration of both pulses is $T = 5$ μs. The linear FM modulated pulse has a narrow peak of the autocorrelation function inversely proportional to its bandwidth. The single carrier RF pulse has a significantly larger peak of the autocorrelation function proportional to the length of the pulse. A frequency mismatch during the filtration of the single carrier RF pulse does not change the position of the maximum, but its amplitude gets smaller and after frequency mismatch $\Delta f = f_d = 1$ MHz the signal is completely filtered out. A frequency mismatch in the filtration of the linear FM pulse results in a time shift of the peak of the autocorrelation function. The shape of the autocorrelation function remains, however, unchanged making the use of linear FM chirps appealing in the presence of the Doppler shift as it will be discussed in Chapter 10.

9.1.5 Binary codes

The binary codes have been widely investigated for ultrasound imaging, and though not used in this work they are worth noting. Several types of codes such as Barker codes [121, 122], and m-sequences [123, 124] have been investigated. However, the most promising seem to be the codes based on Golay pairs which have been tried by a number of groups around the world. In 1979 Takeuchi proposed and described a system using Golay codes [125, 126]. Mayer and colleagues have successfully applied coded excitation based on Golay codes in synthetic aperture imaging [57]. More recently a group from General Electric has been investigating the use of such codes together with spatial Hadamard encoding [127, 128]. The idea behind the use of Golay codes is illustrated in Figure 9.9. Two binary codes are necessary: A and B. Each

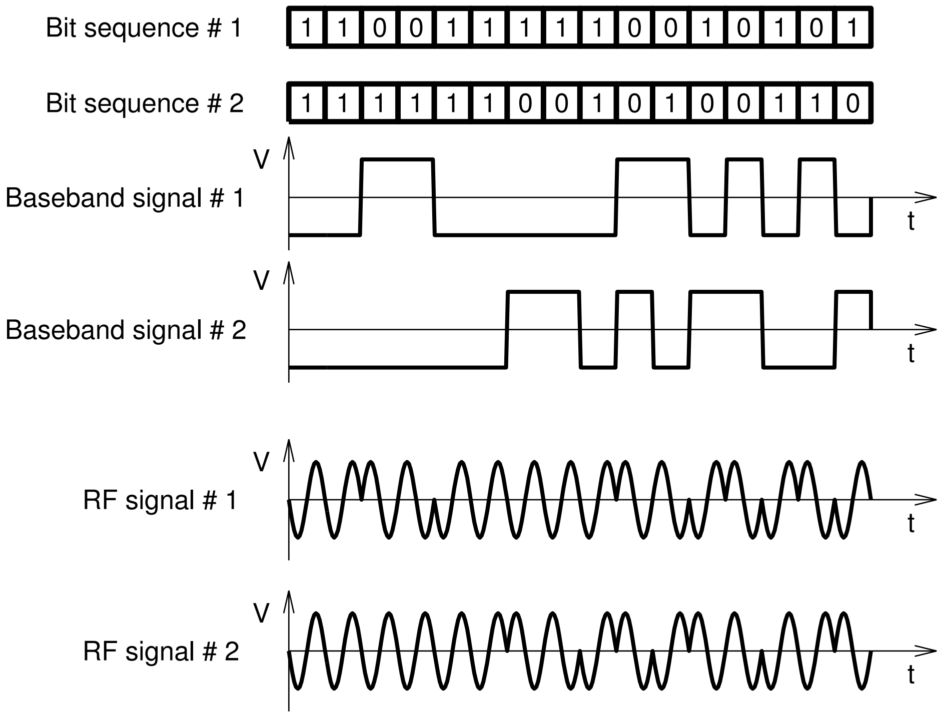


Figure 9.10: Creating waveforms using complementary Golay codes. From top to bottom : the golay pair; base-band signal; RF signal.

of the binary codes has an autocorrelation function, whose side lobes are equal in magnitude and opposite in sign to the corresponding side lobes of the other code. Thus the sum of the autocorrelation functions results in cancellation of the side lobes as shown in the figure. A good description of what Golay sequences are, and how to construct them is given by Đoković [129].

The codes being binary are encoded as a waveform using phased modulation. Zero corresponds to an initial phase 0, and 1 to an initial phase π. The creation of the binary waveform for a 16 bit Golay pair is illustrated in Figure 9.10. The baseband signal contains rectangular pulses with amplitude ± 1, depending on the phase. The RF signal is a 5 MHz phase modulated signal. For each of the bits one period at the carrier frequency is used.

For imaging purposes the use of Golay pairs means to transmit two times in the same direction- the first time using sequence A, the second time sequence B. The received signals are matched-filter with the respective matched filters for each of the transmitted pulses. The beamformed signals from the two emissions are summed together prior to envelope detection and display.

Figure 9.11 shows the performance of the Golay pairs illustrated in Figure 9.10. With black line is given the envelope of the sum of the autocorrelations functions of the RF signals assuming a transducer with impulse response $\delta(t)$. The blue line shows the envelope of the same transmitted pulse, but in the presence of a transducer with two-sided fractional bandwidth of 65 % around 5 MHz. The main lobe gets wider in the presence of a transducer. The red line shows what would happen if only a binary waveform at baseband is transmitted and used for matched filter. The main lobe gets wider and the side-lobes get higher.

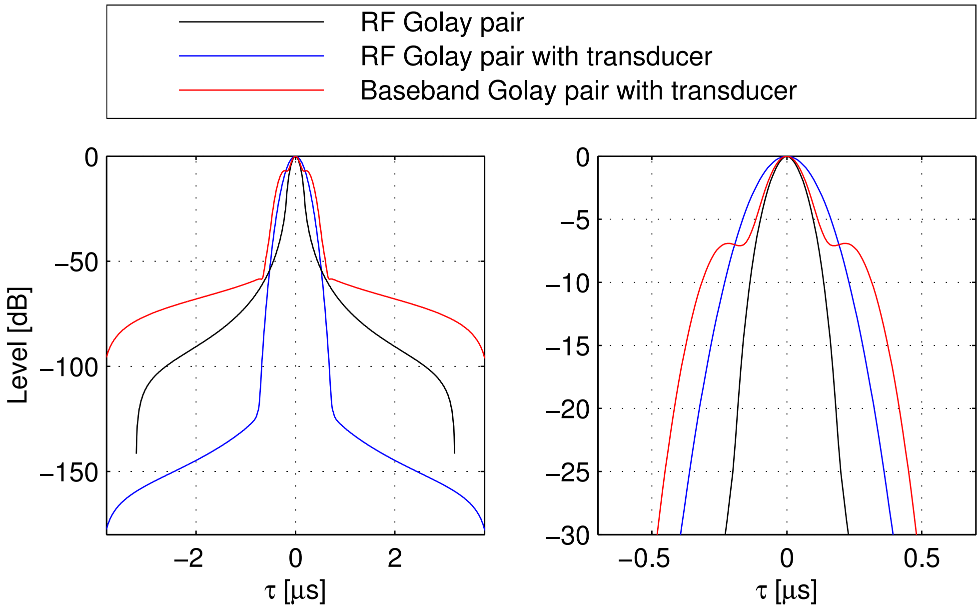


Figure 9.11: Performance of the Golay codes shown in Fig. 9.10.

9.2 Spatial coding

9.2.1 Principle of spatial coding

The spatial encoding was first suggested by Silverstein in [130] for calibration of satellite antennas. Later on, the same year, Chiao, Thomas and Silverstein [3], suggested this method to be used for ultrasound imaging.

The idea behind it is very simple. In synthetic transmit aperture N_{xmt} emissions are necessary to create an image. Every emission is done by using a single virtual source element. The transmitted energy is proportional to the number of elements forming the virtual source, but there is a limit to how many these elements can be. One way to increase the transmitted power is instead of transmitting only by a single virtual source, to transmit with all of the virtual sources at the same time. The beamforming is based on the round trip from the virtual source to the receive elements. It is therefore necessary to distinguish which echo from the emission of which source is created. This can be done by a set of mutually orthogonal signals. Such codes haven't been found yet. Another way is to send N_{xmt} times, every time with all of the N_{xmt} virtual sources with with some kind of encoding applied on the apodization coefficients. Let the signal transmitted by a single source i be:

$$g_i(t) = q_i \cdot g(t), \qquad 1 \le i \le N_{xmt}, \tag{9.34}$$

where $g(t)$ is a basic waveform and q_i is an encoding coefficient.

Assuming a linear propagation medium, the signal $r_j(t)$ received by the jth element can be expressed as:

$$r_j(t) = \sum_{i=1}^{N_{xmt}} q_i \cdot r_{ij}(t), \tag{9.35}$$

where $r_{ij}(t)$ would be the signal received by element j, if the emission was done only by the ith source.

From Eq. (5.8) on page 51:

$$L_i(t) = \sum_{j=1}^{N_{xdc}} a_{ij}(t) r_{ij}(t - \tau_{ij}(t))$$

it can be seen that the components $r_{ij}(t)$ must be found in order to beamform the signal. The received signals can be expressed in a matrix form:

$$\begin{bmatrix} r_j^{(1)}(t) \\ r_j^{(2)}(t) \\ \vdots \\ r_j^{(N_{xmt})}(t) \end{bmatrix} = \begin{bmatrix} q_1^{(1)} & q_2^{(1)} & \cdots & q_{N_{xmt}}^{(1)} \\ q_1^{(2)} & q_2^{(2)} & \cdots & q_{N_{xmt}}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ q_1^{(N_{xmt})} & q_2^{(N_{xmt})} & \cdots & q_{N_{xmt}}^{(N_{xmt})} \end{bmatrix} \begin{bmatrix} r_{1j}(t) \\ r_{2j}(t) \\ \vdots \\ r_{N_{xmt}j}(t) \end{bmatrix} \tag{9.36}$$

where the superscript $^{(k)}$, $1 \le k \le N_{xmt}$ is the number of the emission, $q_i^{(k)}$ is the encoding coefficient applied in transmit on the transmitting source i, and $r_j^{(k)}(t)$ is the signal received by the jth element. It was tacitly assumed that the tissue is stationary, and the building components of the signals $r_j^{(k)}$ are the same for all the emissions:

$$r_{ij}^{(1)}(t) = r_{ij}^{(2)} = \cdots = r_{ij}^{(k)}(t) = r_{ij}(t) \tag{9.37}$$

More compactly, the equation can be written in a vector form as:

$$\vec{r}_j(t) = \mathbf{Q}\vec{r}_{ij}(t), \tag{9.38}$$

where $\mathbf{Q}$ is the encoding matrix. The components $r_{ij}(t)$ can be found by solving the above equation:

$$\vec{r}_{ij}(t) = \mathbf{Q}^{-1}\vec{r}_j(t) \tag{9.39}$$

Every invertible matrix $\mathbf{Q}$ can be used for encoding and decoding. In Silverstein's work [130] prooves are given that the class of equal amplitude renormalized unitary matrices such as the 2-D DFT and Hadamard matrices are optimal for encoding of coherent signals. In the work of Chiao and colleagues [3], the choice was on using bipolar Hadamard matrices. Their advantage is the use of only two coefficients: +1, and -1. The Hadamard matrix is given by its order N (for more of the properties of the Hadamard matrices and the Hadamard transform, one can see [131], Section 3.5.2). The lowest order is $N = 2$, and the order can be only an even number. The Hadamard matrix of order 2 is defined as:

$$\mathbf{H}_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \tag{9.40}$$

The matrix of order $2N$ is recursively defined by the matrix of order N as:

$$\mathbf{H}_{2N} = \mathbf{H}_2 \otimes \mathbf{H}_N \tag{9.41}$$

$$= \begin{bmatrix} \mathbf{H}_N & \mathbf{H}_N \\ \mathbf{H}_N & -\mathbf{H}_N \end{bmatrix}, \tag{9.42}$$

where $\otimes$ is a Kronecker product. The inverse of the Hadamard matrix is the matrix itself:

$$\mathbf{H}_N^{-1} = \frac{1}{N}\mathbf{H}_N, \tag{9.43}$$

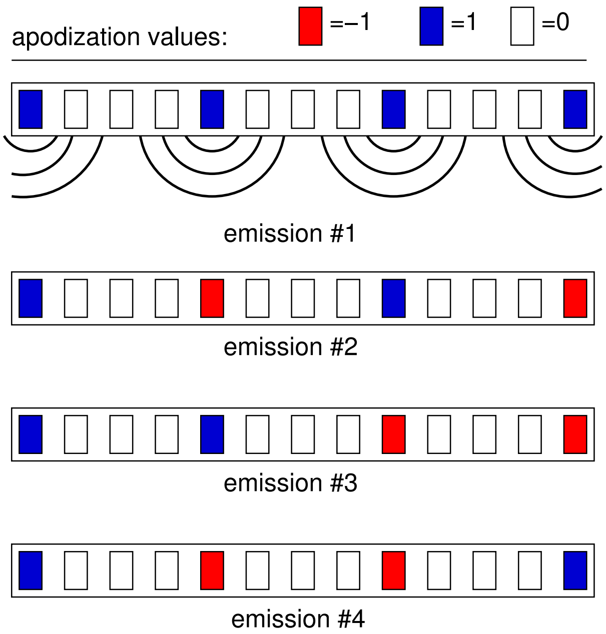


Figure 9.12: Spatially encoded transmits using 4 transmit elements.

which makes the decoding quite easy. The advantages of the Hadamard matrix include: (1) simple encoding mechanism - change in the polarity of the signals, (2) no multiplications are involved in the decoding stage, (3) there exist fast algorithms for the Hadamard transform significantly reducing the number of operations.

Figure 9.12 shows how the spatial encoding is applied for a Hadamard matrix of order 4. The elements are evenly distributed across the whole aperture. These can be either single elements, or a group of elements creating a virtual source, whose apodization coefficients are multiplied with the encoding values of the Hadamard matrix.

Let's consider an example of coding and decoding using a 2×2 Hadamard matrix. The transmitting elements have indices 1 and N_{xmt}. The signal received by a single element j is:

$$\begin{aligned} r_j^{(1)}(t) &= r_{1j}(t) + r_{N_{xmt}j}(t) \\ r_j^{(2)}(t) &= r_{1j}(t) - r_{N_{xmt}j}(t) \end{aligned} \tag{9.44}$$

Getting the sum and the difference of the signals $r_j^{(1)}(t)$ and $r_j^{(2)}(t)$ one gets:

$$r_j^{(1)}(t) + r_j^{(2)}(t) = 2r_{1j}(t) \tag{9.45}$$

$$r_j^{(1)}(t) - r_j^{(2)}(t) = 2r_{N_{xmt}j}(t) \tag{9.46}$$

The resulting components can be directly plugged into the beamformation unit.

9.2.2 Gain in signal-to-noise ratio

The gain in signal to noise ratio in a synthetic transmit aperture system is proportional to the square root of the number of receive elements N_{rcv} and the number of virtual transmit sources

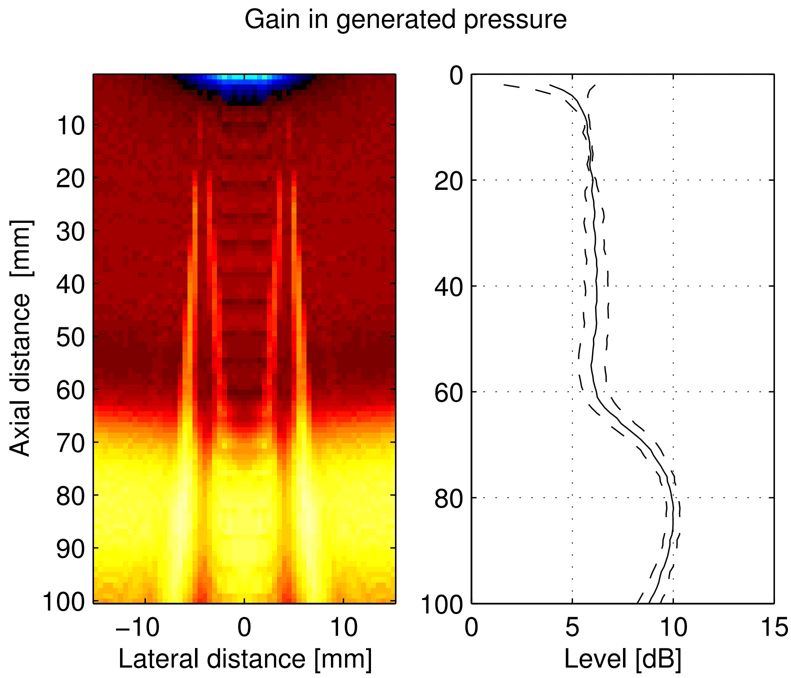


Figure 9.13: The gain in pressure. The plot shows the ratio of the maximum of the amplitude of the generated pressure using a single element emission and a 4 elements Hadamard encoded emission. The plot is obtained by simulation in Field II. The left image shows the spatial distribution and the right shows the magnitude.

N_{xmt}:

$$GSNR_{\mathrm{STA}} \sim SNR_0 \sqrt{N_{xmt} N_{rcv}}, \tag{9.47}$$

where SNR_0 is the signal-to-noise ratio of a single element. This situation is improved by the use of spatial Hadamard encoding. Consider again the case of only two emissions:

$$r_j^{(1)}(t) = r_{1j}(t) + r_{N_{xmt}j}(t) + n^{(1)}(t) \tag{9.48}$$

$$r_j^{(2)}(t) = r_{1j}(t) - r_{N_{xmt}j}(t) + n^{(2)}(t), \tag{9.49}$$

where $n^{(1)}(t)$ and $n^{(2)}(t)$ represent independent white noise received at the two emissions. Summing the equations to get the signal out, one gets:

$$\hat{r}_{1j}(t) = r_{1j}(t) + \frac{1}{\sqrt{2}} n(t), \tag{9.50}$$

where $n(t)$ is noise. The signal to noise ratio of the decoded data set is $10 \log_{10} 2$ decibels better than the dataset obtained from transmitting by the two transmitters separately. The gain in signal-to-noise ratio according to Chiao and colleagues. [3] is:

$$GSNR_{\text{Hadamard STA}} \sim SNR_0 N_{xmt} \sqrt{N_{rcv}}. \tag{9.51}$$

Assuming that the signal-to-noise ratio is inversely proportional to the transmitted energy it is intuitive that the more elements emit, the higher is the transmitted energy, and the bigger the signal-to-noise ratio is. Figure 9.13 shows the gain in transmitted pressure in decibels. This plot

Parameter	Notation	Value	Unit
Center frequency	f_0	7.5	MHz
Sampling frequency	f_s	40	MHz
Pitch	d_x	208	μm
Element width	w	173	μm
Element height	h	4.5	mm
Elevation focus	F	25	mm
Fractional bandwidth	B	50	%
Number of transducer elements	N_{xdc}	192	-
Number of receive channels	N_{rcv}	64	-

Table 9.1: Some parameters of the transducer and the system.

was obtained by simulations with Field II. At every spatial point the pressure field as a function of time was simulated for the case of emitting with 4 evenly distributed across the aperture elements and for the case of a single transmitting element. The maximum of the pressure was taken and the ratio for the case of encoded and non-encoded transmit is displayed.

9.3 Experimental phantom and clinical trials

The measurements, that will be presented in this section are in no case a complete study of the signal-to-noise ratio and its improvement when coded excitations are used. They are a mere demonstration that: (1) using codes gives better images with higher penetration depth, and (2) synthetic aperture can be used in a clinical situation. The measurements were done on a tissue mimicking phantom with a frequency dependent attenuation of 0.5 (dB/(cm MHz)), and on a soon-to-graduate Ph.D. student.

The measurements were carried out using RASMUS (see: Appendix G for more details). The transducer is a linear array transducer type 8804 by B/K Medical A/S, Gentofte, Denmark. The relevant parameters of the transducer and the system are listed in Table 9.1. In transmit up to 128 elements are available. The amplitude of the voltage applied on the transducer elements was 50 V. The maximum gain in receive was 40 dB.

The experiments were intended to try the following combinations of encodings: (1) no encoding, (2) only temporal encoding, (3) only spatial encoding, and (4) temporal and spatial encoding. The number of active elements N_{act} used to create a single virtual element for the different experiments is given in Table 9.2.

In the table N_{xmt} gives the number of emissions used to create a single frame. The conventional pulse, denoted with the abbreviation “sin” is a 3 cycles Hanning weighted sinusoid at the center frequency $f_0 = 5$ MHz. The duration of the FM pulse, denoted with “chirp” is $T = 20$ μs, and the fractional bandwidth is 80 %. The number of elements when spatial encoding is used, varies such that there is a gap of at least one non-active element between two neighboring virtual sources. The use of spatial encoding is symbolized by the suffix “_had”.

The results from the phantom measurements are shown in Figure 9.14. The images are logarithmically compressed and have a 50 dB dynamic range. The number of emissions used to create each of the images is given in the titles of the plots. Consider the first row ($N_{xmt} = 4$).

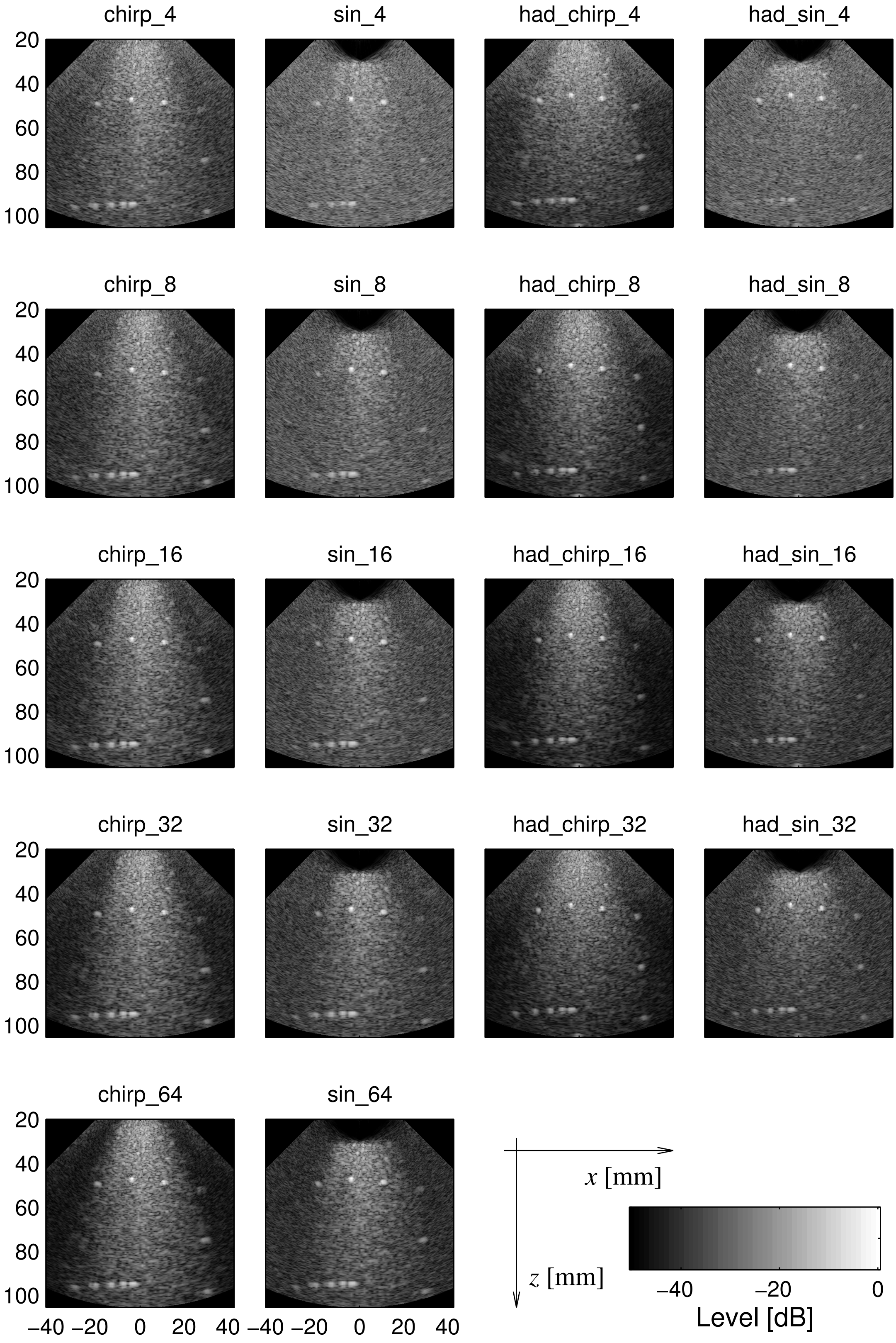


Figure 9.14: B-mode images of a tissue mimicking phantom for different combinations of temporal and spatial encoding schemes, and number of transmissions.

	$N_{xmt}=4$	$N_{xmt}=8$	$N_{xmt}=16$	$N_{xmt}=32$	$N_{xmt}=64$
sin	17	17	17	17	17
chirp	11	11	11	11	11
sin_had	13	5	3	1	-
chirp_had	13	5	3	1	-

Table 9.2: The number of active elements N_{act} used for the different experiments as a function of the number of emissions and the encoding scheme. sin means "conventional pulse", chirp means "FM pulse", and had means "using Hadamard encoding". N_{xmt} is the number of emissions per frame.

At depth z of 95 mm, from $x = -35$ to $x = 0$ mm in axial direction, there are 5 point scatterers. All of them can be seen only on the scans, in which FM pulses were used in transmit. Consider the vertical row of 3 point scatterers, lying at a lateral position $x \approx 35$ mm,. It can not be seen in sin_4, and can be best observed in had_chirp_4. The visibility of the images in ascending order is: sin_4, had_sin_4, chirp_4, and had_chirp_4. An increased number of emissions N_{xmt} increases the image quality for all encoding schemes.

Figure 9.15 shows the results from scanning the carotid artery. The position of the transducer is slightly different for the different scan situations and the results are not directly comparable. The general conclusion from these images is that the synthetic aperture can be used *in-vivo*.

There is a possibility to use a combination of two different excitations, whose cross-correlation is small, and combine them with spatial and temporal encoding using a Hadamard matrix [127]. The transmit sequence is:

$$\begin{array}{cccc} A_1 & A_2 & B_3 & B_4 \\ A_1 & -A_2 & B_3 & -B_4, \end{array}$$

where A and B are two orthogonal excitations. First, Hadamard decoding is used to obtain the signals $A_1 + B_3$ and $A_2 + B_4$ emanating from the different virtual sources. Then correlation is applied on the result to separate A_1 from B_3, and A_2 from B_4. This procedure makes it possible to use 4 virtual sources, and only two emissions. This, however, increases even more the complexity of the system. The author has no doubt that in the future more and more sophisticated encoding schemes will be utilized in the scanners.

9.4 Conclusion

This chapter presented the basics of using coded excitations for ultrasound imaging combined with spatial encoding. In this work they are used only to increase the signal-to-noise ratio, rather than decrease the number of emissions. These encoding schemes have been implemented and used in the experimental scanner RASMUS developed at CFU. B-mode images of phantom and *in-vivo* experiments were given. The general conclusion is that at low number of emissions the image quality is increased by using FM pulses and spatial Hadamard encoding.

The second part of the dissertation deals with blood flow estimation. The estimates are mostly influenced by noise, and therefore the use of coded excitations is a must.

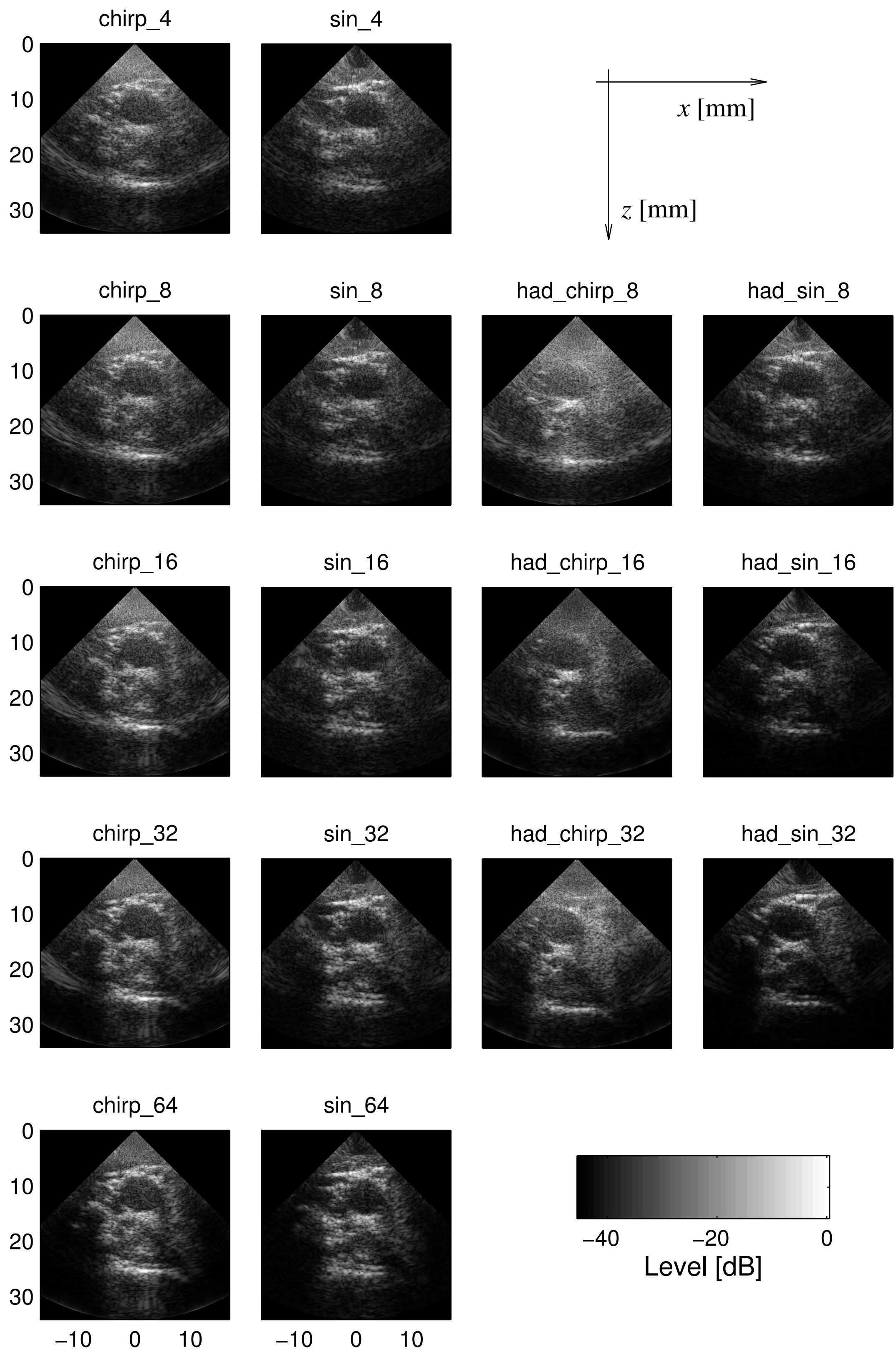


Figure 9.15: B-mode images of a carotid artery for different combinations of temporal and spatial encoding schemes, and number of transmissions.

Part II

Flow imaging

CHAPTER

TEN

Velocity estimation using cross-correlation

A real-time 3D ultrasound system would not be complete without the capability of estimating the blood velocity. The critical parameter of such a system is the acquisition time, and hence the number of emissions per volume scan. The sparse synthetic transmit aperture focusing is a prime candidate. To my knowledge there have not been previous reports on *blood velocity* estimations using synthetic aperture ultrasound imaging. The purpose of the following chapters is to show the specifics when estimating the velocity using ultrasound, rather than investigate the performance of all existing blood flow estimators [1]. Since the number of emissions per volume is limited, it is advantageous to use the same information for both, tissue imaging, and color flow mapping of the blood velocity. A suitable estimator is, thus, the cross-correlation estimator as suggested by Bonnefous and Pasqué [133], Foster [134, 135], Embree [136] and O'Brien [137, 138]. The performance of the estimator is inversely proportional to the length of the transmitted pulse, i.e. the shorter the transmitted pulse, the better the estimation [132, 135], which is in unison with the resolution requirements for B-mode imaging.

This chapter describes the velocity estimation using cross-correlation. The description follows Chapter 8 in "*Estimation of blood velocities using ultrasound. A signal processing approach*"[132]. The idea of this estimator is to measure the distance traveled by the scatterers for a period of time. The traveled distance is estimated by the difference in the arrival times of the signal scattered back by a group of scatterers.

10.1 Measurement principle

The considerations are confined to the simplified one dimensional case illustrated in Figure 10.1. The transducer is assumed to be ideal in all aspects: no frequency dependent sensitivity, linear conversion from electric energy to acoustic energy and vice versa. The medium is assumed perfectly elastic so that the pressure wavefront through the medium is proportional to the emitted signal $p(t)$. At the depth of interest the wavefront is assumed perfectly plane. If a single scatterer positioned in front of the transducer, is insonified then the received echo $r(t)$ is

[1]A rather complete review of the existing methods can be found in "Estimation of blood velocities using ultrasound. A signal processing approach" [132], which was my main source of information on blood-velocity estimation.

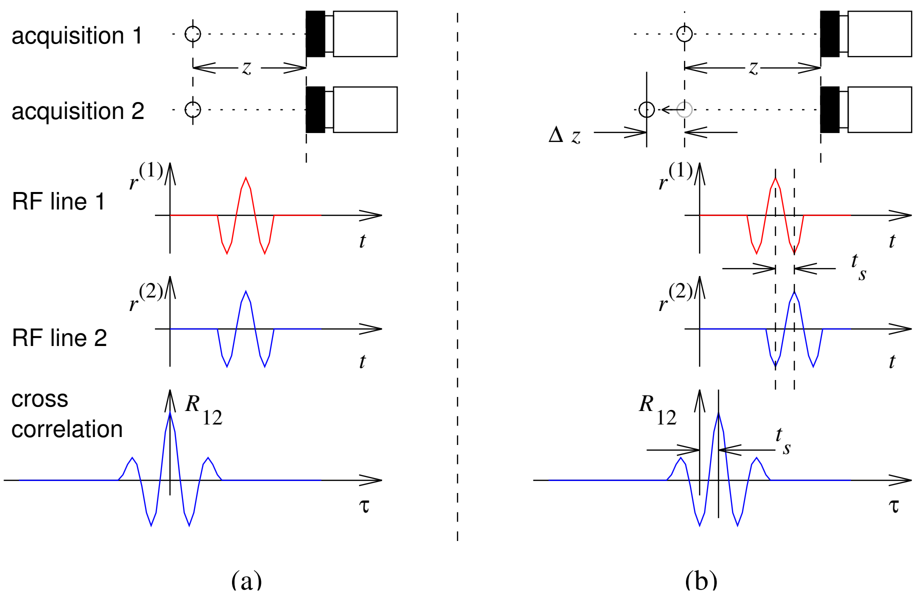


Figure 10.1: Principle of velocity estimation using cross-correlation.

a time-delayed and scaled version of the transmitted pulse $p(t)$:

$$r(t) = p\left(t - \frac{2z}{c}\right), \tag{10.1}$$

where z is the distance to the point scatterer, c is the speed of sound, and t is time relative to the start of the emission. If the scatterer is not moving, then two consecutive RF lines $r^{(1)}(t)$ and $r^{(2)}(t)$ will be identical:

$$r^{(1)}(t) = r^{(2)}(t) = p(t - \frac{2z}{c}). \tag{10.2}$$

The cross correlation of these two consecutively acquired lines yields[2]:

$$\begin{aligned} R_{12}(\tau) &= \frac{1}{2T}\int_T r^{(1)}(t) r^{(2)}(t+\tau)\, dt \\ &= \frac{1}{2T}\int_T p(t - \frac{2z}{c}) p(t - \frac{2z}{c} + \tau)\, dt \\ &= R_{pp}(\tau). \end{aligned} \tag{10.3}$$

Because $p(t)$ is of finite duration the function R_{pp} has a global maximum at $\tau = 0$, as shown at the bottom of Figure 10.1(a).

Consider the case shown in Figure 10.1(b), in which the point scatterer moves at a certain

[2]The notation $\frac{1}{2T}\int_T \bullet\, dt$ is used instead of $\lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} \bullet\, dt$ to simplify the notation as in [132].

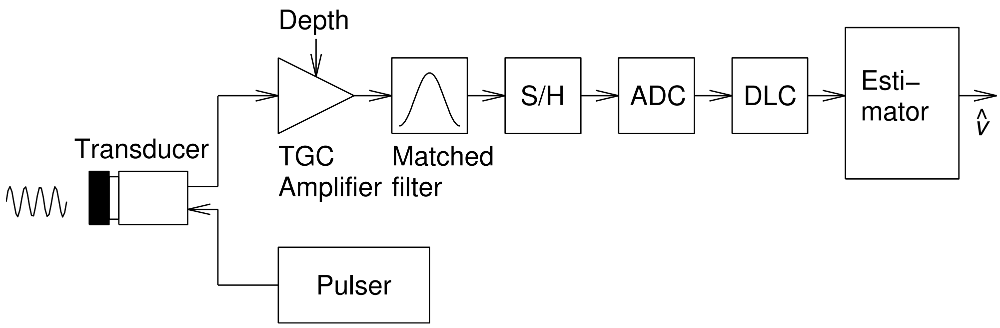


Figure 10.2: System measuring the velocity using cross-correlation.

distance Δz between the two emissions. The received signals then become:

$$r^{(1)}(t) = p\left(t - \frac{2z}{c}\right) \tag{10.4}$$

$$r^{(2)}(t) = p\left(t - \frac{2(z+\Delta z)}{c}\right) \tag{10.5}$$

$$\text{if } \ t_s = \frac{2\Delta z}{c} \ \text{ then} \tag{10.6}$$

$$r^{(2)}(t) = r^{(1)}(t - t_s). \tag{10.7}$$

Taking again the cross-correlation of the two RF lines, one gets:

$$\begin{aligned} R_{12}(\tau) &= \frac{1}{2T}\int_T r^{(1)}(t) r^{(2)}(t+\tau)\, dt \\ &= \frac{1}{2T}\int_T p(t - \frac{2z}{c}) p(t - \frac{2z}{c} - t_s + \tau)\, dt \\ &= R_{pp}(\tau - t_s), \end{aligned} \tag{10.8}$$

where t_s is the time shift due to the movement of the scatterer. $R_{pp}(\tau - t_s)$ has a global maximum at $\tau = t_s$ as shown at the bottom of Figure 10.1(b), and the velocity can be uniquely determined from R_{12}. Estimating the time shift $\hat{t}_s$ one can estimate the axial distance $\hat{\Delta z} = \hat{t}_s c/2$. This distance was traveled by the scatterer for the time between two emissions T_{prf} and the estimated axial velocity becomes:

$$\hat{v}_z = \frac{c}{2}\frac{\hat{t}_s}{T_{prf}}. \tag{10.9}$$

The same considerations can be applied for a group of scatterers as it was done by Foster and colleagues [134, 135] and by Jensen in Chapter 8 of [132]. The scatterers are modeled [139] as a band passed, white, random signal s_c with a mean value of zero, a variance σ_s^2, and a Gaussian amplitude distribution. The received signal can then be written as:

$$r(t) = \int_{-\infty}^{\infty} p(t - t') s_c(t')\, dt', \tag{10.10}$$

where $p(t)$ is now the pulse-echo impulse response and includes the electro-mechanical impulse response of the transducer, and effects of propagation in the tissue. After some math one

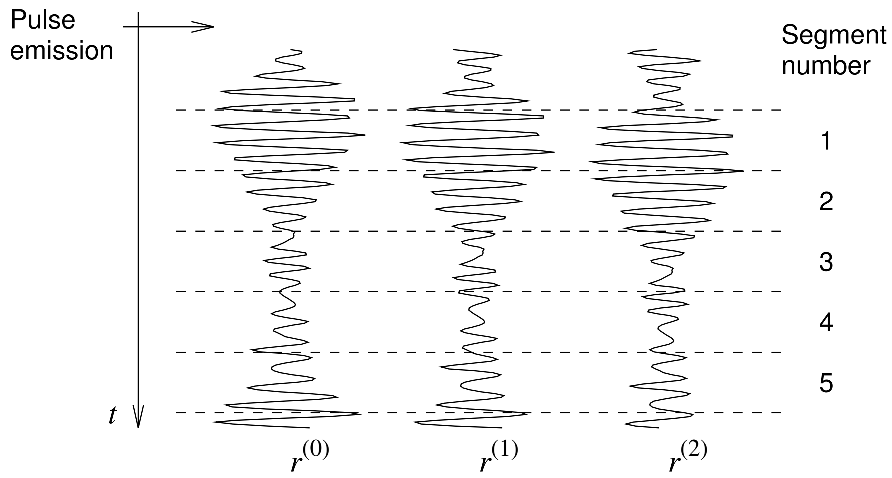


Figure 10.3: Segmentation of RF lines.

gets the following expression for the cross-correlation function $R_{12}(\tau)$:

$$R_{12}(\tau) = \sigma_s^2 R_{pp}(\tau - t_s). \tag{10.11}$$

The cross-correlation function R_{12} is scaled with the variance of the random signal $s_c(t)$ which models the scattering sequence.

A diagram of a cross-correlation system is shown in Figure 10.2. A transmitter transmits a short duration pulse. The received echo is amplified by a TGC amplifier which compensates for the attenuation. The matched filter reduces the noise from the received signal, and ensures that there is no aliasing in the sampling process. The RF signals are sampled by the ADC. The stationary echo is canceled by subtracting the consecutive lines. Finally the velocity is estimated using the estimation procedure described in the next section.

10.2 Calculation of the crosss-correlation

The velocity is estimated from sampled data, and the formulae for the calculation of the cross-correlation are discrete. Every consecutive RF line $r^{(n)}[k]$ obtained from the same direction is divided into segments as shown in Figure 10.3. This segmentation is necessary because the velocity varies across the vessel. From each segment an estimate of the velocity is obtained and mapped onto the screen. The segments shown in the figure are neighboring, but in general they can overlap.

The cross correlation of each pair of segments located at the same depth in two consecutive lines is calculated:

$$\hat{R}_{12}[n, i_{seg}] = \frac{1}{N_s} \sum_{k=0}^{N_s-1} r^{(1)}[k + i_{seg}N_s] r^{(2)}[k + i_{seg}N_s + n], \tag{10.12}$$

where N_s is the number of samples in a segment, and i_{seg} is the segment number. Because in reality two adjacent RF lines are not time shifted versions of each other, the so-estimated cross-correlation function $\hat{R}_{12}$ is not a shifted version of the auto-correlation function $\hat{R}_{11}$ as was assumed in the previous section. The cross-correlation estimate can be improved by averaging

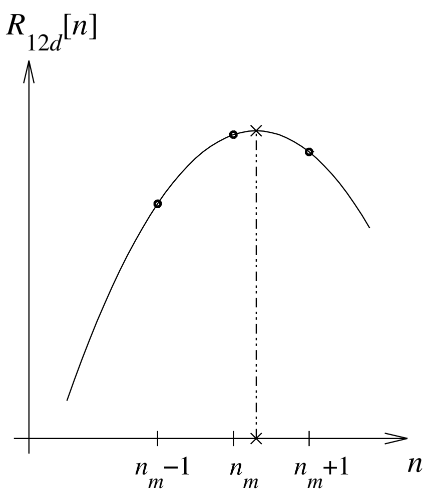


Figure 10.4: Graphical representation showing that the three points near the maximum can be approximated by a parabola.

over several estimates of $\hat{R}_{12}$. The velocity is roughly constant over several lines, so that t_s stays the same. The averaging can be expressed as:

$$\hat{R}_{12d}[n, i_{seg}] = \frac{1}{N_s(N_c - 1)} \sum_{i=0}^{N_c-2} \sum_{k=0}^{N_s-1} r^{(i)}[k + i_{seg}N_s] r^{(i+1)}[k + i_{seg}N_s + n], \tag{10.13}$$

where N_c is the number of lines over which the averaging was done.

Since the only thing that matters is the location of the peak, not its value, the normalization factors are usually skipped. The span of $\hat{R}_{12d}$ determines the maximum detectable velocity. The search range is usually less than the searchable interval $[-N_s, N_s]$. The largest detectable velocity is:

$$v_{max} = \frac{c}{2} N_s \frac{f_{prf}}{f_s}, \tag{10.14}$$

where f_s and f_{prf} are the sampling and pulse repetition frequencies, respectively.

The minimum detectable velocity is :

$$v_{min} = \frac{c}{2} \frac{f_{prf}}{f_s}. \tag{10.15}$$

The velocity resolution can be improved by fitting a second-order polynomial to the three points at the peak of the cross-correlation [134, 135], as shown in Figure 10.4. If the peak is found at a lag n_m, then the interpolated peak is found at:

$$n_{int} = n_m - \frac{\hat{R}_{12d}[n_m + 1] - \hat{R}_{12d}[n_m - 1]}{2\left(\hat{R}_{12d}[n_m + 1] - 2\hat{R}_{12d}[n_m] + \hat{R}_{12d}[n_m - 1]\right)} \tag{10.16}$$

The estimated velocity is then:

$$\boxed{\hat{v} = n_{int} \frac{c}{2} \frac{f_{prf}}{f_s}} \tag{10.17}$$

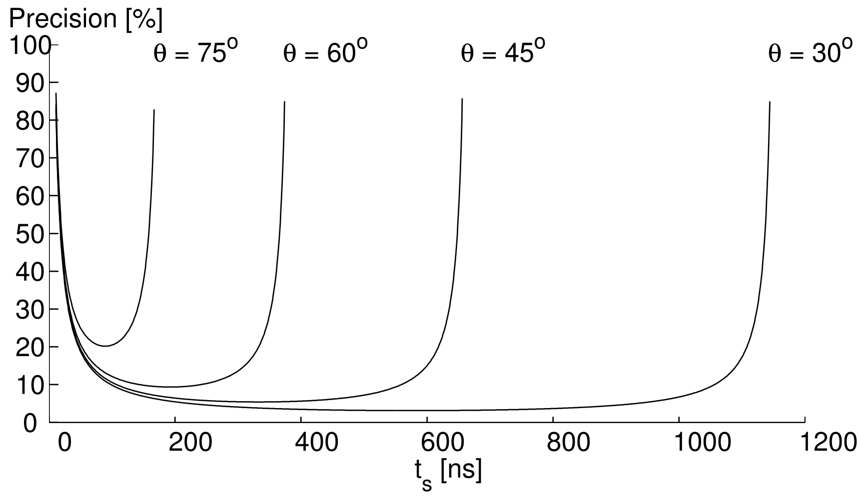


Figure 10.5: Theoretical precision versus time shift and measurement angle. The parameters are set as in [135]: SNR = 20 dB, RMS bandwidth = 2.5 MHz, and beam width = 0.6 mm.

10.3 Sources of errors

There are many sources of error to the velocity estimations based on the time-shift measurement. They have been previously modeled [134, 135, 140] and experimentally explored [141, 142]. In the following some of the errors will be outlined following the analysis made by Foster and Embree [135]. Their error analysis concludes:

> ...In summary, the theoretical precision of the time-domain correlation method depends on the measurement angle, the system RMS bandwidth, the mean time shift between echoes, and SNR.

The equation given by them for the relative precision of the estimate (Precision = $\sigma(\hat{t}_s)/t_s$) is:

$$\text{Precision}[\hat{t}_s] = \frac{1}{2\pi\Delta_f \rho \sqrt{\frac{E_s}{2N_0}}}, \tag{10.18}$$

where Δ_f is the RMS bandwidth (for a definition of RMS bandwidth see (9.25) on page 93), E_s is the energy in the signal, N_0 is the power density of the noise. ρ is the normalized correlation between two received signals that depends on the fact the scatterers move not only axially, but also laterally.

In [135] ρ is a linear function of the lateral distance traveled by a scatterer:

$$\rho = 1 - 1.2a \tag{10.19}$$

$$a = \frac{|\vec{v}|\sin\theta}{\delta_{x3\mathrm{dB}}} T_{prf}, \tag{10.20}$$

where θ is the angle between the ultrasound beam and the direction of motion, and $\delta_{x3\mathrm{dB}}$ is the 3 dB beam width. Figure 10.5 shows a plot of the theoretical relative precision versus the estimated delay. It can be seen that searching for small delays results in high-variance, because of the noise influence. If the velocity $\vec{v}$ is fixed, increasing the angle means that the motion is goes laterally, rather than axially, resulting again in high-variance of the estimate.

The errors associated with the correlation method are:

1. Errors associated with the windowing. As seen only range-gated segments are cross-correlated. A windowed waveform does not necessarily have its maximum correlation at the true shift.

2. The velocity gradient. The range cell contains scatterers moving at different velocities. Their relative positions change and so does the received signal, thus decreasing the correlation.

3. The system impulse response. The performance of the estimator is better at shorter pulses. A long impulse response (narrow band system) results in a higher uncertainty of the estimates.

4. Error associated with intensity profile across the sound beam (beam-width modulation). The lateral motion of the scatterers results in echoes with different amplitude, thus decreasing the correlation between consecutive echoes from the same spatial position.

There are different situations, in which each of the enumerated error sources dominate. For a good SNR, the error associated with the beam profile exceeds significantly the errors from the correlation effects. This error is dominating for the case when all the scatterers move with the same velocity. In the case of a velocity profile (for example parabolic flow) the error introduced by the velocity gradient is the most significant one.

The SNR is an essential source of error. The stationary echo canceling decreases this ratio even more as shown by Jensen [143]. As shown in Section 9.1.3, the signal-to-noise ratio can be increased by using a frequency modulated signal (chirp).

In the following the use of chirps for velocity estimations will be discussed.

10.4 Using frequency modulated signals for cross-correlation velocity estimation.

The use of frequency modulated signals for velocity estimation and the associated specifics were investigated by Wilhjelm in his Ph.D.thesis [144]. Some of the work was also published in a couple of papers [145, 146]. In these publications the interaction between the linear FM chirp and the blood scatterers is described and it is demonstrated that the velocity can be estimated. In this work a rather traditional approach is used: (1) the returned echoes are compressed by matched filtering as described in Section 9.1.3; (2) using the cross-correlation velocity estimator the *difference of the shifts of the peaks* of the matched filtered signal is determined as shown in Figure 10.6.

Figure 10.6 shows again the simplified case depicted in Figure 10.1, but this time using linear FM chirps. Two acquisitions are made. For the time T_{prf} the scatterer increases its distance to the transducer with Δz meters. The velocity is denoted with v. The time instances of the first

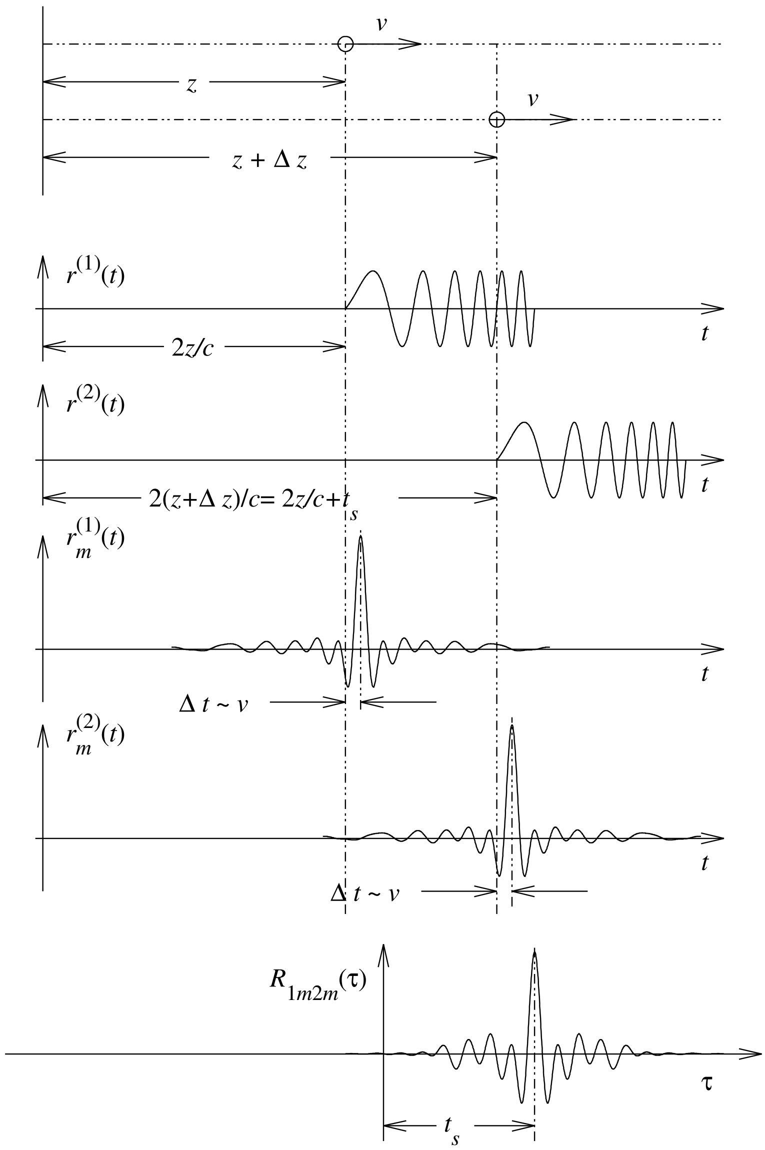


Figure 10.6: Principle of velocity estimation using cross-correlation and FM chirps.

received echo from the *actual position*[3] of the scatterer are $t_1 = 2z/c$ and $t_2 = 2(z+\Delta z)/c = t_1 + t_s$ for the first and second received signals, $r^{(1)}(t)$ and $r^{(2)}(t)$, respectively. The time shift due to the motion is $t_s = 2\Delta z/c$. The signals $r_m^{(1)}(t)$ and $r_m^{(2)}(t)$ are obtained by filtering the signals $r^{(1)}(t)$ and $r^{(2)}(t)$ with a filter which is matched to the transmitted signal. The velocity is estimated by the offset t_s of the peak of the cross-correlation between $r_m^{(1)}(t)$ and $r_m^{(2)}(t)$.

During their interaction with the scatterer, the chirp signal experiences a Doppler shift, and a change of the slope β (for the definition of β see (9.29) on page 94), with which it sweeps from the start frequency f_1 to the end frequency f_2. The slope β of the received signal is proportional to $\left(\frac{c+v}{c-v}\right)^2 \beta_0$ [144], where β_0 is the sweep slope of the transmitted pulse. This change leads to an increased duration of the main-lobe of the matched-filtered responses $r_m^{(1)}(t)$ and $r_m^{(2)}(t)$.

[3] In some papers it is shown that the the echo from a moving scatterer arrives at a time instance $t = \frac{2z}{c+v}$ as opposed to the echo of a stationary scatterer $t = \frac{2z}{c}$. For the sake of simplicity we assume that at the moment of interaction the scatterer *is* at depth z.

Assuming $c = 1540$, and $v = 1$ m/s, one gets that $\beta = 1.0026\beta_0$, and this effect has a negligible influence on $r_m(t)$.

The Doppler shift, experienced by the chirp leads to an offset of the peak of the matched filtered response with Δt (see the relation between the frequency and time shifts in Figure 9.8 (right)). Assuming constant velocity over the period of observation, the transient effects are the same for both $r_m^{(1)}(t)$ and $r_m^{(2)}(t)$. Hence, the cross-correlation estimator will pick only the time shift t_s which is due to the motion of the scatterer.

The conclusion is that the same procedure for estimation of velocity can be used on the matched filtered responses from "normal" RF pulse and a linear FM chirp. The use of a linear FM chirp has the advantage of higher signal-to-noise ratio.

CHAPTER

ELEVEN

Effects of motion

The decorrelation of the RF signals and low signal-to-noise ratio are major causes of the uncertainty in the velocity estimations. In the conventional ultrasound systems the decorrelation is due mainly to the velocity gradient. In synthetic aperture ultrasound imaging another source is added: the motion artifacts. To successfully make velocity estimations, one can either compensate for them as shown in Chapter 12, or circumvent their influence as done in Chapter 13. Both approaches require knowledge of the essence of the motion artifacts.

This chapter tries to develop an intuitive feeling of the nature of motion artifacts, based on simple geometric transforms. Similar considerations using the frequency representation of the system (*k*-space [54, 55]) have been previously done for artifacts due to tissue shearing [147] and for decorrelation due to flow gradients [148]. Following the path paved by these papers, the *k*-space is used in the following sections to create a simple model of the relation between the point spread functions of the low-resolution images and the final high resolution one. This model is then extended to incorporate the artifacts from motion on the low-resolution images and on the high-resolution one. The frequency domain representation is especially suitable for phased array sector images, and only these are considered in the following sections.

11.1 Modeling the low resolution image

In this section an attempt is made to derive a simple model of the point spread function of the low-resolution images. To test the model an analytical approximation using separable function will be created. This model will be compared with a simulated point spread function using the program Field II.

In the following, the term *point spread function* (PSF) will be used to denote a **2D image** of a point lying in the azimuth plane. The lateral coordinate x sometimes will be interchanged with θ, which is the azimuth angle, and is a natural coordinate base for the sector phased array scans.

The pulse echo response of a single scatterer (see Sections 3.3.2 and 3.4.2) is given by [24]:

$$p_r(\vec{x},t) = v_{pe}(t) \underset{t}{*} \delta(\vec{x}-\vec{x}_1) \underset{x}{*} h_{pe}(\vec{x},t), \tag{11.1}$$

where $\vec{x}_1$ is the spatial position of the point, $\delta(\vec{x})$ is the Dirac delta function in space, $h_{pe}(\vec{x},t)$ is the spatial impulse response in space, and $v_{pe}(t)$ includes the transducer excitation and the impulse responses of the transmitting and the receiving apertures. Such a response is given in Figure 11.1. The response is normalized. The top left graph gives the lateral slice of the

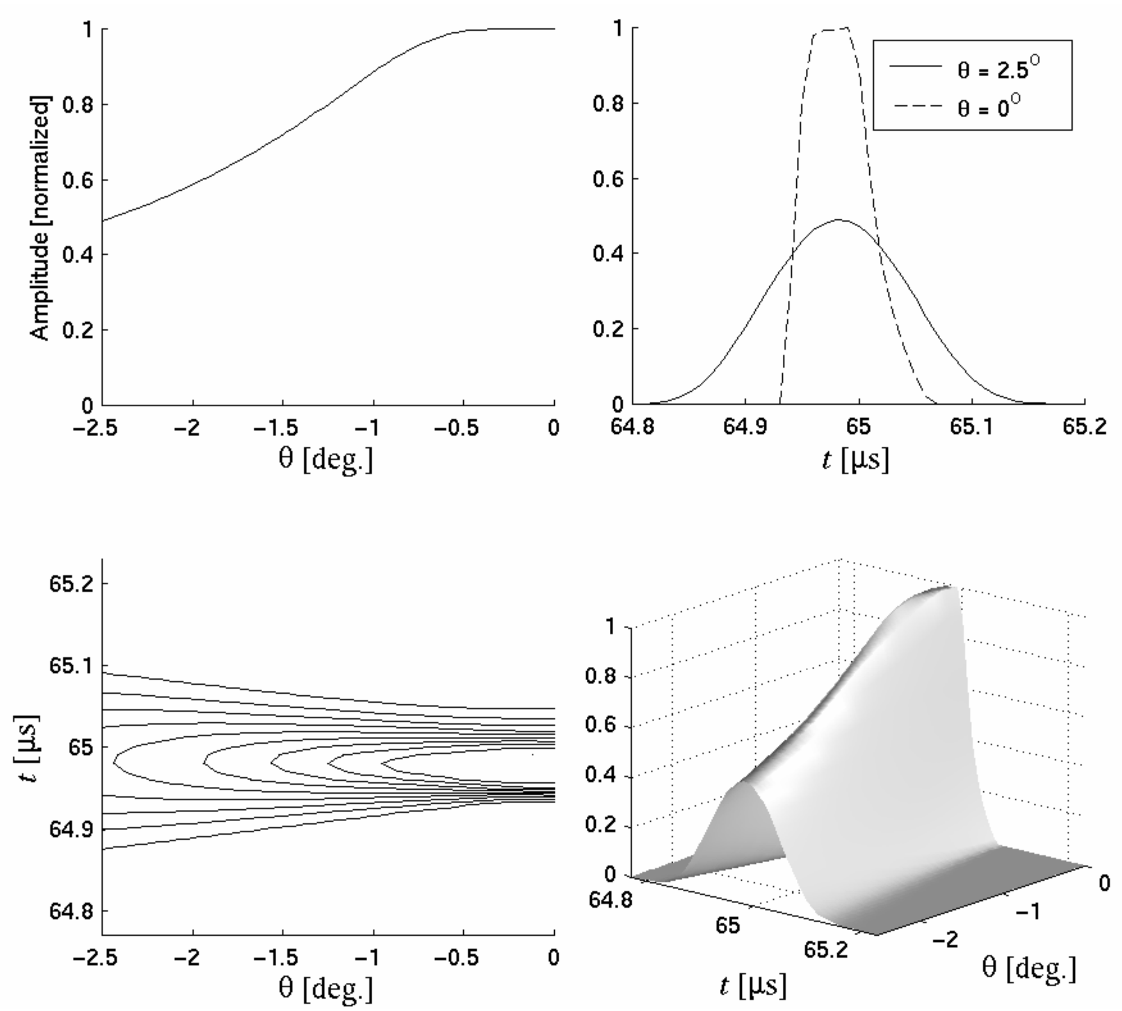


Figure 11.1: Simulated pulse echo field. The number of elements in transmit and receive are $N_{xmt} = 1$ and $N_{rcv} = 64$, respectively.

response at the maximum. The top right plot shows the response in time along two directions defined by the azimuth angles $\theta = 0°$ and $\theta = 2.5°$ respectively. The bottom left plot gives a view from "above". The contours are taken at levels from 0.1 to 1 at steps of 0.1. The bottom right plot represents the 3D graph of the response. The simulations are done for a linear array transducer with a center frequency $f_0 = 5$ MHz. The speed of sound is assumed to be $c = 1540$ m/s. The pitch of the transducer is $d_x = \lambda/2$, and the width of an element $w = 0.95d_x$. The number of transmitting and receiving elements is $N_{xmt} = 1$ and $N_{rcv} = 64$, respectively. For the plot $v_{pe}(t)$ is assumed to be $v_{pe}(t) = \delta(t)$. At this moment the following approximation is introduced:

$$h_{pe}(\theta,t) = h_{pe1}(\theta) \cdot h_{pe2}(t), \tag{11.2}$$

meaning that the spatial impulse response will be approximated with a separable function. It is expressed as the product of the two function: $h_{pe1}(\theta)$, given in top left plot of Figure 11.1, and $h_{pe2}(t)$ given in the top right plot of the same figure. From the figure it can be seen that with increasing angle, the response in time $h_{pe}(t)$ becomes longer. In the approximation $h_{pe2}(t)$ will be the response along the central line ($\theta = 0$). This introduces an error at the outskirts of the approximated PSF. In order for the reader to get a better feeling of this error Figure 11.2 shows

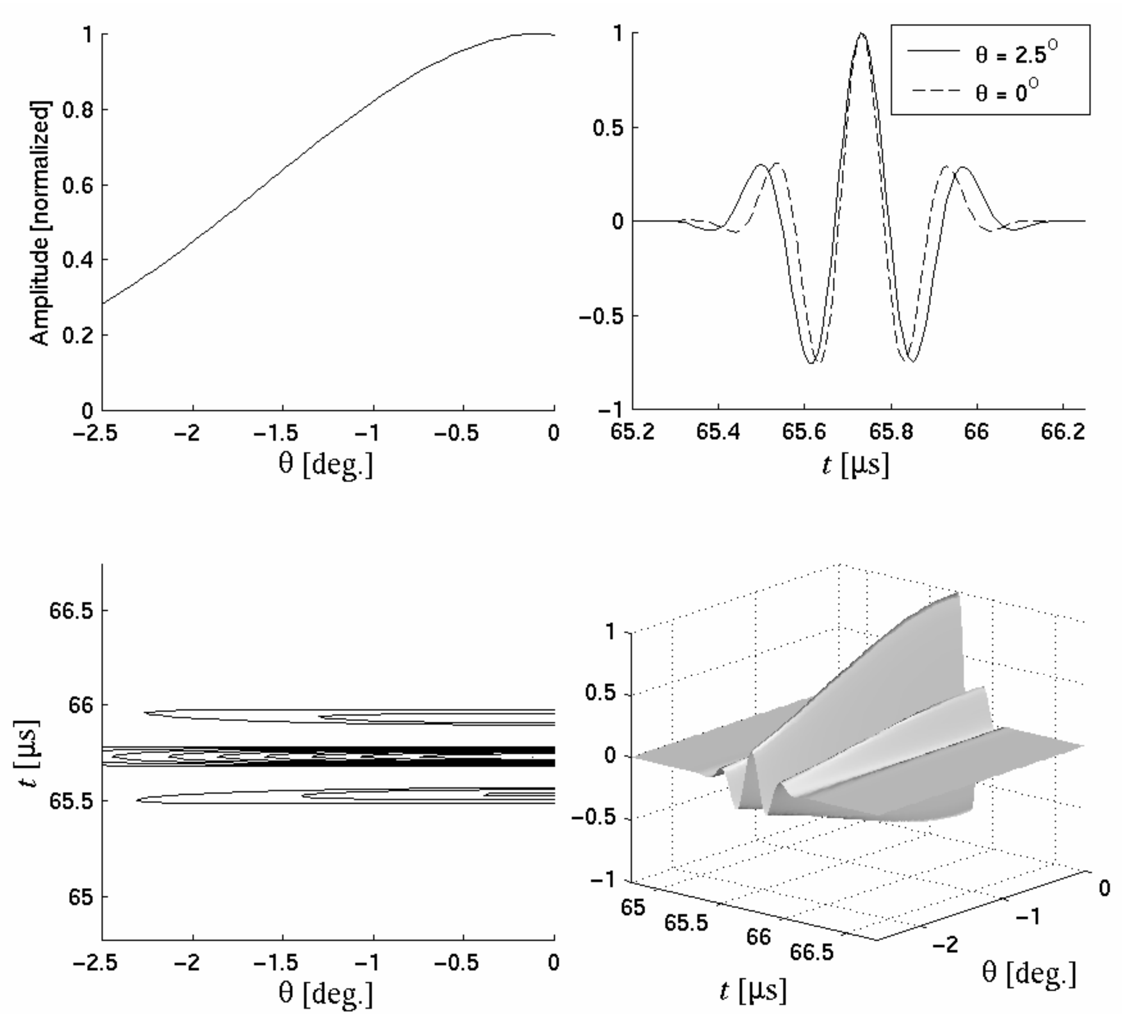


Figure 11.2: The pulse echo response of a single point for a transducer with 60 % fractional bandwidth. The number of transmitting and receiving elements is $N_{xmt} = 1$ and $N_{rcv} = 64$, respectively.

a more realistic situation.

The figure illustrates the case when the excitation $g(t)$ is a 3 cycles Hamming windowed sinusoid, and the two-way impulse response of the system $h_{xmt} \underset{t}{*} h_{rcv}$ is a 60 % pulse with Gaussian envelope. Both, the excitation, and the impulse response have a center frequency of $f_0 = 5$ MHz. The difference between Figures 11.1 and 11.2 is that in the top right plot of the latter, $h_{pe}(t)$ is normalized to the maximum in each direction. It can be seen that the errors will be exhibited at lower signal levels. The efforts in the rest of the section will be directed towards finding suitable expressions for h_{pe1} and h_{pe2}.

As shown in Appendix A the lateral radiation pattern $\mathsf{A}_t(x_1; z_f)$ of a focused transducer is the spatial Fourier transform of the apodization function $a_t(x)$. Both, the apodization function and the radiation pattern are defined on confocal spherical surfaces, which for the 2D case are represented by confocal arcs as shown in Figure 11.3(a). The top arc represents the transducer surface and the shaded region symbolizes the apodization function, which in the depicted case is a rectangular window. The dotted lines connect the edges of the transducer with the focal point, which lies at coordinates $x = 0$ and $z = z_f$. The radiation pattern is calculated along an

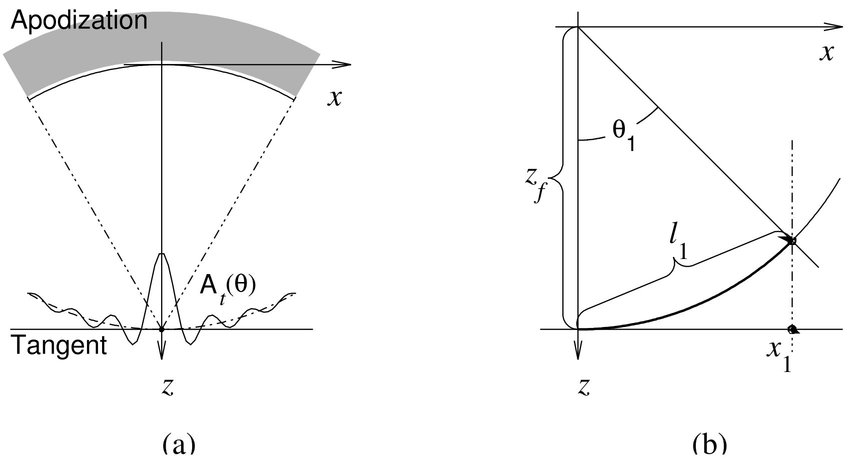


Figure 11.3: Illustration of the Fourier relation between the apodization function of a focused transducer and the line spread function.Sub-figures (a) and (b) show the imaging situation and the geometric relation between the coordinates, respectively.

arc which passes through the focal point. The tangent to the arc is parallel to the transducer surface. A_t lies on the arc, and has for an argument the length of the arc to the focal point l_1. This arc is shown with a thick line in Figure 11.3(b). The approximations are valid for small angles θ, and for these angles:

$$\theta_1 = \frac{l_1}{z_f}$$
$$x_1 = z_f \sin\theta$$
$$\sin\theta \approx \theta$$
$$\Downarrow$$
$$l_1 \approx x_1$$

where x and z are the lateral and axial distances defined in a Cartesian coordinate system. These relations are illustrated in Figure 11.3(b). The point-spread-function can be expressed either as a function of $k\sin\theta$ or equivalently in kx/z, where $k = 2\pi/\lambda$ is the wavenumber, $\lambda = c/f$ is the wavelength at frequency f, and c is the speed of sound. The frequencies f are within the bandpass region of the electro-mechanical impulse-response of the transducer. The Fourier relation between the radiation pattern and the apodization function of the transducer must be calculated over the whole frequency range and the result integrated.

The two dimensional radiated pressure field $p_t(x,t;z_f)$ for a fixed depth z_f as a separable function is given by:

$$p_t(x,t;z_f) = \mathsf{A}_t(x;z_f)p(t), \tag{11.3}$$

where $p(t)$ is the transmitted pressure pulse. It includes the excitation $g(t)$, the electro mechanical impulse response of the transducer $h_{xmt}(t)$, and $h_{pe2}(t)$:

$$p(t) = g(t) \underset{t}{*} h_{xmt}(t) \underset{t}{*} h_{pe2}(t). \tag{11.4}$$

The function $\mathsf{A}_t(x;z_f)$ is the radiation pattern which is given by the Fresnel diffraction formula (see Appendix A) for the focal point:

$$\mathsf{A}_t(x;z_f) = \int_f \mathcal{F}^{-1}\left\{a_t\left(\frac{x}{z_f c/f};0\right)\right\}\, df. \tag{11.5}$$

The integration over f will be skipped further, and A_t will be calculated only for the center frequency f_0.

Taking the two dimensional Fourier transform of $p_t(x,t;z_f)$ one gets the k-space representation of the imaging system:

$$P_t(f_x, f; z_f) = \mathcal{F}\left\{p_t(x,t;z_f)\right\}. \tag{11.6}$$

Since $p_t(x,t;z_f)$ is assumed to be a separable function, then the spectrum is also a separable function:

$$P_t(f_x, f; z_f) = \mathcal{F}\{\mathsf{A}_t(x)\} \cdot \mathcal{F}\{p(t)\} \tag{11.7}$$

$$P_t(f_x, f; z_f) = a_t\left(\frac{x}{z_f\lambda}\right) \cdot P(f) \tag{11.8}$$

If a system works in a pulse-echo mode, then the two-way radiation pattern of the system is the product of the radiation/directivity patterns of the transmit and receive apertures:

$$\mathsf{A}_{t/r}(x;z_f) = \mathsf{A}_t(x;z_f) \cdot \mathsf{A}_r(x;z_f) \tag{11.9}$$

The pulse must pass through the transducers and the electronics of both systems. The temporal component of h_{pe} will be assumed to be a delta function for simplicity, and the pulse echo response at depth z_f, in the vicinity of a point centered below the transducer becomes:

$$p_r(x,t;z_f) = \mathsf{A}_t(x;z_f) \cdot \mathsf{A}_r(x;z_f) \cdot v_{pe}(t). \tag{11.10}$$

In order to distinguish further in the thesis the approximated point spread function from the "real one", an alternative notation will be used. The small bold $\mathbf{p}(x,t;z_f)$ will be used for the point spread function of a low-resolution image. The capital bold $\mathbf{P}(x,t;z_f)$ will be used to denote the point spread function of a high-resolution image.

Figure 11.4(a) shows the formation of $\mathsf{A}_t(x;z_f)$. The transmit $a_t(x)$ and receive $a_r(x)$ apodization functions are convolved giving the apodization of the effective aperture $a_{t/r}(x)$. The emission is done by a single element, which is assumed to be omnidirectional, and can be represented by a delta function. The effective aperture is therefore equal to the receive aperture. Then the argument is scaled to make it suitable for inverse Fourier transform: $f_x = x/(\lambda z)$. The two way radiation pattern is found by the inverse Fourier transform:

$$\mathsf{A}_{t/r}(x;z_f) = \mathcal{F}^{-1}\{a_{t/r}(f_x)\} \tag{11.11}$$

Figure 11.4(b) shows the generation of the temporal component of the point spread function. The transmitted pulse $g(t)$ is convolved with the impulse response of the transducer twice - once in transmit and once in receive.

Figure 11.4(c) shows a comparison between a Field II simulation and the approach just described. The contours are taken on the normalized RF lines at levels from -0.9 to 0.9 from the maximum value at steps of 0.2. The time is given relative to the time instance at which the peak value occurs. The same approach of the display of the point-spread-function will be used in the rest of the chapter. Table 11.1 lists the relevant parameters for the Field II simulation. The apodization function in receive is defined as:

$$a_r\left(x = \left(i - \tfrac{N_{xdc}+1}{2}\right) \cdot d_x\right) = \exp\left(-\frac{1}{2}\frac{\left(\frac{i-(N_{xdc}+1)/2}{N_{xdc}-1}\right)^2}{\sigma_x^2}\right), \quad \text{for} \quad 1 \le i \le N_{xdc}, \tag{11.12}$$

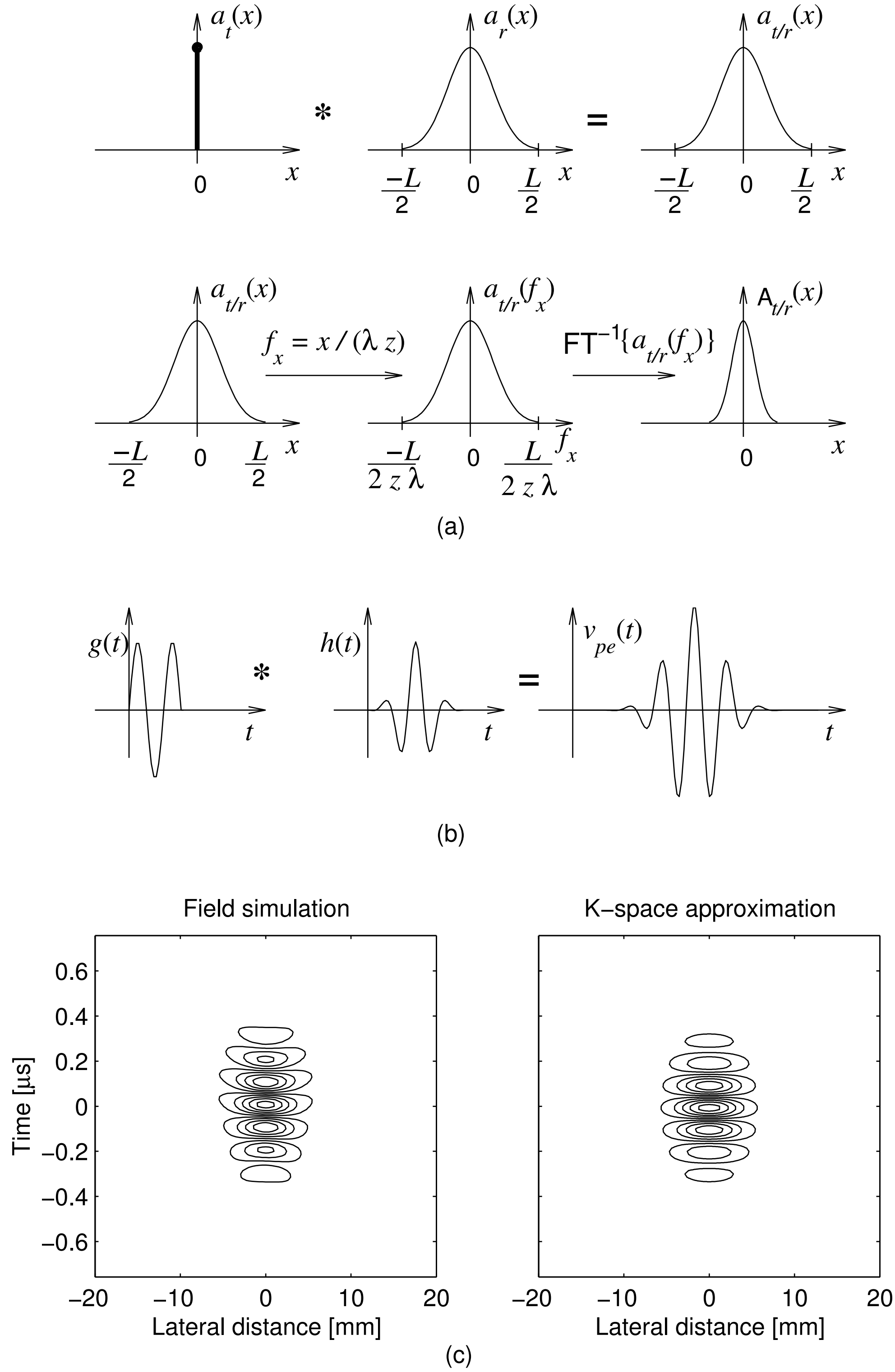


Figure 11.4: Step-by-step building of the approximation of the point-spread-function of a low-resolution image. The bottom plots show a comparison between the *k*-space approximation and a Field II simulation.

Parameter	Notation	value	unit
Speed of sound	c	1540	m/s
Center frequency	f_0	5	MHz
Sampling frequency	f_s	70	MHz
Wavelength	λ_0	308	μm
Fractional bandwidth	BW	66	%
No elements	N_{xdc}	65	-
Pitch	d_x	308	μm

Table 11.1: Simulation parameters used for the comparison in Figure 11.4.

where i is the element index and σ_x defines the -6 dB fall of the curve. In the particular example σ_x is equal to 0.3 guaranteeing that the apodization curve almost reaches zero at the edges of the transducer as shown in Figure 11.4(a). The advantage of using an exponential apodization function (popularly known as a "Gaussian") is that its Fourier transform is also an exponential function, and is easy to derive.

The transmitted pulse is a 1.5 cycles RF pulse at the central frequency f_0 as shown in Figure 11.4(b). The impulse response is modeled using a 3 cycles RF pulse at the central frequency with a Gaussian envelope. Comparing the numerical results shows reasonably good correspondence. The duration of the point-spread-function is $\approx 0.8\ \mu$s. Substituting $z = 0.8 \cdot 1540/2$ gives a length of 0.6 mm. For this small duration, one can assume that there is no significant change of the size of the point spread function, and by freezing the time, one can get a 2-dimensional spatial point spread function in the vicinity of z_f:

$$\mathbf{p}(x,z;z_f)\big|_{z=ct/2} = \mathbf{p}(x,t;z_f). \tag{11.13}$$

In the rest of the chapter these two variables, z and t are used interchangeably using the relation between them $z = ct/2$.

The point-spread-function of the low-resolution image obtained by emitting with the central element of the receive aperture will be used as a reference and denoted with $\mathbf{p}_0(x,t;z_f)$. The relation between this reference point-spread-function and the point-spread-functions of the low-resolution images obtained by emitting with other elements of the aperture will be derived in the following.

Figure 11.5 (a) shows by the means of simple geometry relations that the rotation of a focused transducer round the focal point results in the rotation of the point-spread-function. The solid and the dashed lines show the original and the rotated imaging situations, respectively. The rotation angle between the two point-spread-functions can be determined by the angle between the two tangent lines, which is equal to the angle of rotation of the whole setup. The rotation is with respect to a local coordinate system centered at the focal point. The same effect can be observed if a linear array (Figure 11.5 (b)) is translated at a distance Δx, while focusing at the same point. Notice that the reference point for calculating the delays must be the same for both positions (before and after the translation) of the linear array. The operation can be expressed as:

$$\mathbf{p}'(x,z;z_f) = \mathcal{R}[\beta;\vec{x_f}]\{\mathbf{p}(x,z;z_f)\} \tag{11.14}$$

where $\mathcal{R}[\beta;\vec{x_f}]$ means "rotation with β around the coordinates $\vec{x_f} = (x_f, y_f, z_f)$". In this case $x_f = y_f = 0$, and therefore only the depth z_f is given for $\mathbf{p}$. The rotation is given in a simple

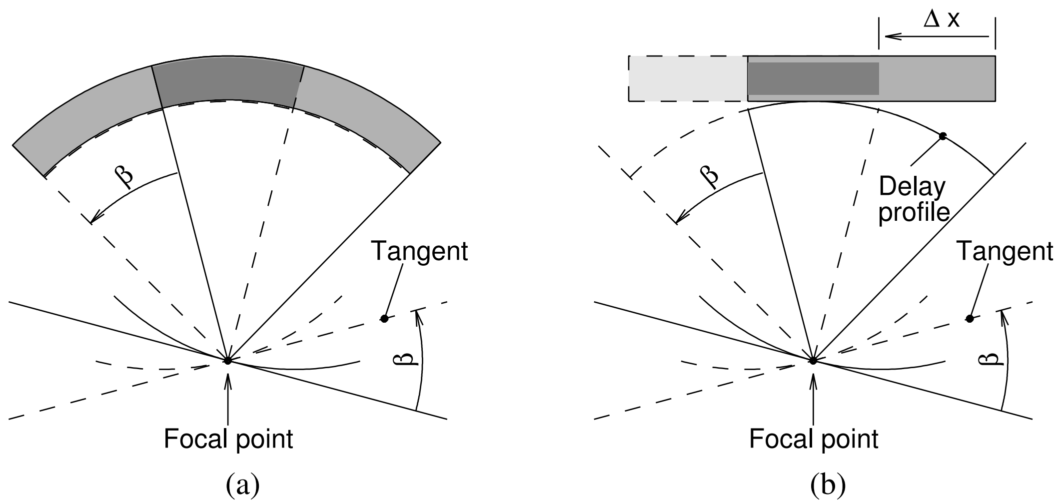


Figure 11.5: The rotation of a focused transducer round the focal point (a) and the translation of a linear array (b) result in rotation of the point-spread-functions.

matrix[1] form as:

$$\mathbf{p}(x', z'; z_f) = \mathbf{p}(x, z; z_f) \tag{11.15}$$

$$\begin{bmatrix} x' \\ z' \end{bmatrix} = \begin{bmatrix} \sin\beta & \cos\beta \\ \cos\beta & -\sin\beta \end{bmatrix} \begin{bmatrix} x - x_f \\ z - z_f \end{bmatrix} + \begin{bmatrix} x_f \\ z_f \end{bmatrix} \tag{11.16}$$

where the angle β is determined by:

$$\beta = \arctan\frac{\Delta x}{z_f}. \tag{11.17}$$

The rotation of the point spread function is caused by the translation of the transducer, respectively the effective aperture. This rotation results in rotation of the k-space representation of the imaging system (this can be shown by the properties of the 2-D Fourier transform [131]). The lateral extent of the k-space of the system is determined by the convolution of the transmitting and receiving apertures. Changing the position of either one results in a rotation of the point spread function. Figure 11.6 shows several equivalent transmit-receive cases. The scan lines along which the focusing is done are shown with dots, and their origin coincides with the origin of the coordinate system. The region in light gray outlines the position of the receive aperture and the region in dark gray represents the transmit element. The purpose of the figure is to show through simulations that the k-space transforms (which are a mere approximation) actually hold. All of the results plotted were obtained using Field II. The simulation parameters are listed in Table 11.1. Figure 11.6(a) is obtained by transmitting with element $i = 64$ ($1 \leq i \leq 65$), and receiving with all the elements. The whole transducer is translated to the left with $\Delta x = -32d_x$. Figure 11.6(b) is obtained by transmitting with the left most element $i = 1$. The transducer this time is centered above the point scatterer, $\Delta x = 0$. The last plot, Figure 11.6(c) is obtained by transmitting with the central element $i = 33$, but the whole aperture is translated with $\Delta x = -16d_x$. Using the more formal approach of k-space , the above experiments can be

[1]Notice that the chosen coordinate system is left handed.

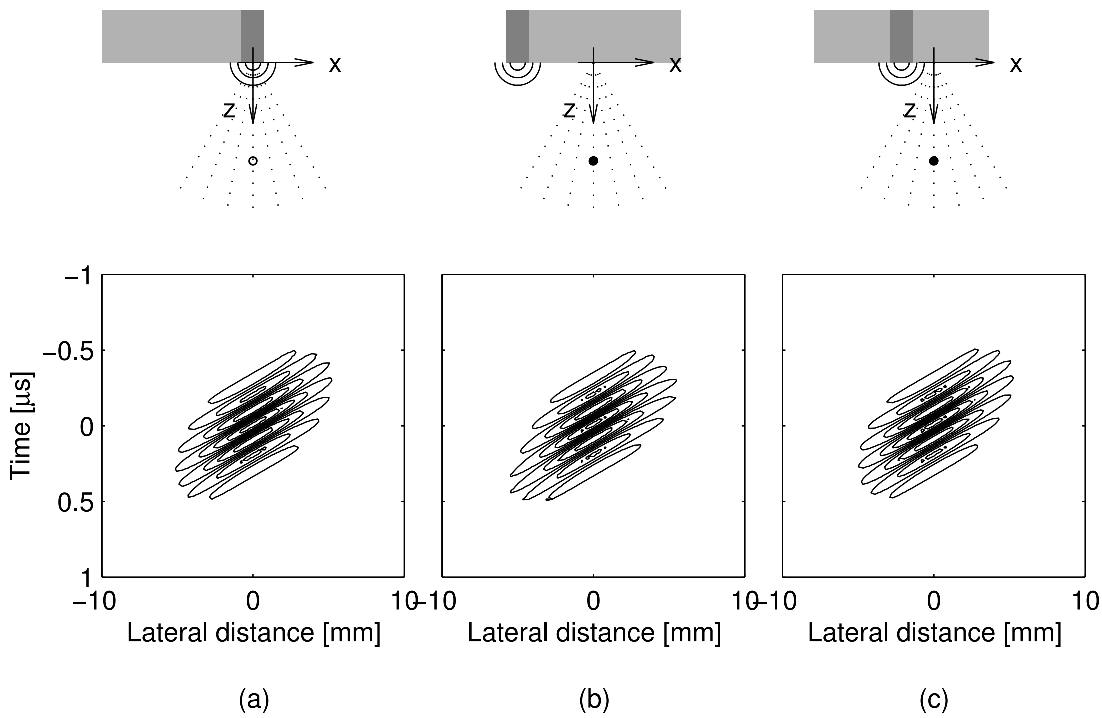


Figure 11.6: Illustration of several combinations of transmit and receive aperture locations resulting in the same inclination of the point-spread-function.

summarized by the properties of the convolution:

$$\begin{aligned} a_{t/r}(x-\Delta x) &= a_t(x) \underset{x}{*} a_r(x-\Delta x) \\ &= a_t(x-\Delta x) \underset{x}{*} a_r(x) \\ &= a_t(x-\Delta x/2) \underset{x}{*} a_r(x-\Delta x/2). \end{aligned} \tag{11.18}$$

11.2 Building a high-resolution image

In the previous section it was shown that translating both, the transmit and receive apertures at a distance Δx results in a rotation of the point-spread-function at an angle $\beta = \arctan(\Delta x / z_f)$, where z_f is the distance to the transducer (the focal point is assumed to lie on the z axis, i.e. $x_f = 0$). It was also shown that translating the transmit aperture at a distance Δx, without moving the receive aperture, gives a point-spread-function which could be obtained by translating both apertures at a distance $\Delta x/2$. If $\mathbf{p}_i(x,z;z_f)$ is the point-spread-function obtained by emitting with element i, $1 \le i \le N_{xdc}$, it can be approximated using the reference point-spread-function $\mathbf{p}_0(x,z;z_f)$ and a simple rotation:

$$\mathbf{p}_i(x,z;z_f) = \mathcal{R}[\beta_i; z_f]\{\mathbf{p}_0(x,z;z_f)\}, \tag{11.19}$$

where d_x is the pitch of the transducer. The angle of rotation β_i is related to the transmitting element i via:

$$\beta_i = \arctan \frac{\left(i - \frac{N_{xdc}+1}{2}\right) d_x}{2 z_f}, \quad \text{for } 1 \le i \le N_{xdc}. \tag{11.20}$$

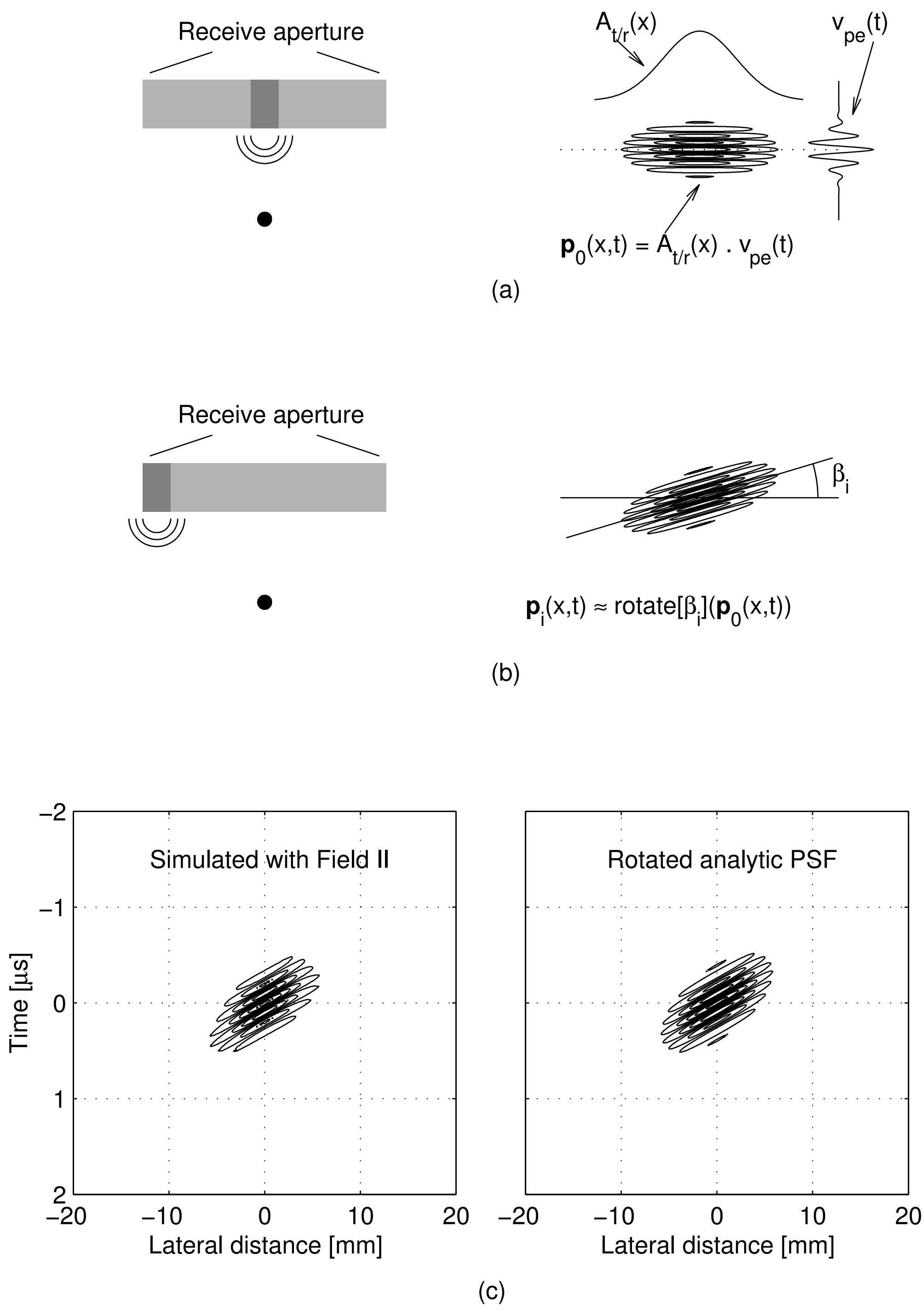


Figure 11.7: Comparison between the PSF obtained using Field II, and a analytically created and rotated PSF.

Figure 11.7 shows how the approximation of such a point spread function can be obtained: (1) $\mathbf{p}_0(x,z;z_f)$ is created using the Fourier relation between the aperture apodization function and the lateral shape of the point-spread-function, and (2) the reference point-spread-function is rotated at an angle β corresponding to the position of the transmitting element. Figure 11.7(c) shows a comparison between a Field II simulation and the analytically approximated point-spread-function. The simulation parameters are the same as the ones listed in Table 11.1 on page 118. The emission was done by the left-most transducer element.

The high-resolution image is obtained by summing the RF lines of the low-resolution ones. The resulting point-spread-function can then be expressed as:

$$\mathbf{P}(x,z;z_f) = \sum_{i=1}^{N_{xdc}} \mathbf{p}_i(x,z;z_f). \tag{11.21}$$

Using the relation between $\mathbf{p}_i$ and $\mathbf{p}_0$, this procedure can be expressed as a sum of rotated point-spread-functions:

$$\mathbf{P}(x,z;z_f) = \sum_{i=1}^{N_{xdc}} \mathcal{R}[\beta_i;z_f]\mathbf{p}_0(x,z;z_f) \tag{11.22}$$

So far all considerations were made for a focal point which is centered below the transducer (on-axis). They can be repeated for an off-axis point, and the fundamental result is the same. To make the considerations, one must use a coordinate system, whose z axis lies on the line connecting the point in question and the center of the transducer. The point-spread-function is slightly broader and asymmetric because of the decreased *apparent*[2] aperture which changes as a function of angle.

11.3 The motion artifacts

In order to continue the considerations, yet more assumptions will be made. Let's assume a maximum speed of scatterers v_{max} of 1 m/s and a pulse repetition frequency f_{prf} of 5000 Hz. The maximum distance traveled by the scatterer between two emissions then is 0.143 mm. The following considerations will be made for relatively large distances, such as 55 mm away from a transducer which is 11 mm wide ($f_\# = 5$). The assumption is that the point-spread-function of the low-resolution images does not change significantly for small changes in the axial and lateral position[3]:

$$\mathbf{p}_0(x,z;\vec{x}_f) = \mathbf{p}_0(x,z;\vec{x}_f + \Delta\vec{x}). \tag{11.23}$$

Because the distance traveled by the point scatterer is small relative to the distance between the scatterer and the transducer, (11.23) will be used to approximate the point-spread-function of the point at the new location. The point spread function at position $(x_f + \Delta x, z_f + \Delta z)$ will be merely a translated version of the point spread function at position (x_f, z_f):

$$\mathbf{p}_i(x,z;\vec{x}_f + \Delta\vec{x}) = \mathcal{T}[\Delta\vec{x}]\{\mathbf{p}_i(x,z;\vec{x}_f)\}. \tag{11.24}$$

$\mathcal{T}[\Delta\vec{x}]\{\mathbf{p}_i(x,z;\vec{x})\}$ is used to denote translation of the point-spread-function:

$$\mathbf{p}_i(x',z';\vec{x}_f + \Delta\vec{x}) = \mathbf{p}_i(x,z;\vec{x}_f) \tag{11.25}$$

$$\begin{bmatrix} x' \\ z' \end{bmatrix} = \begin{bmatrix} x \\ z \end{bmatrix} + \begin{bmatrix} \Delta x \\ \Delta z \end{bmatrix}. \tag{11.26}$$

Figure 11.8 shows the comparison between point-spread-functions of high-resolution images, obtained without and with motion. The top row displays the imaging situation. The emissions are done consecutively with the outermost elements, first with the leftmost and then with

[2]By apparent I mean the visible size of the aperture when looking from the focal point. Off axis the aperture is tilted with respect to the focal point. In some sources it is called "projected aperture".

[3]The considered distances are on the order of a milliliter.

the rightmost one. The positions of the scatterer and the point-spread-functions of the low-resolution images obtained at the first and the second emissions are given in blue and red colors, respectively. The middle row shows superimposed the low resolution images obtained at each of the emissions. The thick lines outline the envelope detected data at -1 and -20 dB. The thin and pale lines are contours of the raw RF signal, and are presented to show the direction and frequency of the oscillations. The distance which the point scatterer passes between the two positions is chosen to be $\lambda/4$ at a center frequency f_0 of 5 MHz. The point-spread-functions were obtained using Field II, and the simulation parameters are listed in Table 11.1. The bottom row shows contour plots of the envelope detected high-resolution images. The contours are drawn at levels -6, -10 and -20 dB from the peak value. It can be seen that the axial motion results in a distortion of the point-spread-function. The PSF becomes asymmetric and its maximum shifts away from the central line. The lateral motion introduces only minor change in the PSF. The reason for this is the anisotropy of the PSF - the lateral size is several times larger than the axial one. The process of creating a high-resolution image in the presence of motion becomes:

$$\mathbf{P}(x,z;\vec{x}_f) = \sum_{i=1}^{N_{xmt}} \mathcal{T}[i\Delta\vec{x}]\left\{\mathcal{R}[\beta_i;\vec{x}_f)]\{\mathbf{p}_0(x,z;\vec{x}_f)\}\right\}, \tag{11.27}$$

where β_i is the angle relate to the transmitting element i via the relation (11.20):

$$\beta_i = \arctan\frac{\left(i-\frac{N_{xdc}+1}{2}\right)d_x}{2z_f}, \quad \text{for } 1\leq i\leq N_{xdc}.$$

It can be shown that if the order of emissions gets reversed, then the peak of the point-spread-function of the high resolution image will be shifted in the opposite direction. The conclusion is that the motion artifacts depend not on the absolute motion of the scatterers, but on the relative motion between the scatterer and the change in position of the transmitting element. Creating high-resolution images with different transmit sequences results in different shapes of the point-spread-function and a decreased correlation between them. This fact will be used in Chapter 13 when modifying the cross-correlation velocity estimator.

The inverse operation of the translation is the translation in opposite direction. A motion compensated point-spread-function is given by:

$$\begin{aligned}\mathbf{P}_{\text{compensated}}(x,z;\vec{x}_f) &= \sum_{i}^{N_{xmt}} \mathcal{T}[-i\Delta\vec{x}]\left\{\mathcal{T}[i\Delta\vec{x}]\left\{\mathcal{R}[\beta_i;\vec{x}_f]\{\mathbf{p}_0(x,z;\vec{x}_f)\}\right\}\right\}\\ &= \sum_i \mathcal{R}[\beta_i;\vec{x}_f]\{\mathbf{p}_0(x,z;\vec{x}_f)\}\\ &= \mathbf{P}(x,z;\vec{x}_f).\end{aligned} \tag{11.28}$$

In practice such an ideal motion compensation is nearly impossible since the point-spread-function varies with the change in position. This simple relation means, however, that it is possible to obtain motion-compensated high-resolution images by shifting the samples in the *beamformed* RF lines of the low-resolution images, rather than changing the delay profiles on the individual channels. This approach will be tested in Chapter 12, in which the motion-compensated high-resolution images will be used for velocity estimation.

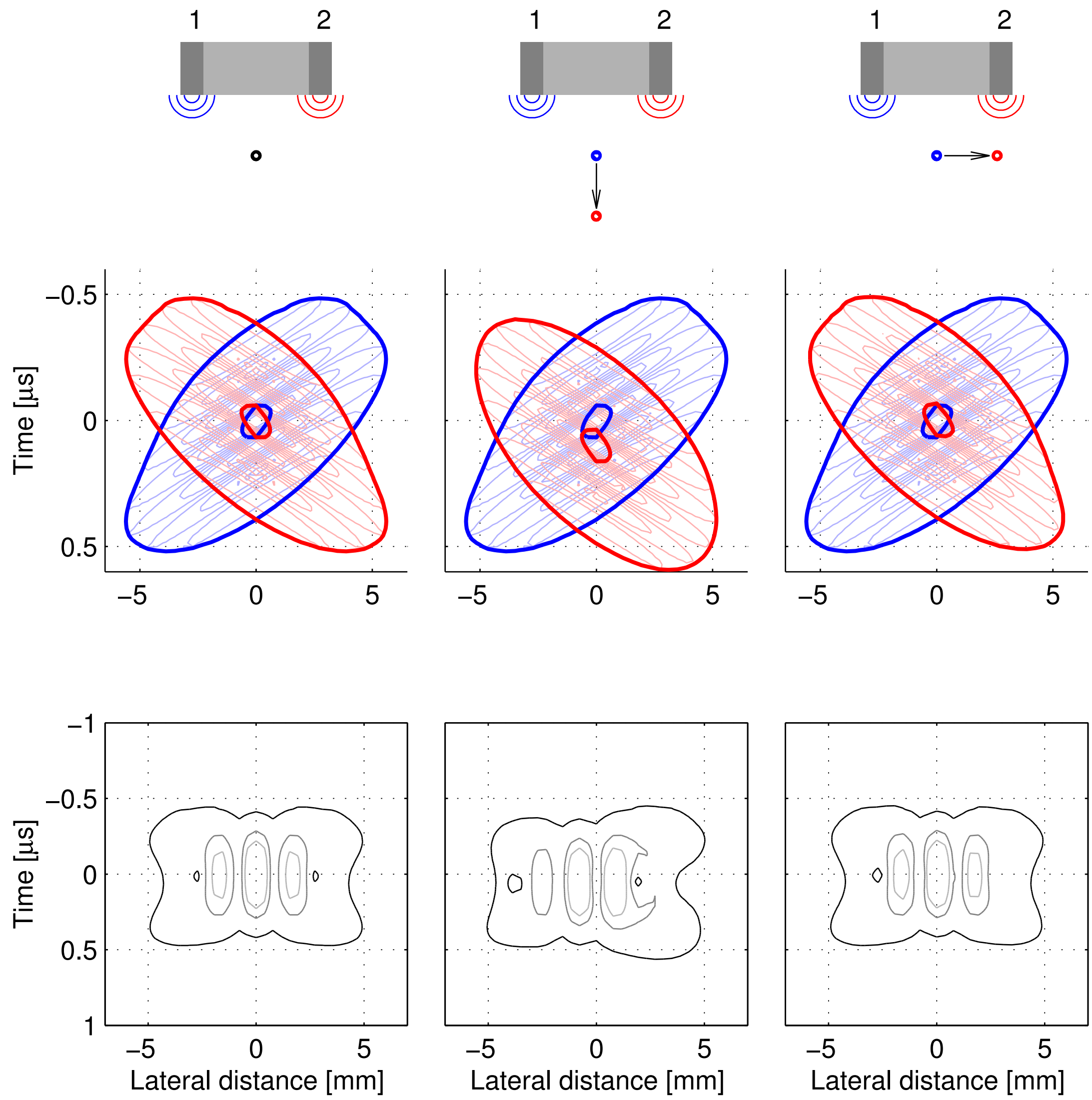


Figure 11.8: Making a high-resolution image with only two emissions. The emissions are done first with the leftmost element and then with the rightmost one. The left column shows the case when the point scatterer is still, the center column when the scatterer moves away from the transducer and right column when it moves parallelly to the transducer.

CHAPTER

TWELVE

Motion compensation

Synthetic transmit aperture images are acquired over a number of emissions. Tissue motion prohibits the coherent summation of the low resolution images formed at every emission. The latter results in a decrease of the contrast and resolution. The phases of the beamformed lines get distorted and the velocity of blood cannot be estimated using the traditional approaches.

This chapter deals with the motion compensation which must overcome the introduced artifacts. In the beginning an overview of the methods developed prior to this work is given. They involve either cross-correlation between the signals received by the individual channels, or cross-correlation between partially beamformed images.

Further a new method for motion compensation is suggested. It is based on motion estimation using the high resolution images. In the end the results from an experimental verification of the suggested motion compensation schemes are discussed.

12.1 Overview of the existing methods for motion compensation

One can consider the motion artifacts as caused by phase aberration, as was done by the group led by Prof. Trahey at Duke University. In several papers [69, 70, 71] they investigated the motion compensation applied to synthetic receive aperture ultrasound imaging (SRAU). They explored the possibility to correctly detect motion based on the received signals by several sub-apertures.

Figure 12.1 shows the measurement setup used in their work. The example is based on four receive sub-apertures. Four transmissions by all transducer elements are necessary in order to beamform a single RF line. Every time the reception is carried out by one of the four receive sub-apertures (illustrated in Figure 12.1 with different colors). The left and right sub-figures shows the cases when the sub-apertures are neighboring, and interleaved, respectively. The arcs above the transducers show the delayed echo from the point scatterer: the black arc shows the delays when there is no motion, and the colored arcs show the delayed echoes of the individual sub-apertures in the presence of motion. On the top of the figure the delay difference $\Delta\tau$ between the arrival time of the moving target and a stationary one is plotted. This difference must be added to the delay profile of the beamformer in order to sum the signals coherently. Between every two emissions the scatterers in the imaged direction move at a distance Δl. The distance traveled by a single scatterer between the first and the last emission is $l = 3\Delta l$. For abdominal imaging the maximum assumed velocity is 0.06 m/s [71] , resulting in a maximum delay error $3\Delta\tau = 50$ ns, which is a quarter of a period at 5 MHz, for a speed of sound $c = 1500$ m/s. The spatial distance between two elements with the same relative position

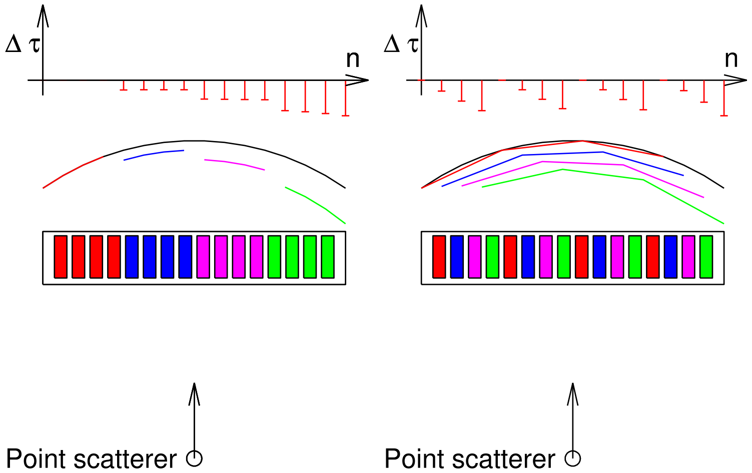


Figure 12.1: Delay errors in synthetic receive aperture imaging using 4 sub-apertures. The left sub-figure shows the case of neighboring sub-apertures and the right sub-figure shows interleaved sub-apertures.

in the sub-aperture is smaller for the interleaved sub-apertures. Because the elements are closer, the correlation between the received signals is higher and the time shift $\Delta\tau$ can be detected more reliably. Nock and Trahey [70] explain this phenomenon using the k-space representation of the system as shown in Figure 12.2. The figure shows the lateral k-space representation, where $k_x = k \sin\theta$ is the spatial frequency, θ is the steering angle, $k_{xs} = k\frac{D}{z}$, D is the lateral size of the whole aperture, and z is the imaging depth. The figure depicts the creation of a synthetic receive aperture of 32 elements by using two sub-apertures of 16 elements each. In the left sub-figure the sub-apertures are neighboring, and in the right sub-figure they are interleaved. The correlation between the beamformed signals from the two sub-apertures is proportional to the common area between the two k-space representations. Clearly the overlapping area is bigger when the sub-apertures are interleaved. The results in [70, 71] confirm these considerations.

A similar approach was undertaken by the group led by Karaman [2, 79, 149]. Their work is focused on synthetic transmit and receive aperture imaging. Every transmit is performed using a single defocused sub-aperture and the reception is performed again by it, as shown in Figure

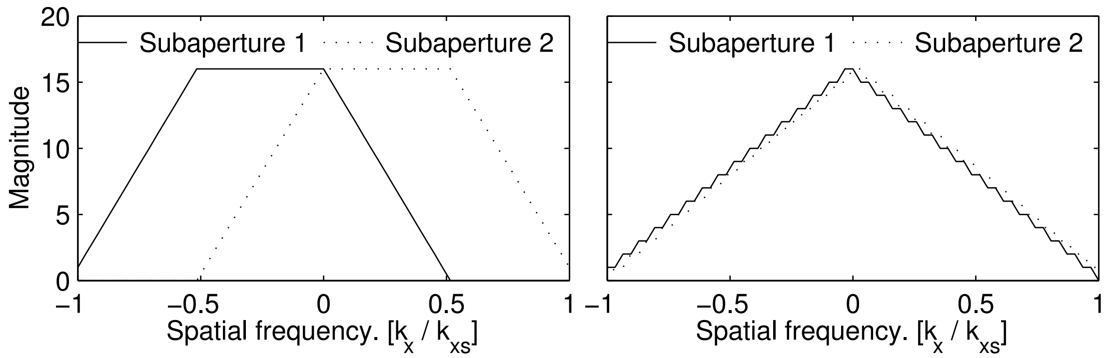


Figure 12.2: The k-space representation of the SRAU. The left graph shows the case of neighboring and the right graph of interleaved sub-apertures, respectively.

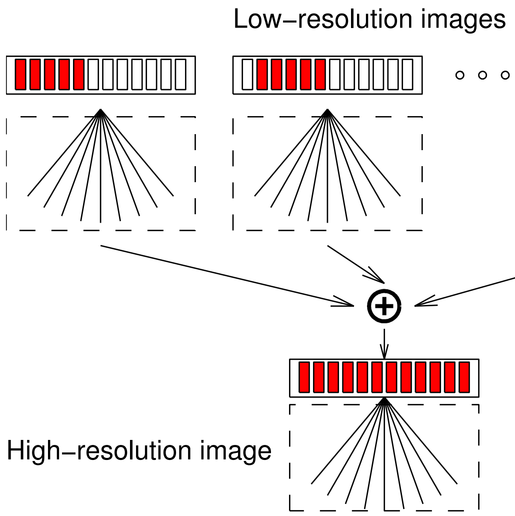


Figure 12.3: Forming a high-resolution image from low-resolution sub-aperture images as in [2].

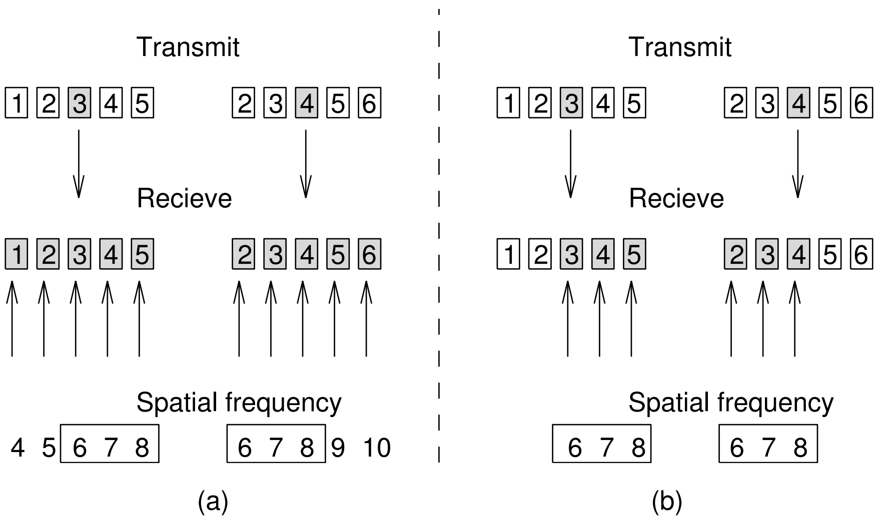


Figure 12.4: Sub-apertures used for the motion estimation using (a) partially, and (b) fully common spatial frequencies. The shaded squares symbolize the active elements. The common spatial frequencies are surrounded by rectangles.

12.3. At every transmit a low-resolution image is formed. In the end all low-resolution images are added together. The receive sub-apertures are centered around the transmitting elements as shown in Figure 12.4. The figure shows two consecutive transmit events using transmit elements with indexes 3 and 4. Transmitting with element i and receiving with element j results in the creation of an element from the effective aperture (or equivalently a spatial frequency) with index $i+j$. Thus, transmitting with element 3 and receiving with elements $1 \div 5$ creates elements from the virtual aperture with indexes $4 \div 8$, as shown in Figure[1] 12.4. Two different configurations for the motion estimation from the low-resolution images were investigated in [2, 149]: low resolution images formed using the signals from the whole receive sub-aperture,

[1]The notation in Figure 12.4 reads "spatial frequency" following the notation from the original paper [2].

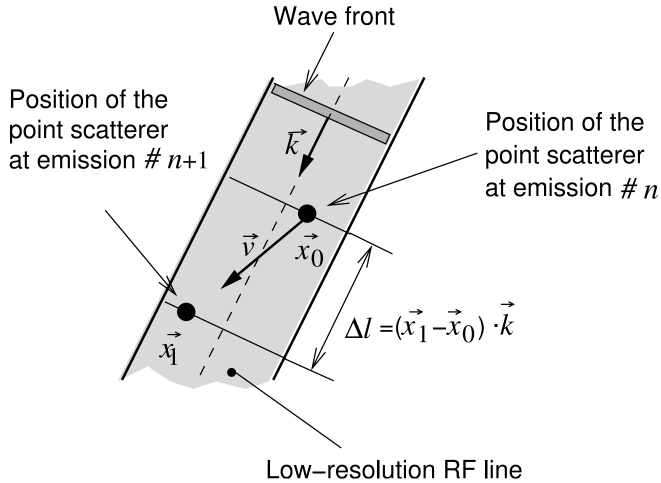


Figure 12.5: Model for motion compensation for synthetic aperture ultrasound imaging without spatial encoding.

and using only the signals from the elements that form the same effective aperture elements (or put in other words, have the same spatial frequencies). The signals that participate in the formation of the images with partially and fully common spatial frequencies are shown with arrows in Figure 12.4(a) and (b), respectively. The group demonstrates in [2] and [149] that using the signals with fully common spatial frequencies result in higher correlation between the low-resolution images. The correlation coefficient is on the order of 0.94, which according to their experimental results suffices for the motion estimation.

As demonstrated by the previous work in motion compensation of synthetic aperture ultrasound imaging the motion can be estimated more reliably by using images that have high correlation between them. The images that exhibit the highest correlation are the high-resolution images, which led to the method suggested by the author in [150]. The idea can be summarized as follows: precise velocity estimates can be obtained from the high resolution images created at every emission using recursive ultrasound imaging (RUI). These velocity estimates are used for motion compensation in the beamformation process. The high-resolution images obtained in this way retain the phase information necessary to perform velocity estimation, and so on.

In the next sections the model for the motion compensation is developed and the performance of estimating the velocity is given.

12.2 Models for motion compensation

Because of the different steps involved in the image formation the models for motion compensation are divided in two: without and with spatially encoded transmits.

12.2.1 Synthetic transmit aperture imaging

During the first stage of the beamformation process low-resolution images are created using dynamic receive focusing. The assumption is that within one scan line L_t the wavefront propagates as a plane wave, as illustrated in Figure 12.5. Figure 12.5 shows the movement of a

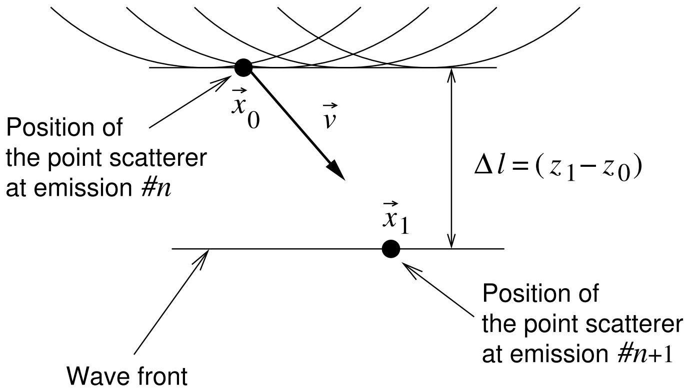


Figure 12.6: Model for motion compensation for synthetic aperture ultrasound imaging with spatial encoding.

single point scatterer within the limits of one scan line. The scatterer moves with velocity $\vec{v}$ from position $\vec{x}_0$ to a new position $\vec{x}_1$ for the time T_{prf} between two emissions. The movement across the beam (perpendicular to the direction $\vec{k}$) determines the strength of the backscattered energy, while the movement along the beam determines the time instance when the backscattering occurs.

For the case depicted in Figure 12.5 the difference in time when the backscattering occurs for the positions $\vec{x}_0$ and $\vec{x}_1$ is:

$$t_s = \frac{2\Delta l}{c}, \tag{12.1}$$

where Δl is the distance traveled by the scatterer for the time from one emission to the next. This distance is given by:

$$\Delta l = \vec{v} \cdot \vec{k} T_{prf}, \tag{12.2}$$

where $\vec{v} \cdot \vec{k}$ gives the velocity component along the scan line.

12.2.2 Synthetic transmit aperture imaging with spatially encoded transmits

Figure 12.6 shows the model adopted for motion compensation in the presence of spatially encoded signals. Several transducer elements across the whole span of the aperture are used in transmit. The sum of the emitted waves creates a planar wave propagating in a direction perpendicular to the transducer surface. A point scatterer moves for one pulse repetition period T_{prf} from position $\vec{x}_0$ to a new position $\vec{x}_1$. The difference between the time instances, when the scattering occurs is:

$$t_s = \frac{2\Delta l}{c}, \tag{12.3}$$

where Δl is the distance traveled by the point scatterer in *axial direction* (along z):

$$\Delta l = v_z T_{prf} \tag{12.4}$$

In the above equation v_z is the velocity component normal to the transducer surface.

The difference between (12.1) and (12.3) is that in the former case the shift is along a *beamformed* line $L(t)$, while in latter case the shift is in the received signals $r_j(t)$.

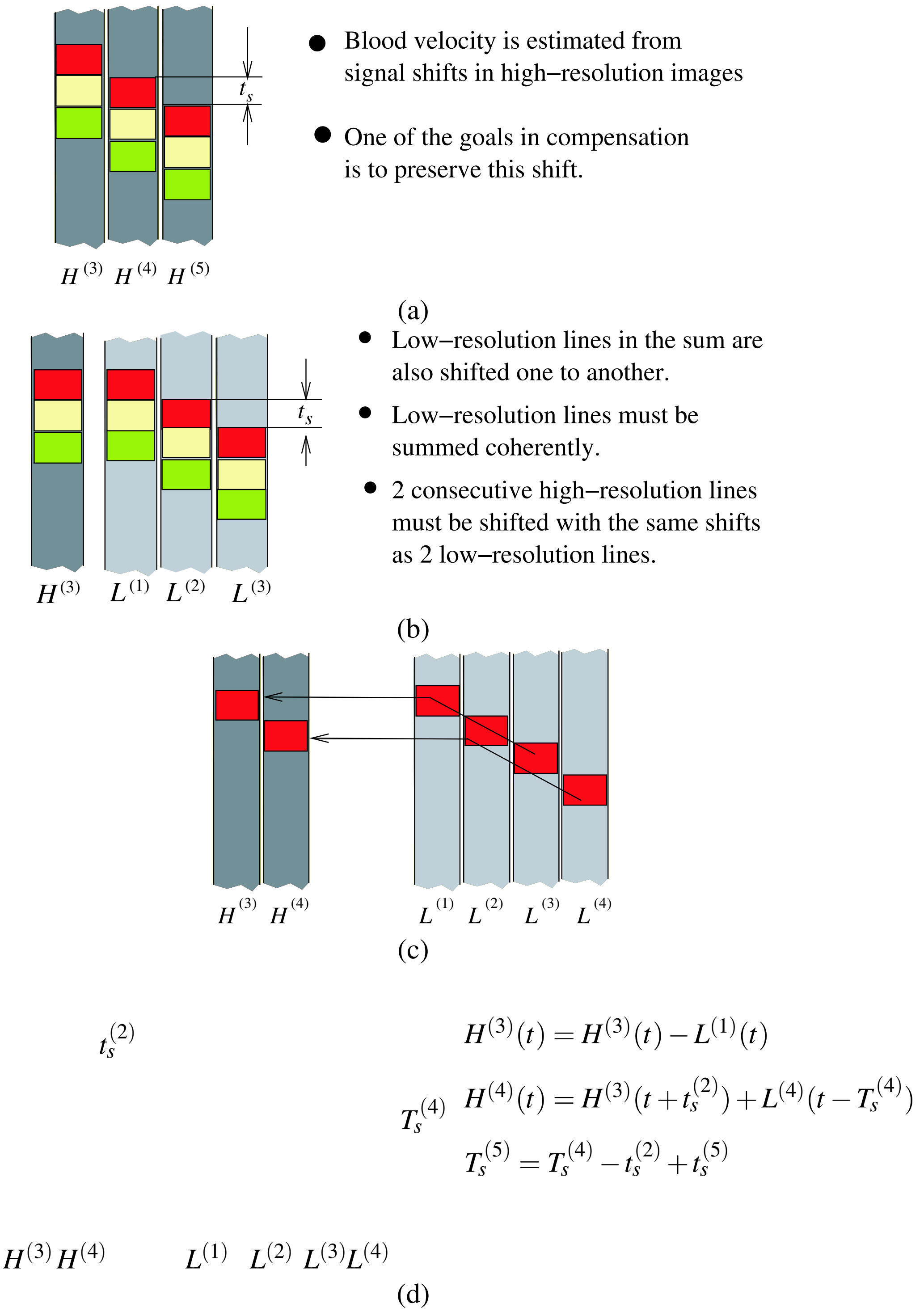


Figure 12.7: Summary of the considerations and steps involved in the motion compensation.

12.3 Motion compensation in recursive imaging

12.3.1 Recursive ultrasound imaging without spatial encoding

All of the considerations and steps in the motion compensation process are summarized in Figure 12.7. The main goal of motion compensation is to preserve the phase information in the high-resolution beamformed RF lines. This is necessary because the motion must be estimated from the high-resolution images. Figure 12.7(a) shows that the first step in motion compensation is to estimate the velocity. This estimation can be done by finding the time shift t_s between two consecutive lines as shown in sub-figure (a) (cross-correlation velocity estimation), or by finding phase shifts (auto-correlation velocity estimation), or by any other method. The RF lines from high-resolution images have higher correlation than the RF lines in the low-resolution images, thus the time shift can be detected more reliably. The high-resolution RF lines are better focused and the spread of velocities within one resolution volume is lower, resulting in a lower bias of the estimates. Figure 12.7(a) shows that the high resolution high-resolution scan line $H^{(n+1)}(t)$ at emission $(n+1)$ is a time shifted version of the high-resolution scan line formed at emission $H^{(n)}(t)$:

$$H^{(n+1)}(t) = H^{(n)}(t - t_s). \tag{12.5}$$

This is a standard assumption used in blood velocity estimation [132]. Usually the velocity is estimated using several RF lines to decrease the uncertainty of the estimate. Thus the velocity estimation gives the mean velocity for the number of lines used in the estimation. The shift t_s present in high resolution images is present also in the low resolution ones. For two consecutive low-resolution scan lines $L^{(n)}(t)$ and $L^{(n+1)}(t)$ this relation is written as:

$$L^{(n+1)}(t) = L^{(n)}(t - t_s), \tag{12.6}$$

and is illustrated in Figure 12.7(b).

Figure 12.7(c) shows how the low-resolution scan lines $L^{(n)}(t)$ are summed coherently in the high-resolution scan lines $H^{(n)}(t)$. The red blocks symbolize a segment of the signal corresponding to a single point scatterer that moves at a constant speed away from the transducer. Creating the high resolution image $H^{(3)}(t)$ at emission 3 all the segments are aligned to the low-resolution image $L^{(3)}(t)$. Creating $H^{(4)}(t)$, all signals are aligned with respect to $L^{(4)}(t)$.

Up to this moment it was assumed that the velocity is constant in depth and time (emission number, or slow time). This is, however, not true. Acceleration in the motion is usually present. The time shifts t_s are a function of the pulse emission. The time shift $t_s^{(n+1)}$ at emission $n+1$ can be estimated from the high resolution images up to emission n. The time shift is also a function of depth, or equivalently of the time t from the trigger of the current emission, $t_s^{(n)} = t_s^{(n)}(t)$.

Let the number of low-resolution images used to make a high-resolution one be N_{xmt}. Without motion compensation the beamformation process would be:

$$H^{(n)}(t) = \sum_{i=n-N_{xmt}+1}^{n} L^{(i)}(t) \tag{12.7}$$

Using the motion compensation the beamformation process is expressed as:

$$\begin{aligned} H^{(n)}(t) =& L^{(n-N_{xmt}+1)}\big(t\big) + L^{(n-N_{xmt}+2)}\big(t - t_s^{(n-N_{xmt}+2)}(t)\big) + \\ & + L^{(n-N_{xmt}+3)}\big(t - (t_s^{(n-N_{xmt}+2)}(t) + t_s^{(n-N_{xmt}+3)}(t))\big) + \cdots \\ & + L^{(n)}\big(t - \sum_{j=n-N_{xmt}+2}^{n} t^{(j)}(t)\big) \end{aligned} \tag{12.8}$$

The sum of the delays $t_s^{(i)}$ from the first low-resolution image entering the sum to the last low-resolution image is called *accumulated time shift* and is denoted $T_s^{(n)}$:

$$T_s^{(n)}(t) = \sum_{j=n-N_{xmt}+2}^{n} t_s^{(j)}(t). \tag{12.9}$$

Figure 12.7 shows the case of using $N_{xmt} = 3$ low resolution images to create a high-resolution one. It can be seen that the high-resolution images $H^{(n-1)}(t)$ and $H^{(n)}(t)$ are shifted one to another with a time shift $t_s^{(n-N_{xmt}+1)}$. The high-resolution image $H^{(n-1)}(t)$ and the low-resolution image $L^{(n-N_{xmt})}$ are aligned. The recursive procedure for beamformation (see Eq. (6.13) on page 57) becomes:

$$H^{(n-1)}(t) = H^{(n-1)}(t) - L^{(n-N_{xmt})}(t) \tag{12.10}$$

$$H^{(n)}(t) = H^{(n-1)}(t + t_s^{(n-N_{xmt}+1)}) + L(n)(t - T_s^{(n)}) \tag{12.11}$$

$$\text{from } H^{(n)}(t), H^{(n-1)}(t) \text{ ... estimate the velocity } \vec{v}^{(n+1)}(t) \tag{12.12}$$

$$t_s^{(n+1)}(t) = \frac{2\vec{v}^{(n+1)}(t)\cdot\vec{k}T_{prf}}{c} \tag{12.13}$$

$$T_s^{(n+1)}(t) = T_s^{(n)}(t) + t_s^{(n+1)}(t) - t_s^{(n-N_{xmt}+2)}(t). \tag{12.14}$$

12.3.2 Recursive imaging with spatially encoded signals

The motion compensation for the recursive imaging with spatially encoded transmits is performed on the received RF data. Considering only the emissions from 1 to N_{xmt}, and using the simplified model the received RF data by a single element, j can be expressed as:

$$\begin{aligned} r_j^{(1)}(t) &= \sum_{i=1}^{N_{xmt}} q_{i1} \cdot r_{ij}(t - T_s^{(1)}(t)) \\ r_j^{(2)}(t) &= \sum_{i=1}^{N_{xmt}} q_{i2} \cdot r_{ij}(t - T_s^{(2)}(t)) \\ &\vdots \\ r_j^{(N_{xmt})}(t) &= \sum_{i=1}^{N_{xmt}} q_{iN_{xmt}} \cdot r_{ij}(t - T_s^{(N_{xmt}-1)}(t)), \end{aligned} \tag{12.15}$$

where q are the encoding coefficients and $r_{ij}(t)$ is the signal that would be received by element j after the emission of element i in the absence of motion. The time shifts $t_s^{(n)}(t)$ are assumed to be due only to the motion perpendicular to the transducer surface and $t_s^{(1)}(t)$ is set to 0. The compensation and the reconstruction can be expressed as:

$$r_{ij}(t) = \sum_{n=1}^{N_{xmt}} q_{in} \cdot r_j^{(n)}(t - T_s^{(n)}(t)), \quad \text{for } i \in [1, N_{xmt}] \tag{12.16}$$

A more precise compensation scheme can be derived using the time shifts t_s, which based not only on estimated velocity but also on the position of the transmitting and receiving elements. This means to replace the plane wave from the assumed model with a spherical wave.

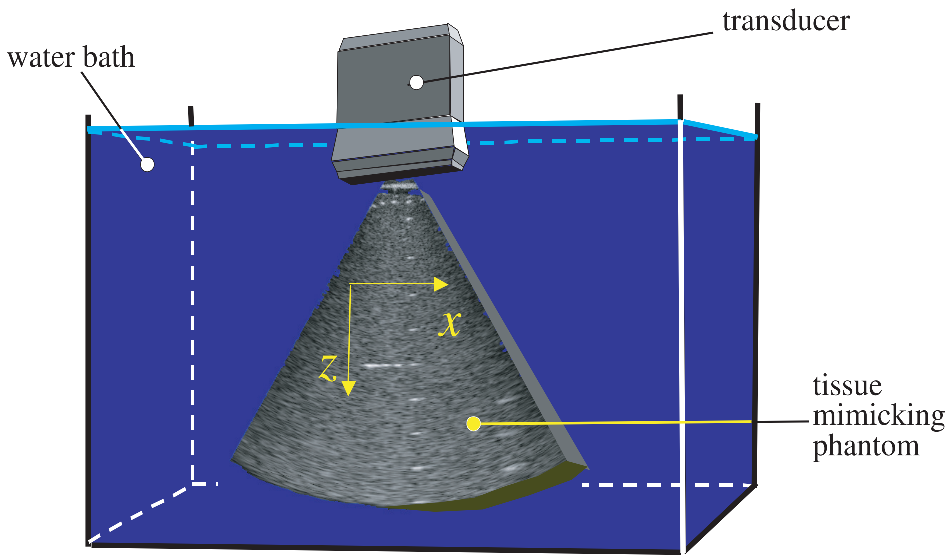


Figure 12.8: Measurement setup.

12.4 Experimental verification

The procedure for motion compensation was verified experimentally.

12.4.1 Experimental setup

The measurements were done using the XTRA system (see Appendix J). A tissue mimicking phantom with frequency dependent attenuation of 0.25 dB/(cm · MHz) and speed of sound $c = 1540$ m/s was scanned at 65 positions in a water bath as shown in Figure 12.8. From position to position the phantom was moved 70 μm at an angle of 45° relative to the transducer surface. Assuming a pulse repetition frequency of $f_{prf} = 1/T_{prf} = 7000$ Hz, this movement corresponds to a plug flow with velocity $|\vec{v}| = 0.495$ m/s. At every position a data set from 13 emissions was recorded. For each position a reference high-resolution image was formed using all of the 13 emissions. The recursive ultrasound images were formed by using only a single emission from a position. In this way any shot sequence could be realized. The system is not capable of using negative apodization coefficients, and therefore the encoding was performed in software for each position as in [3].

A precision translation system (the XYZ system) was used for the movement of the phantom. The precision of motion in the axial and lateral directions was: $\Delta z = 1/200$ mm, and $\Delta x = 1/80$ mm, respectively.

12.4.2 Experimental results

A simple visual inspection of the B-mode images reveals that the motion compensation procedure using the aforementioned motion compensation models and procedures works. Two cases were investigated: (1) using 13 emissions and classical recursive imaging ($N_{xmt} = 13$, see Section 6.1), and (2) using 4 emissions and spatial encoding using Hadamard matrices.

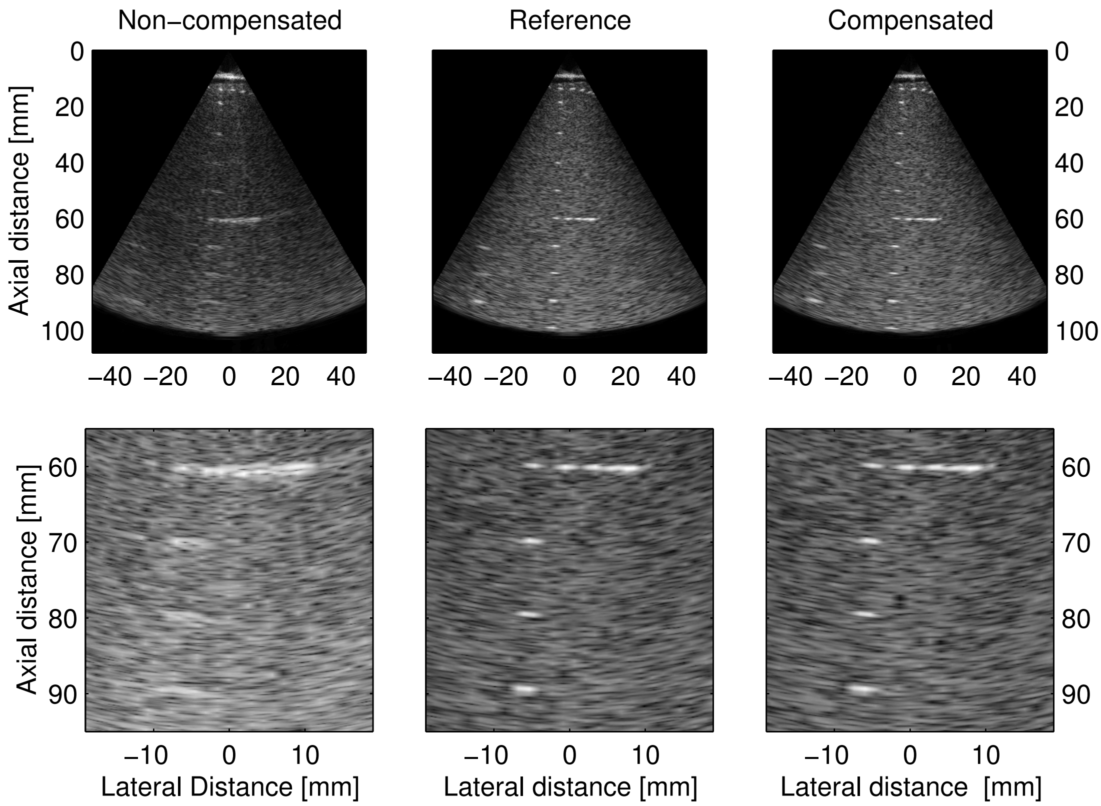


Figure 12.9: Comparison of B-mode images. The left column shows an image obtained using recursive ultrasound imaging without motion compensation. The center column shows a reference image obtained at one position. The right column shows an image using recursive ultrasound imaging and motion compensation.

Figure 12.9 shows B-mode images with a dynamic range of 50 dB. The left and the right columns show the high-resolution synthetic aperture images created at emission $n = 15$. The left column image is created without motion compensation and the image in the right column is created using the motion compensation scheme as described in Section 12.3.1. The center column shows a reference image created at one position, using 13 emissions. From the bottom row it can be seen that the motion compensated image is more "focused" than the non-compensated one and is visually closer to the reference image. The visual impression can be also confirmed by the auto-covariance analysis shown in Figure 12.10. The top plot shows the projection of the maxima in axial direction of the 2D auto-covariance analysis. As it can be seen the plots of the reference and motion-compensated images are almost identical. The level of the non-compensated synthetic image is at the end of the plot is almost 10 dB higher. This means that the contrast of the image is lower and dark regions are hard to image. The bottom plot shows the -6 dB contours of the 2D auto-correlation function of the images taken around a point scatterer (the point at 80 mm). The contours of the compensated and the reference images coincide.

Another measure for the coherent summation of the image is the energy contained in it. Figure 12.11 shows the RMS values of the lines from the images in Figure 12.9. The success of the motion compensation can be judged by the fact the RMS values of the lines in the motion compensated image are comparable to those in the reference image. The mean of the RMS

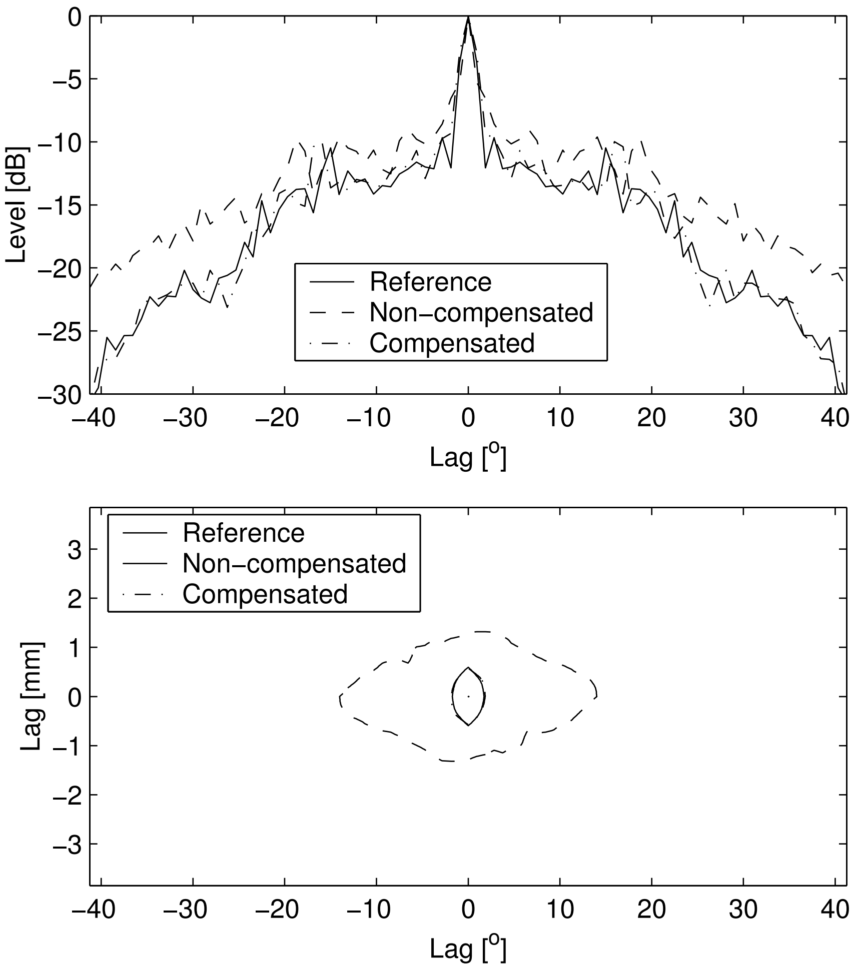


Figure 12.10: Resolution of the images. The upper plot shows the maximum along the RF lines of the 2D auto-covariance of a speckle region. The lower plot shows the 2D auto-correlation functions in a region of a single point scatterer.

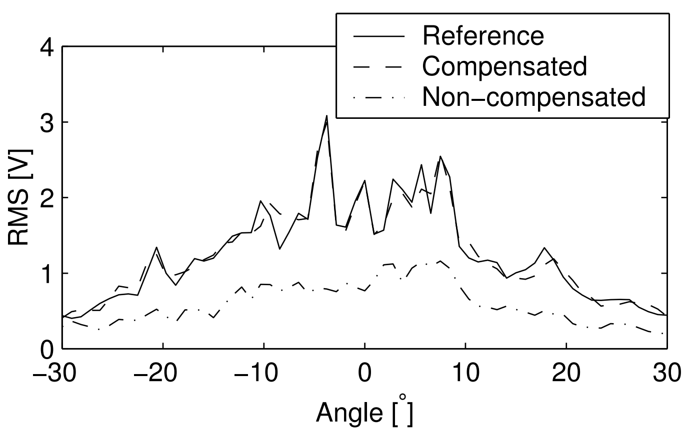


Figure 12.11: The RMS values of the images obtained using recursive ultrasound imaging as a function of angle.

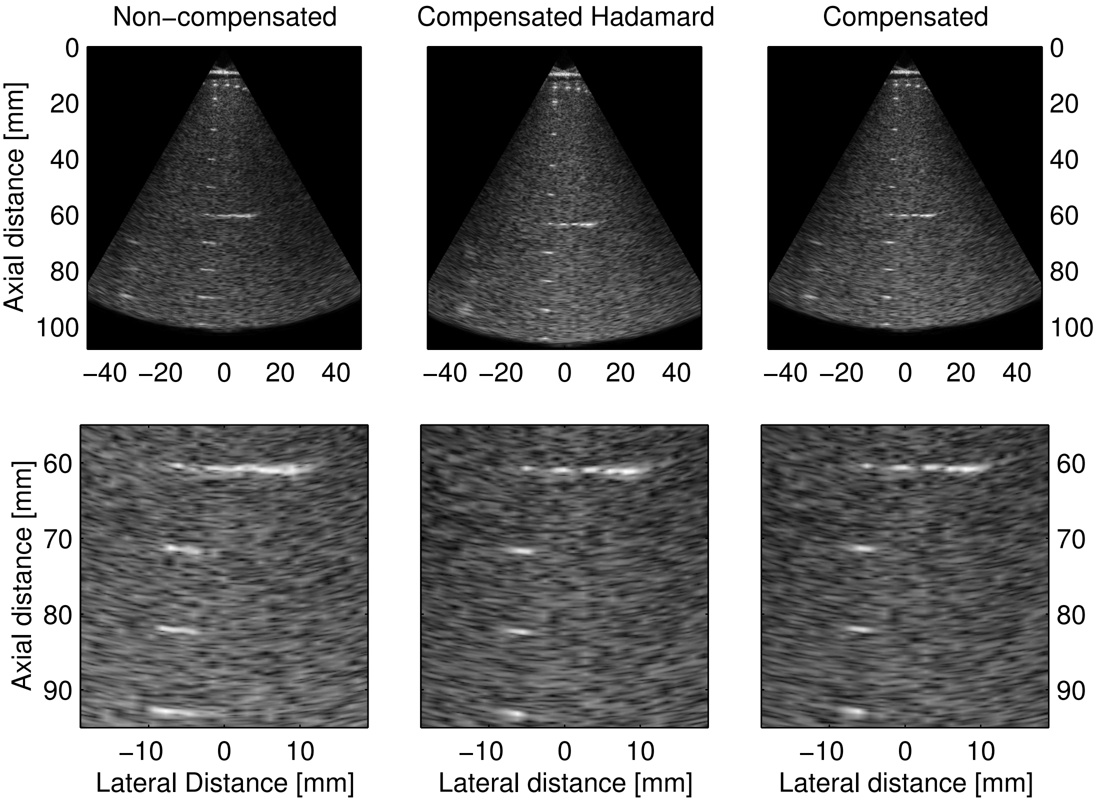


Figure 12.12: Comparison of B-mode images. The left column shows an image obtained using recursive ultrasound imaging without motion compensation and spatial encoding. The right column shows an image obtained using motion compensation and spatial encoding using Hadamard matrix.

values for the reference, motion compensated and non-compensated images is 1.258, 1.256 and 0.588 V, respectively.

So far it was shown that the motion compensation scheme works for the case of synthetic aperture imaging (or recursive ultrasound imaging). The remaining question is whether the motion model and the motion compensation schemes for spatially encoded transmits are adequate. Figure 12.12 shows three images with 45 dB dynamic range: two formed without and one with spatial encoding. All the images are acquired with only 4 transmit events. The left column shows a non motion-compensated synthetic aperture image and the right column shows a motion-compensated one. The middle image is made using 4×4 Hadamard encoding matrix. The motion compensation was done prior to the decoding as described in Section 12.3.2. It can be observed that the two compensated images look alike, and that the non-compensated one has a lower resolution. This is confirmed by the plots in Figure 12.13. The difference in the resolution of the images is not as big as for the case of 13 emissions. The shorter the time, the smaller the traveled distance and motion artifacts are. Figure 12.14 shows the RMS values of the signals for the three cases depicted in Figure 12.12. The mean RMS values for the non-compensated, the compensated without and with spatial Hadamard encoding are: 0.399, 0.447, and 0.455, respectively. The conclusion is that the compensation for the Hadamard encoding using 4 emissions is as good as the motion compensation for non-encoded synthetic aperture images obtained with the same number of emissions.

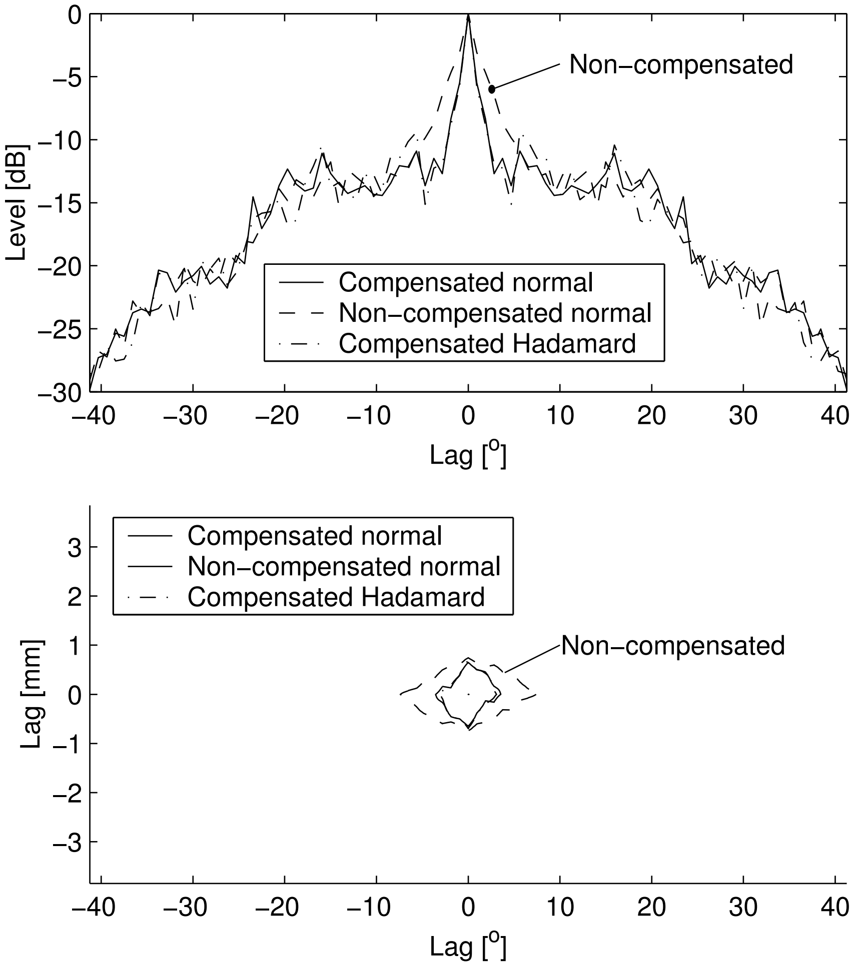


Figure 12.13: Resolution of the images. The upper plot shows the maximum along the RF lines of the 2D auto-covariance of a speckle region. The lower plot shows the 2D auto-correlation functions in a region of a single point scatterer.

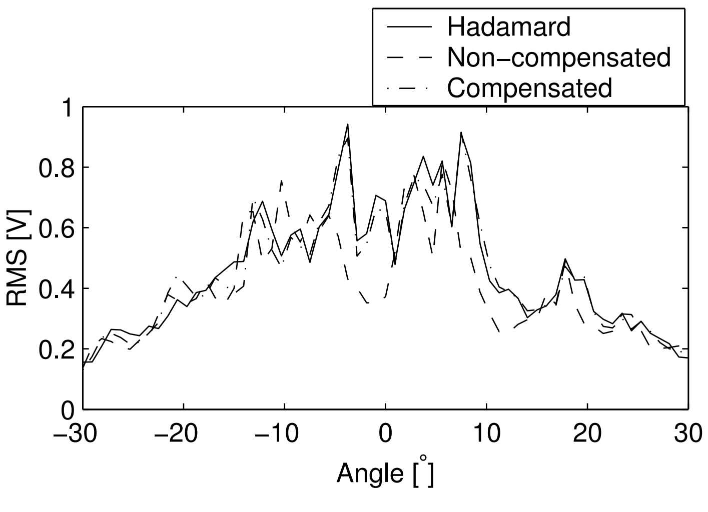


Figure 12.14: The RMS values of the images with and without Hadamard encoding and motion compensation.

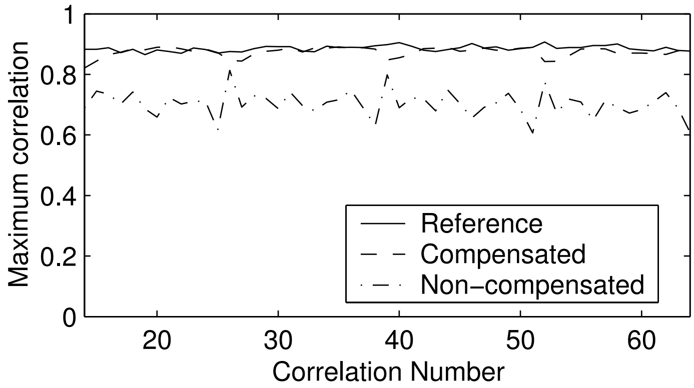


Figure 12.15: The maximum cross-correlation value between two consecutive frames.

There exist other approaches to evaluate the success of the motion compensation, such as the one used by Maurice and Bertrand in [147]. The motion artifact in their work is estimated from the cross-correlation of the original, ”premotion” image and the motion compensated one.

One of the goals in the motion compensation is to create images which not only have nice appearance, but can be also used for extracting velocity information necessary for the motion compensation. The more alike two consecutive frames are, the lower the bias and the variance of the estimates are. From each high-resolution image seven regions were taken, and the peaks of the 2D normalized cross-correlation with the same regions from the next high-resolution images were found. Then the average of these peaks was plotted in Figure 12.15. The normalization was performed with the peak of the auto-correlation functions of the respective regions. The average maximum of the cross-correlation function between two consecutive high-resolution images for the reference, compensated and non-compensated case without spatial encoding are: 0.89, 0.87, and 0.7. This means that the velocity from the motion-compensated images can be estimated almost as well as from the reference images obtained at every position of the phantom. To test this the velocity from the motion compensated images was estimated.

12.4.3 Velocity estimation

Three sets of velocity estimates were obtained from the scanned data: (1) reference (2) without spatial encoding, and (3) with spatial encoding. In order to obtain a reference estimate of the velocity, at each position of the phantom a high resolution image was created using 13 emissions per image. The velocity was estimated using a cross-correlation velocity estimator as suggested by Bonnefous and Pesqué [133] and is described in Chapter 10. Although any estimator can be used for the velocity estimation, this estimator was chosen because it is suitable for the broad-band pulses used in the acquisition process. The correlation length was equal to the length of the transmitted pulse. The number of lines, over which the correlation function was averaged, was 8. The search length was $\pm\lambda/2$.

The velocity estimates without spatial encoding were obtained from the same data set. The number of emissions N_{xmt} was set to 13. Only a single emission at each position was used in the formation of the image. The emissions used were cyclically rotated, i.e. after the emission with element number 64, the emission with element number 1 was used. The compensation was performed using a 32 coefficients FIR interpolation filter for the fine delays. The fine delays

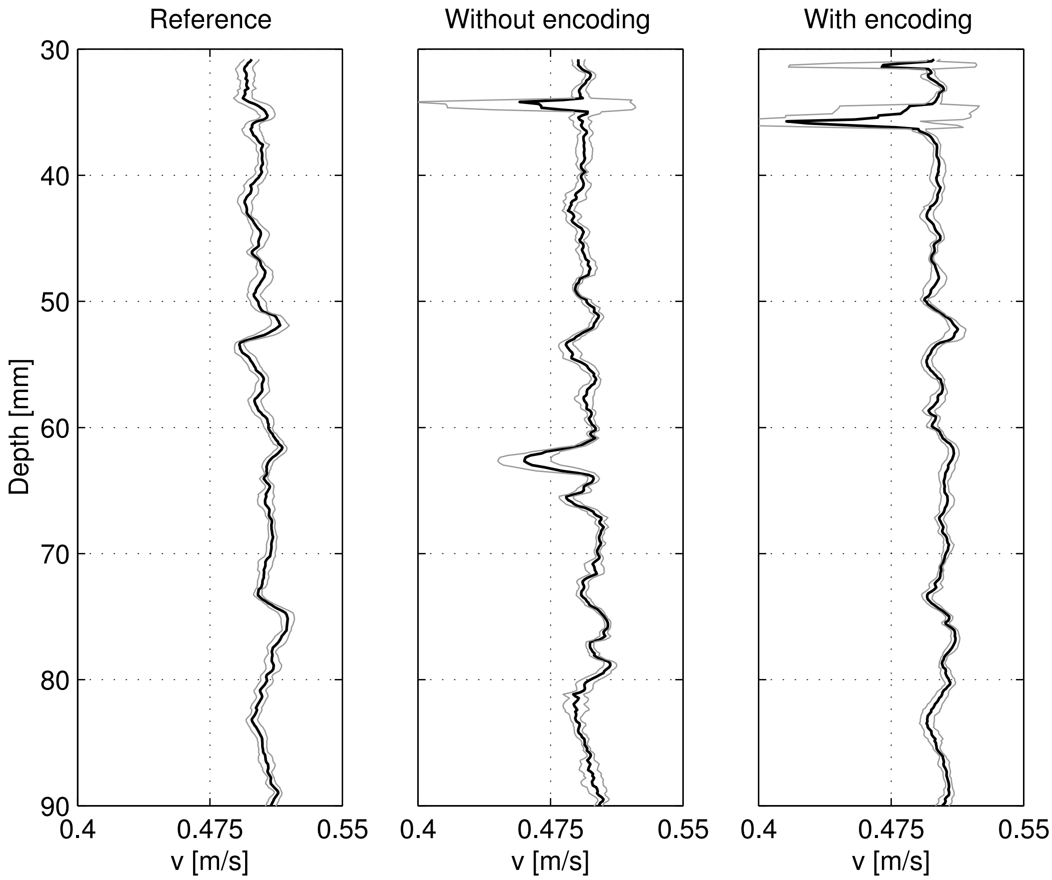


Figure 12.16: The mean velocity and the velocity at $\pm$ one standard deviations as a function of depth.

were set to $1/(100f_s)$, which corresponds to a sampling rate of 400 MHz. For the first eight high resolution images, the knowledge of the actual movement of the system was used.

The high-resolution images using spatially encoded transmits were obtained using only $N_{xmt} = 4$ elements in transmit. Then the encoded data from 4 consecutive emissions was used to obtain the individual decoded data sets. Prior to the decoding the motion compensation was performed. The filter bank used in the motion compensation is the same as for the case without spatial encoding.

Figure 12.16 shows the velocity estimates as a function of depth for the scan line at 45° to the flow direction ($\angle(\vec{v},\vec{k}) = 45°$). Because the whole phantom was moved the velocity was constant for all depths. The real velocity is 0.495 m/s. The mean and the standard deviations were calculated over 52, 39, and 48 estimates for the reference velocity, the velocity for the recursive imaging without and with spatial encoding, respectively. The number of available estimates varies with the number of positions used to create a high-resolution image. The estimates are not independent. It can be seen that estimates using recursive imaging do not deviate significantly from the estimates used as a reference.

The angular dependency of the velocity estimates is shown in Figure 12.17. The dashed lines show the velocity at $\pm\sigma$, where σ is the standard deviation. This angular velocity distribution was obtained by calculating the velocity for all the scan lines in the image, not by changing the motion direction of the phantom.

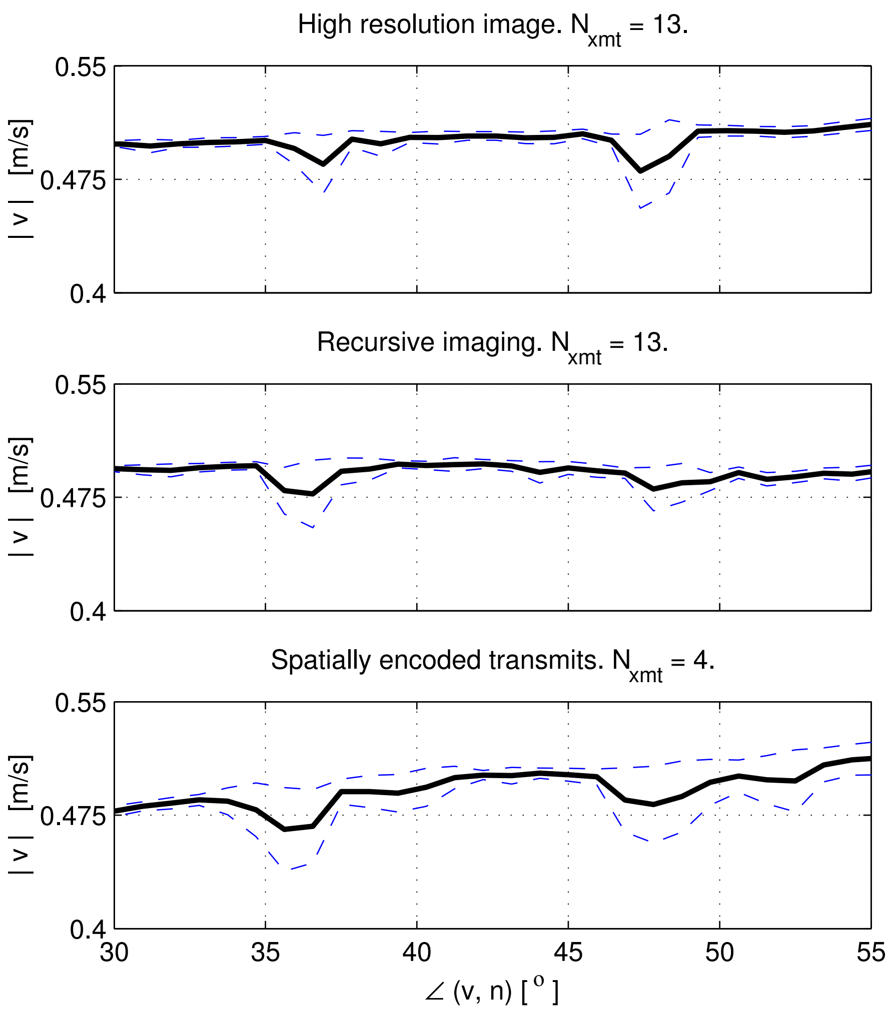


Figure 12.17: The mean velocity and the velocity at ± one standard deviations as a function of angle.

	Reference	Spatially encoded	Non encoded
mean(v) [m/s]	0.496	0.486	0.479
std(v) %	2.3	2.2	1.8

Table 12.1: Results from the velocity estimation at an angle 45°.

It can be seen that the reference velocity estimation exhibits a smaller bias than the velocity estimation using recursive imaging. This could potentially lead to a worse motion compensation over a longer period of time.

The recursive imaging using spatially encoded transmits exhibits angular dependence. At angles 42° – 45° it has a low bias and standard deviation comparable to that of the reference velocity estimation. There are two possible reasons for this dependence: (1) the low number of emissions resulting in higher side and grating lobes, and (2) the compensation model. As was pointed out in Section 12.3.2, only the velocity component normal to the transducer surface is compensated for. This is sufficient for the region below the transducer, but obviously is not an adequate model for the echoes coming from outer parts of the scanned region.

Table 12.1 shows the average mean velocity and standard deviation of the estimates obtained at an angle of 45 $^{\circ}$ giving concrete numbers to confirming the discussion of the results.

CHAPTER

THIRTEEN

Velocity estimation using recursive ultrasound imaging

In Chapter 12 a way to estimate the velocity and to perform motion compensation was given. The method relies on the fact, that using long data sequences from the practically uninterrupted data flux, the velocity can be estimated with a lower bias and standard deviation. Precise velocity estimation leads to precise motion compensation, which on its turn leads to precise velocity estimation. Breaking this circle, however, would lead to the opposite effect: bad motion estimation leads to bad motion compensation, which on its turn worsens the motion estimation. Thus, the method is not robust[1].

This chapter introduces a modified cross-correlation estimator making it possible to estimate the velocity out of the non motion-compensated high-resolution images. First the rationale behind the method is given and then the results of the velocity estimation with the new method are compared with the existing methods.

13.1 Derivation of the method.

13.1.1 Measurement principle

Consider the simple experiment simulated in Field II and shown in Figure 13.1. A single scatterer moves towards the transducer. The image is formed using four transmissions. Initially the scatterer is centered at 55 millimeters away from the transducer. Between two emissions it moves at a distance $\Delta z = -\lambda/4$. The simulated transducer has 64 elements, a center frequency of 5 MHz and a pitch of λ, with assumed speed of sound 1540 m/s. The left and right images are formed transmitting sequentially with elements $\{1, 22, 43, 64\}$ and $\{22, 43, 64, 1\}$, respectively. In both cases the scatterer starts from the same position. The middle row of images in Figure 13.1 show 50 dB B-mode images of the obtained point-spread-functions. They are not symmetric and differ in shape. Comparing the two RF lines (given in yellow on the B-mode images) reveals that the echoes are shifted and different in shape. Because of the change in shape the cross-correlation is low and the velocity cannot be estimated with a high precision. If the synthetic aperture is to be used for velocity estimation of *blood flow*, one must consider the velocity gradient present in the flow, which decreases the correlation between the consecutive RF signals even more. The conclusion is that the consecutive high-resolution non-motion-compensated RF lines are not suited for velocity estimations.

[1]For me a robust method is a method not sensitive to small errors, and without memory of previous errors.

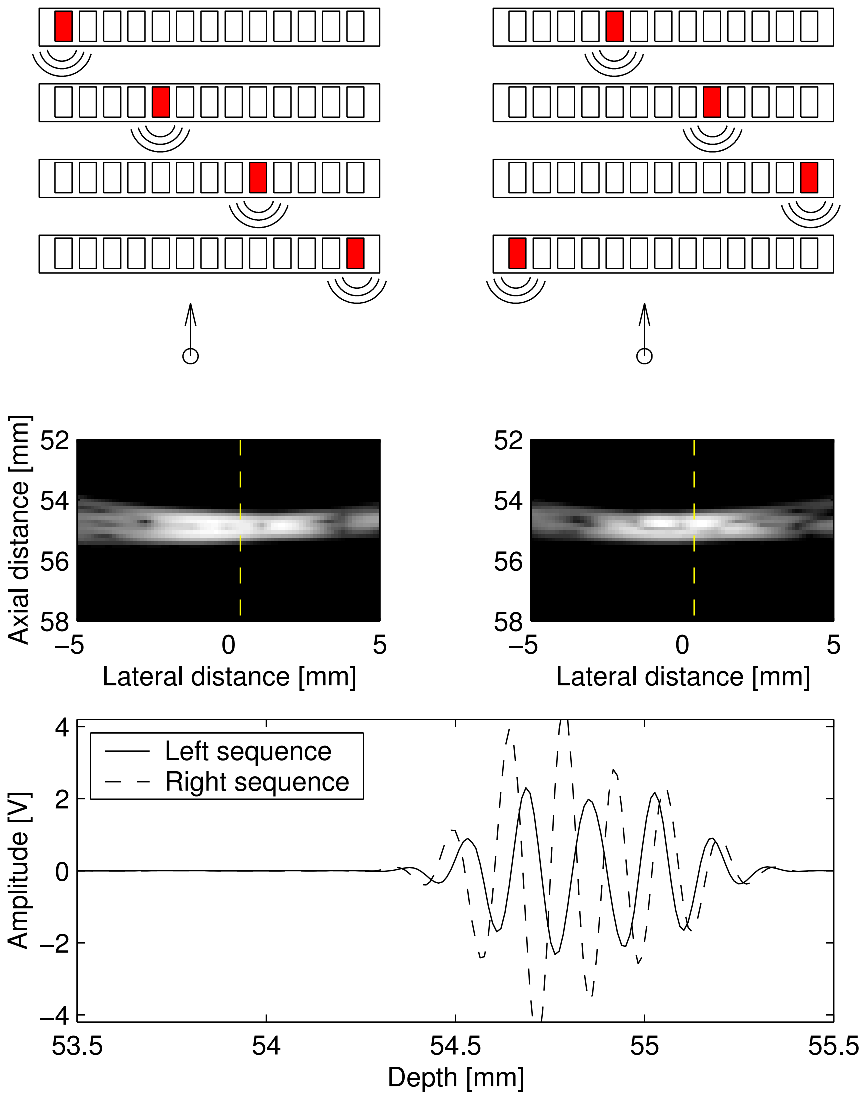


Figure 13.1: Point spread functions obtained using different transmit sequences.

The reason for this is that the motion artifacts are caused by the change of position of both, the scatterer and the transmitting element as discussed in the end of Chapter 11. This is confirmed by Figure 13.1 which shows that for the same displacement of the scatterer, and different transmit sequence the motion artifacts differ.

The velocity estimation [132] using conventional beamformation methods, in essence, compares RF signals from the same range (depth). Because the sound travels through the same tissue layers, the data samples taken at the same spatial position have undergone the same distortions (phase aberration, refraction, shift in mean frequency, attenuation, etc.). The estimation of blood is based on comparing the changes (phase shifts) in the signal at the particular point, from which the velocity is estimated. The same principle can be applied to estimating the velocity using synthetic aperture imaging.

Consider Figure 13.2 which shows the creation of 3 high resolution images using two emissions per image. The transmit sequence is illustrated in the top left part of the figure. In the right side of the figure the B-mode images of the created point-spread-functions are given. It can be clearly seen that they have different geometries. The bottom plot shows a single RF line from

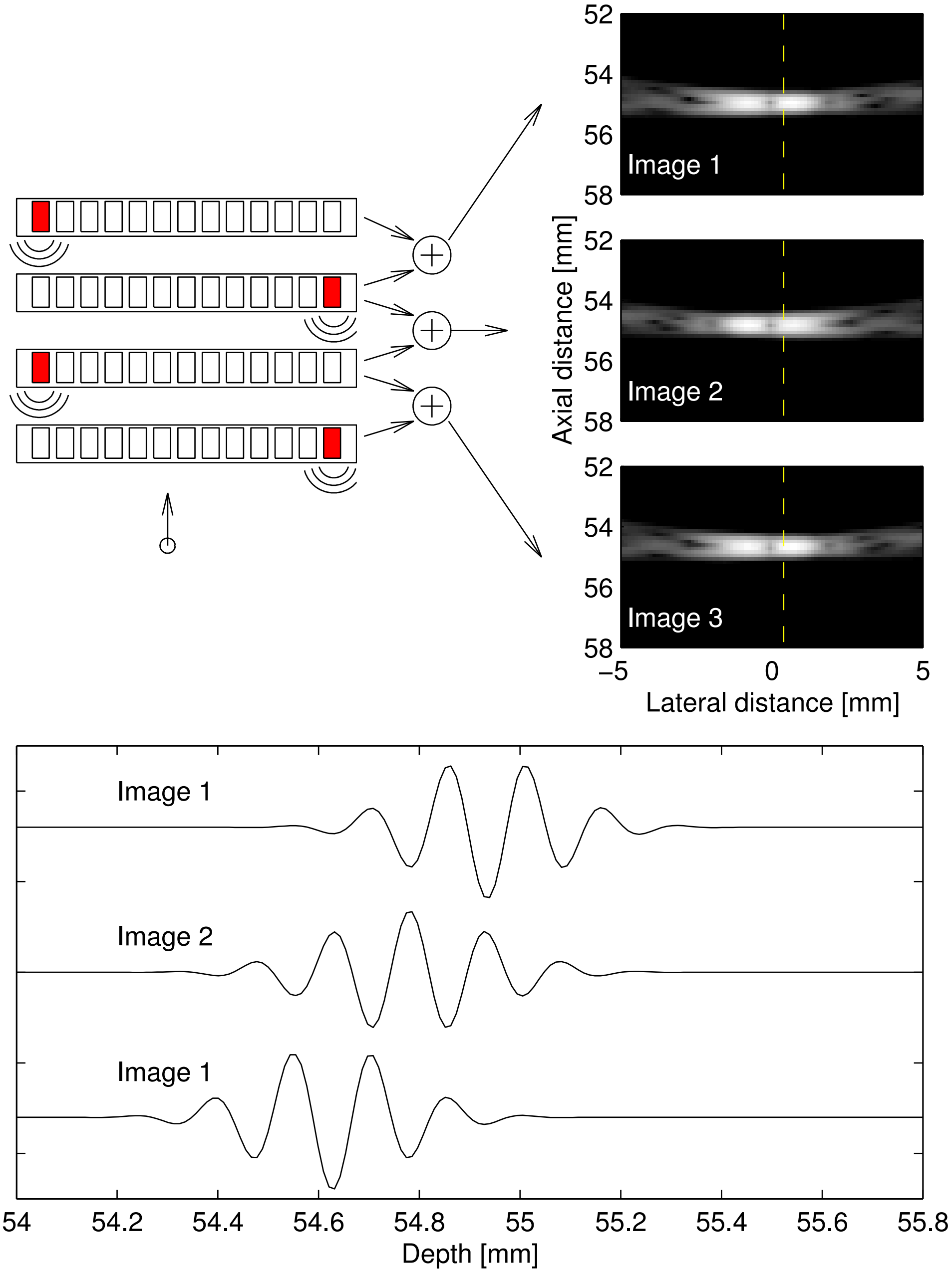


Figure 13.2: Comparison between three high-resolution images of a moving point scatterer.

each of the images (the line is indicated with yellow on the B-mode images). The RF lines from the first and the third images are time shifted versions of each other, while the RF line from the second image has a different appearance.

Using the notation used in Chapter 11 the point spread-function of a high-resolution image $\mathbf{P}(x,z;\vec{x}_f)$ is obtained from the point-spread-functions of the low resolution images $\mathbf{p}_i(x,z;\vec{x}_f)$ as:

$$\mathbf{P}(x,z;\vec{x}_f) = \sum_i \mathcal{R}[\beta_i]\mathbf{p}_0(x,z;\vec{x}_f), \tag{13.1}$$

where $\vec{x}_f = (x_f, z_f)$ is the position of the point scatterer, and β_i is the angle, at which the point spread function of the low resolution image is inclined, and the subscript 0 of $\mathbf{p}_0$ means that this point-spread-function is parallel to the transducer surface. The angle β_i is related to the transmitting element i with the relation:

$$\beta_i = \arctan \frac{\left(i - \frac{N_{xdc}+1}{2}\right) d_x}{2z_f},$$

where d_x is the pitch of the transducer and N_{xdc} is the number of transducer elements. The same approximations as in Chapter 11 will be used:

$$\mathbf{p}_0(x,z;\vec{x}_f + \Delta\vec{x}) = \mathcal{T}[\Delta\vec{x}]\{\mathbf{p}_0(x,z;\vec{x}_f)\} \tag{13.2}$$

The considerations for the velocity estimations will be made for a single point scatterer as was done in Chapter 10. The signals for a number of emissions from the same scatterer will be needed. The superscript $^{(n)}$ will be used to denote the current emission. The point spread function at emission n will be denoted with $\mathbf{P}^{(n)}(x,z;\vec{x}_f^{(n)})$, where $\vec{x}_f = (x_f^{(n)}, z_f^{(n)})$ is the current position of the scatterer. The considerations will be restricted only to the axial motion and the x_f coordinate will be skipped in the notation further. At position n, the formation of the current point spread function is expressed as:

$$\mathbf{P}^{(n)}(x,z;z_f^{(n)}) \sum_{k=n-N}^{n} \mathcal{R}[\beta_i; z_f^{(k)}]\mathbf{p}_0(x,z;z_f^{(k)}) \tag{13.3}$$

Let for the time span from $(n-2N)T_{prf}$ to nT_{prf}[2] the point scatterer has a constant mean axial velocity v_z. The distance traveled by the scatterer between two emissions is constant $\Delta z = v_z T_{prf}$. Using the aforementioned assumptions one gets:

$$\begin{aligned}
&\mathbf{P}^{(n)}(x,z;z_f^{(n)}) = \\
&= \sum_{k=n-N+1}^{n} \mathcal{T}[0,(k-(n-N+1))\Delta z]\mathcal{R}[\beta_i; z_f^{(n-N+1)}]\{\mathbf{p}_0(x,z;z_f^{(n-N+1)})\}
\end{aligned} \tag{13.4}$$

Considering the point spread function obtained at time instance $n-N$ one gets:

$$\begin{aligned}
&\mathrm{P}^{(n-N)}(x,z;z_f^{(n-N)}) = \\
&= \sum_{k=n-2N+1}^{n-N} \mathcal{T}[0,(k-(n-2N+1))\Delta z]\mathcal{R}[\beta_i; z_f^{(n-2N)}]\{\mathbf{p}_0(x,z;z_f^{(n-2N)})\}
\end{aligned} \tag{13.5}$$

Realizing that:

$$\mathbf{p}_0(x,z;z_f^{(n-N)}) = \mathcal{T}[0,N\Delta z]\mathbf{p}_0(x,z;z_f^{(n-2N)}) \tag{13.6}$$

and

$$\mathcal{R}[\beta_i,(x,z)]\mathcal{T}[\Delta z,\Delta x] \equiv \mathcal{T}[\Delta x,\Delta z]\mathcal{R}[\beta_i;(x-\Delta x, z-\Delta z)] \tag{13.7}$$

[2] N is used instead of N_{xmt} in this chapter. The reason is that the subscript "$_{xmt}$" is too long and the equations cannot fit on a single line.

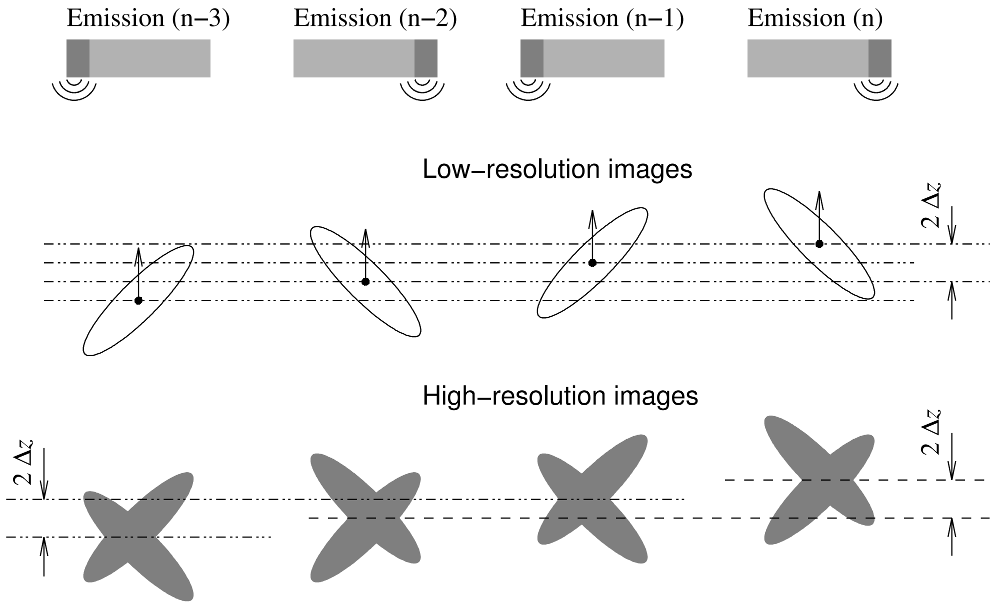


Figure 13.3: Illustration of the fact that images of a moving scatterer acquired at different time instances but using the same transmit order are translated versions of each other.

one gets:

$$\begin{aligned}
&\mathbf{P}^{(n)}(x,z;z_f^{(n)}) = \\
&= \sum_{k=n-N+1}^{n} \mathcal{T}[0,(k-(n-N+1))\Delta z]\mathcal{R}[\beta_i;z_f^{(n-N+1)}]\{\mathbf{p}_0(x,z;z_f^{(n-N+1)})\} \\
&= \sum_{k=n-2N+1}^{n-N} \mathcal{T}[0,N\Delta z]\mathcal{T}[0,(k-(n-2N+1))\Delta z]\mathcal{R}[\beta_i;z_f^{(n-N)}]\{\mathbf{p}_0(x,z;z_f^{(n-2N+1)})\} \qquad (13.8)\\
&= \mathcal{T}[0,N\Delta z] \sum_{k=n-2N+1}^{n-N} \mathcal{T}[0,(k-(n-2N))\Delta z]\mathcal{R}[\beta_i;z_f^{(n-N+1)}]\{\mathbf{p}_0(x,z;z_f^{(n-2N+1)})\} \\
&= \mathcal{T}[0,N\Delta z]\mathbf{P}^{(n-N)}(x,z;z_f^{(n-N)})
\end{aligned}$$

The above equation might look somewhat overwhelming but it can be summarized as follows. Let a scatterer move at a constant speed. Using the same order of emissions, but starting at different initial moments, one can obtain images that have the same shape but at different spatial positions. This process is illustrated in Figure 13.3.

The two-dimensional point spread function $\mathbf{P}(x,z;z_f)$ is obtained by range-gating the beamformed RF lines from the high-resolution image around the time instance t corresponding to the image depth z_f. Equation (13.8) means that for small distances (up to several wavelengths) there is a segment in the high-resolution line $H^{(n)}(t)$, which is a time-shifted version of a segment from the high-resolution line $H^{(n-N)}(t)$:

$$H^{(n)}(t) = H^{(n-N)}(t-t_s). \qquad (13.9)$$

The two scan lines are beamformed in the same direction and the motion is along the scan line itself. There is a limit, however, on the time elapsed between the two scan-lines due to the

Vessel	Peak velocity cm/s	Mean velocity cm/s
Ascending aorta	20 - 290	10 - 40
Descending aorta	25 - 250	10 - 40
Abdominal aorta	50 - 60	8 -20
Femoral artery	100 - 120	10-15
Arteriole	0.5 - 1.0	
Capillary	0.02 - 0.17	
Inferior vena cava	15 - 40	

Table 13.1: Typical velocities in the human vascular system (data taken from [132], originally published in [151]).

change in the relative position of the scatterers, the change in the shape and size of the point-spread-function etc. Friemel and colleagues [148] have studied the decorrelation in speckle and found that for the velocities present in the body the peak cross-correlation value falls to 0.9, if the translation in elevation direction is Δz_{max} 0.37 mm. The maximum axial speed that can be measured in this case is:

$$v_{max} = \frac{\Delta z_{max} f_{prf}}{N}. \tag{13.10}$$

Assuming $N = 4$ and $f_{prf} = 7000$ Hz, one gets $v_{max} \approx 0.65$ m/s. The value for the lateral shift Δx_{max} for which the correlation value falls to 0.9 is reported to be 2.8 mm. Measuring at 45 degrees the maximum velocity which results in a peak correlation value of 0.87 – 0.9 is 0.9 m/s. Table 13.1 shows some typical values of velocities in the human circulatory system. Most of the velocities given are within the range, for which the peak of the cross-correlation[3] function is around 0.9.

The shift in position of the high-resolution images can be found from their cross-correlation:

$$\begin{aligned} R_{n-N,n}(\tau) &= \frac{1}{2T}\int_T H^{(n-N)}(t)H^{(n)}(t+\tau)\,dt \\ &= \frac{1}{2T}\int_T H^{(n-N)}(t)H^{(n-N)}(t+\tau-t_s)\,dt \\ &= R_{n-N,n-N}(\tau-t_s). \end{aligned} \tag{13.11}$$

The peak of the cross-correlation function is located at $\tau = t_s$. The velocity can be found from the time shift:

$$v = \frac{c}{2}\frac{t_s}{NT_{prf}}. \tag{13.12}$$

As a summary: the measurement principle is to estimate the time shifts between two high-resolution scan lines. The images from which the time shifts can be estimated must be acquired using the same transmit sequence, so they experience the same motion artifacts.

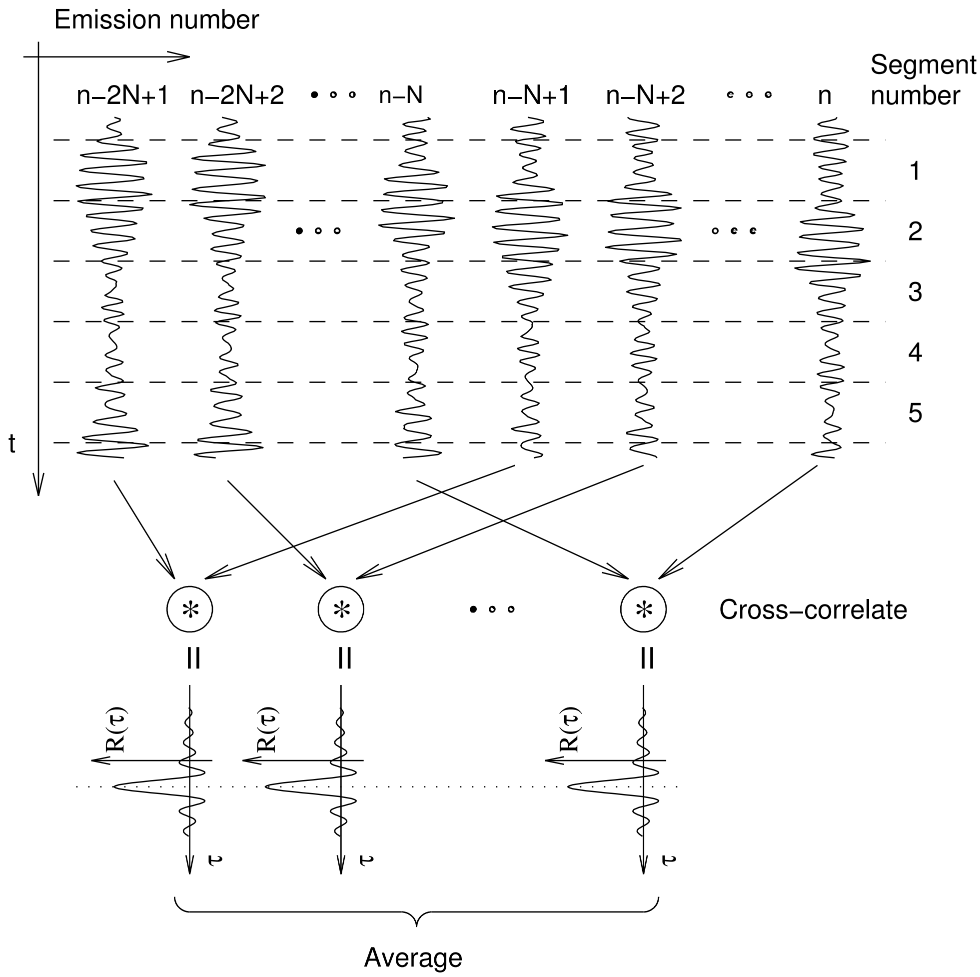


Figure 13.4: Estimation of the cross-correlation function from a number of high-resolution RF lines.

13.1.2 Estimation of the cross-correlation function

Figure 13.4 shows the process of estimating the cross-correlation function. Only a single RF line (along the same direction) from consecutive high-resolution images is shown. Two adjacent RF lines have decreased correlation due to motion artifacts caused by different transmit sequences. The same transmit sequence is used for RF lines that are separated by $N-1$ emissions. The estimation of the cross-correlation is like the standard procedure described in Chapter 10. The RF lines are divided into segments and the segments from the same depth are cross-correlated. The discrete version of the cross-correlation is:

$$\hat{R}_{1N}[n, i_{seg}] = \frac{1}{N_s} \sum_{k=0}^{N_s-1} H^{(1)}[k + i_{seg}N_s] H^{(N)}[k + i_{seg}N_s + n], \tag{13.13}$$

where N_s is the number of samples in the segment and i_{seg} is the number of the segment. To improve the estimate, some averaging must be done. Assuming that the velocity does not change significantly, the time shift t_s estimated from $\hat{R}_{1N}$ should be the same as the time shift

[3]Remember that the higher the peak cross-correlation, the bigger the accuracy and the smaller the standard deviation of the estimates are.

of the peak of the cross-correlation function $R_{2,N+1}$ as shown in Figure 13.4. The averaging can be expressed as:

$$\hat{R}_{1Nd}[n, i_{seg}] = \frac{1}{N_s(N_c-1)} \sum_{i=0}^{N_c-2} \sum_{k=0}^{N_s-1} H^{(i)}[k + i_{seg}N_s] H^{(i+N)}[k + i_{seg}N_s + n], \tag{13.14}$$

where N_c is the number of lines over which the averaging is done. The span of $\hat{R}_{12d}$ determines the maximum detectable velocity. In order to detect the same velocity range as the "conventional" cross-correlation estimator, the length of the interval in which the peak is searched must be N times bigger. If the search length is within the interval $[-N_s, N_s]$, then the largest detectable velocity becomes:

$$v_{max} = \frac{c}{2} N_s \frac{f_{prf}}{f_s}. \tag{13.15}$$

The minimum velocity is:

$$v_{min} = \frac{c}{2N} \frac{f_{prf}}{f_s}, \tag{13.16}$$

which is N times smaller than the minimum detectable velocity in the conventional cross-correlation estimators. In order to improve the estimates a second order polynomial is fitted to the three point around the peak of the cross-correlation function.

13.1.3 Stationary echo canceling

Usually the echo canceling is done by subtracting the adjacent lines [132, 143] or by using a high-pass filter with a very short impulse response. This way of processing is chosen because of the short data sequences available for the motion, typically $N_c \sim 8, 10$. Using a synthetic aperture imaging with recursive beamformation gives an uninterrupted data stream and filters with long impulse responses can be used. In this thesis the stationary echo canceling was done by subtracting the mean from the signal at a given depth:

$$H^{(k)}(t) = H^{(k)}(t) - \frac{1}{N_c+1} \sum_{i=k-\frac{N_c}{2}}^{k+\frac{N_c}{2}} H^{(i)}(t) \tag{13.17}$$

When the cross-correlated lines are not obtained in successive emissions, only every Nth line is used:

$$H^{(k)}(t) = H^{(k)}(t) - \frac{1}{N_c+1} \sum_{i=-\frac{N_c}{2}}^{\frac{N_c}{2}} H^{(k+iN)}(t). \tag{13.18}$$

13.2 Results

The performance of the suggested approach was tested for a flow with constant velocity as was done in Chapter 12 and for a flow which has a parabolic profile.

Parameter	Value	Unit
Transducer center frequency	7.5	MHz
Relative transducer bandwidth	66	%
Number of elements	64	
Excitation center frequency	5	MHz
Pulse repetition frequency	7000	Hz
Propagation velocity	1540	m/s
Center wavelength	308	μm
Excitation length	468	μm
Lines for one estimate	20	
Correlation length	3.1	mm
Search length	$\pm$468	μm
Step between two estimates	1	mm
Velocity	0.495	m/s

Table 13.2: Relevant parameters for the experimental verification of the new method.

13.2.1 Constant velocity

The performance of the method was tested on data measured using the XTRA system (see Appendix J for description of the system). The measurement setup was described in Section 12.4.1. The velocity estimation was applied on the very same data set. Only four transmit elements evenly distributed across the receive aperture ($N = 4$) were used in transmit. Two cases were investigated: (1) the standard cross-correlation estimation when the cross-correlation is done between successive lines, and (2) the new cross-correlation scheme in which the cross-correlation is done between lines that are separated by 3 emissions.

The parameters used in the velocity estimation are summarized in Table 13.2. These parameters were used for both cases of velocity estimation.

Figure 13.5 shows the values of the estimates as a function of emission number and depth. As it can be seen it completely misses the true estimates. The estimates very about the zero in both positive and negative directions. Figure 13.6 shows that the standard deviation is relatively small, which shows that the errors are due to the motion artifacts which are not random in nature. Figures 13.7 and 13.8 show the results from the velocity estimation using the modified cross-correlation estimation procedure. There are two relatively large errors in the beginning of the profile (between 15 and 20 mm) and at 60 mm. Looking at the reference image shown in Figure 12.9 on page 131 one can see that these two regions correspond to the strong wire scatterers that are placed along a line parallel to the transducer surface. As discussed in Chapter 10 the dominating source of error in the absence of noise, when all of the scatterers move with the same velocity is the error associated with the beam width modulation. When the scattering is diffused as in the case of speckle this phenomenon is less pronounced compared to the case when a single strong scatterer enters and exits the beam. The relative standard deviation is around 1 %, which is less even than the standard deviation of the reference velocity estimation in Chapter 12. The reasons for this are: (1) the time delay t_s is 4 times larger for the modified cross-correlation procedure and can be determined with a higher precision (see Figure 10.5), (2) the correlation interval is larger giving a better estimate of the cross-correlation function, and (3) more lines are used to average the estimated cross-correlation function. Having successfully

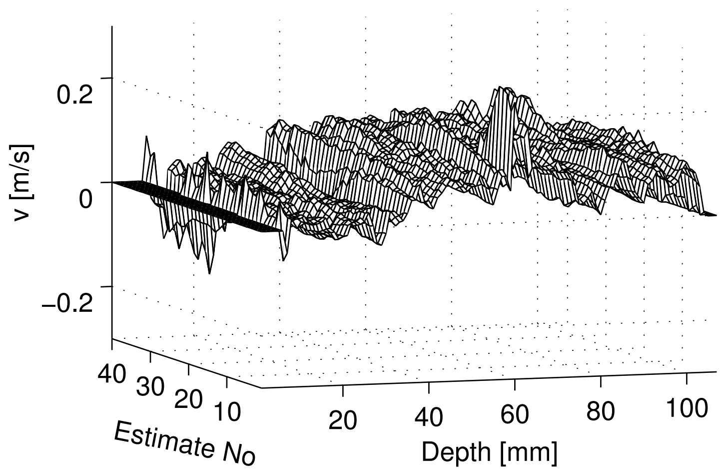


Figure 13.5: The result of velocity estimation using non-motion-compensated high-resolution images and a standard cross-correlation velocity estimation procedure. The estimated velocity as a function of emission number and depth is shown.

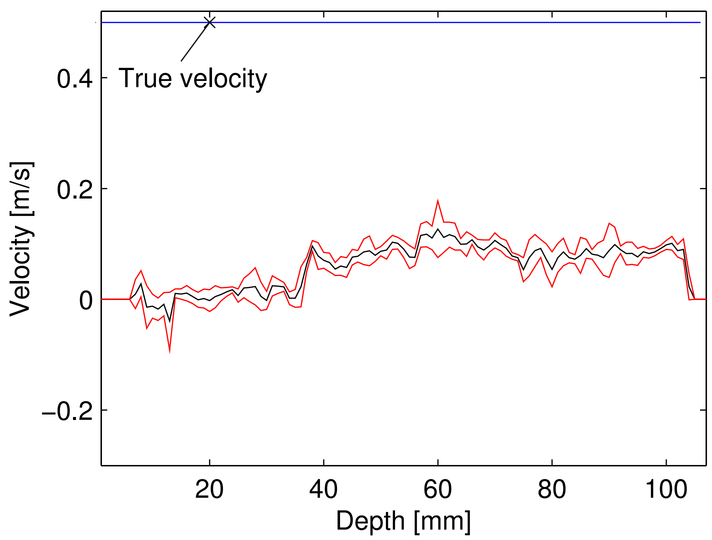


Figure 13.6: The result of velocity estimation using non-motion-compensated high-resolution images and a standard cross-correlation velocity estimation procedure. The figure shows the mean velocity and the region of $\pm$ one standard deviation.

estimated the velocity of a constant flow, the ability of the modified cross-correlation estimation procedure using synthetic aperture imaging for laminar flow with parabolic profile was tested.

13.2.2 Parabolic velocity profile

Developing a full model of the blood flow in the human circulatory system is beyond the scope of the thesis. Most articles on this subject assume a parabolic velocity profile and a steady flow

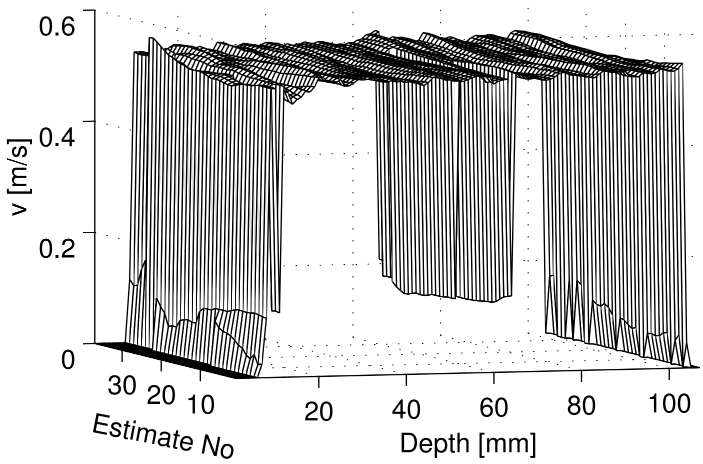


Figure 13.7: The result of velocity estimation using non-motion-compensated high-resolution images and the modified cross-correlation velocity estimation procedure.The estimated velocity as a function of emission number and depth is shown

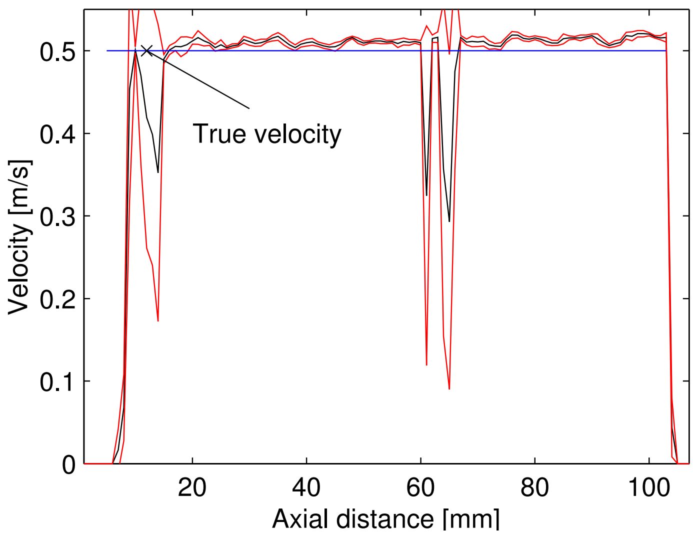


Figure 13.8: The result of velocity estimation using non-motion-compensated high-resolution images and the modified standard cross-correlation velocity estimation procedure. The figure shows the mean velocity and the region of $\pm$ one standard deviation.

(i.e. $\partial v/\partial t = 0$ at one particular position). This type of blood flow will be investigated here. Such velocity profile is developed in long and in narrow (with small radius) tubes. A comprehensive treatment of the subject can be found in Chapter 3, “Flow physics”, in *Estimation of blood velocities using ultrasound. A signal processing approach* by Jensen [132]. Figure 13.9 shows a typical measurement situation for measuring blood flow. A vessel with radius R is placed in front of a transducer. The angle between the vessel and the normal vector to the transducer surface is β. In this case the normal to the transducer surface coincides with the z

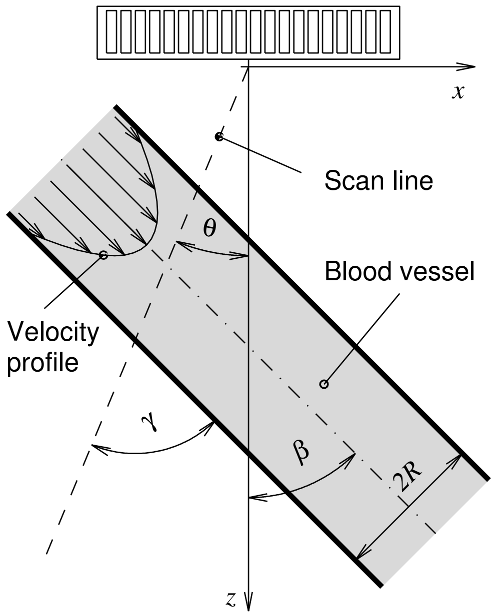


Figure 13.9: The simulation setup.

axis of the coordinate system. The blood flow is measured along a scan line that is defined by an azimuth angle θ (this is the angle between the scan line and the normal vector to the transducer surface). The velocity has a parabolic profile :

$$v(r) = v_{max}\left(1 - \left(\frac{r}{R}\right)^2\right), \tag{13.19}$$

where r is the distance to the center of the vessel and v_{max} is the maximum velocity in the center of the vessel. The distribution of the magnitude of the velocity is depicted in Figure 13.9 with arrows. The estimation procedure (13.12) gives only the velocity component along the scan line. The true velocity is found by:

$$\hat{v}_{true} = \frac{\hat{v}}{\cos\gamma}, \tag{13.20}$$

where γ is the angle between the scan line and blood vessel. The angle can be determined by using simple geometric relations as:

$$\gamma = \beta - \theta. \tag{13.21}$$

One must be careful with the definition of the angles. In the case shown in Figure 13.9 the positive angles are defined in the direction from the z axis towards the x axis.

Table 13.3 shows typical sizes of the larger vessels in human's vascular system. Their diameter $2R$ is close to 1 cm, and this is the size of the vessel simulated using Field II.

The vessel that was simulated had a diameter of 1 cm and a length of 5 cm. It closed an angle β of 45 $^\circ$ with the z axis. The center of the vessel was positioned on the z axis ($x = 0$, $y = 0$) 50 mm away from the transducer. To simulate the interaction with the blood 10 scatterers per cell of size $1 \times 1 \times 0.5$ mm were generated giving a total of 78549 scatterers. The positions of the scatterers were randomly selected with an uniform (white) distribution within the boundaries of the vessel. Their amplitudes had a Gaussian distribution with a zero mean [132]. Each

Vessel	Internal diameter [cm]	Wall thickness [cm]	Length [cm]
Ascending aorta	1.0 - 2.4	0.05 - 0.08	5
Descending aorta	0.8 - 1.8	0.05 - 0.08	20
Abdominal aorta	0.5 - 1.2	0.04 - 0.06	15
Femoral artery	0.2 - 0.8	0.02 - 0.06	10
Carotid artery	0.2 - 0.8	0.02 - 0.04	10-20

Table 13.3: Typical dimensions of some vessels in the human vascular system (data taken from [132], originally published in [151]).

Parameter	Value	Unit
Transducer center frequency	7.5	MHz
Relative transducer bandwidth	66	%
Number of elements	64	
Transducer pitch	205	μm
Transducer height	4	mm
Excitation type	chirp	
Excitation start frequency	0	
Excitation end frequency	17	MHz
Excitation duration	20	μs
Sampling frequency	100	MHz
Pulse repetition frequency	7000	Hz
Propagation velocity	1540	m/s
Maximum velocity	0.18	m/s

Table 13.4: Relevant parameters for the simulation of the parabolic flow.

scatterer had a velocity associated with its radial position (relative to the central axis of the vessel). Between every two emissions the scatterers were translated in a direction parallel to the walls of the transducer. Every scatterer was displaced at a distance corresponding to its velocity. The scatterers that “fell out” of the vessel were put in again from the opposite side (cyclically moved).

The most important parameters of the simulation are listed in Table 13.4.

One of the serious problems in synthetic aperture imaging is the poor signal to noise ratio. This makes it almost imperative to use coded excitation, and that is why such an excitation was used in the simulations. The excitation was a linear FM chirp, sweeping the frequency range from 0 to 17 MHz in 20 μs. Such a chirp cannot pass unaltered through the transducer. The applied filter is not the matched filter, but is rather a filter optimized with respect to minimizing the range side-lobe level [119]. The chirp, the filter, and the logarithmically compressed envelope of the output of the filter are given in Figure 13.10. The output of the filter was created by convolving the excitation with the impulse response of the transducer twice, and then convolving it with the impulse response of the filter. The low level of the range side lobes guarantees that there will be a minimum influence on the estimates from signals coming from the stationary echoes

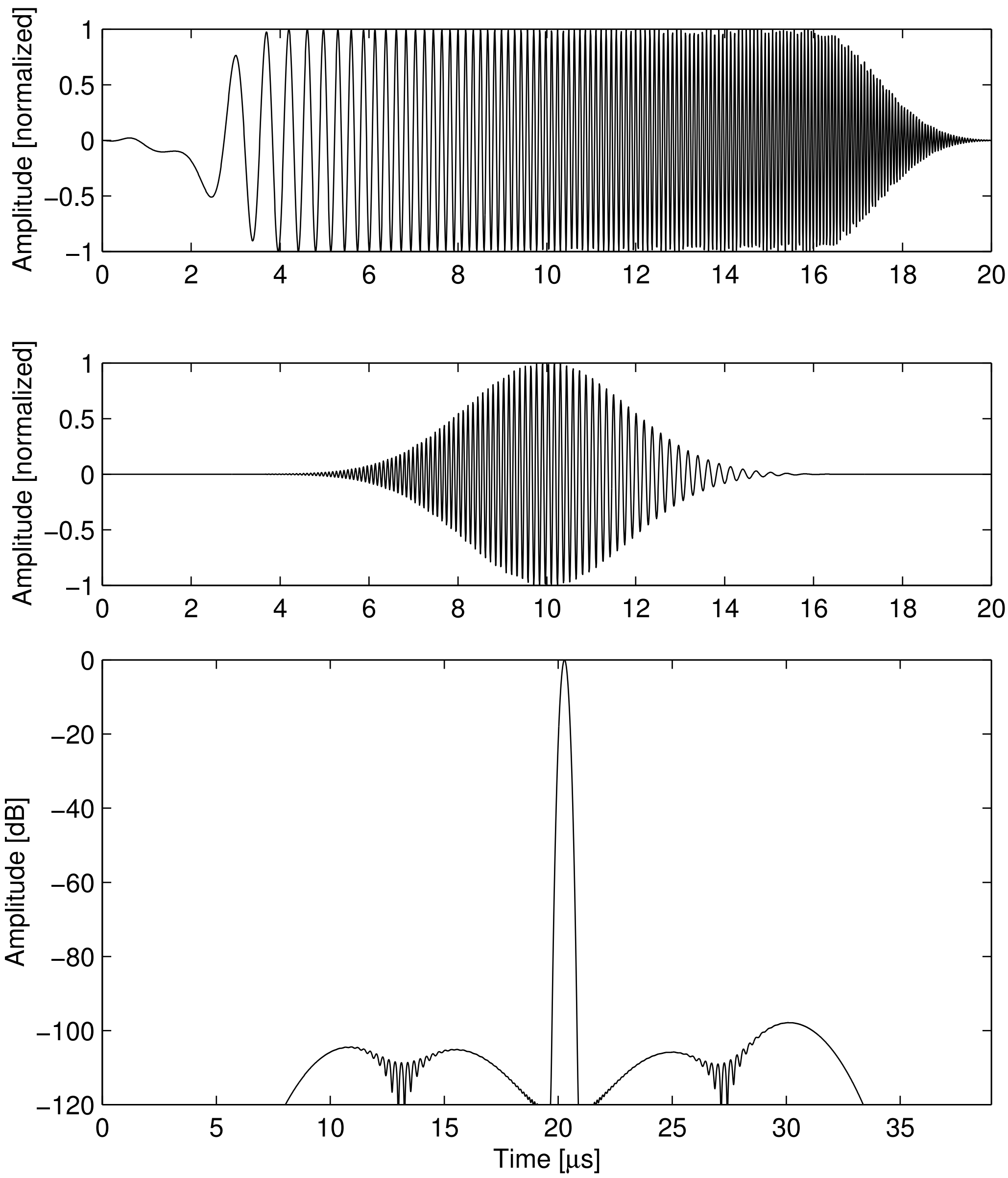


Figure 13.10: The excitation used in the simulation (top), the applied filter (middle), and the resulting output (bottom) are given. The output is a result of applying the filter on a signal obtained by convolving the excitation with the impulse response of the transducer twice.

and from other segments of the blood vessel.

The estimates using cross-correlation can be improved by two means: (1) using longer data segments (in depth, or equivalently fast time) and (2) using more lines to average the cross-correlation estimate. In the former case the correlation length was set to 5λ, where λ is the wavelength at the center frequency of the transducer. In the simulation λ is equal to 205 μm. The number of lines over which the cross-correlation function is averaged is $N_c = 4$. The interval in which the maximum of the cross-correlation function is sought is restricted to $\pm 2.5\lambda$. In the latter case, when $N_c = 24$, the correlation length is 2.5λ and the search interval is set to $\pm 1.7\lambda$.

The high-resolution images were obtained using only four transmit elements evenly distributed

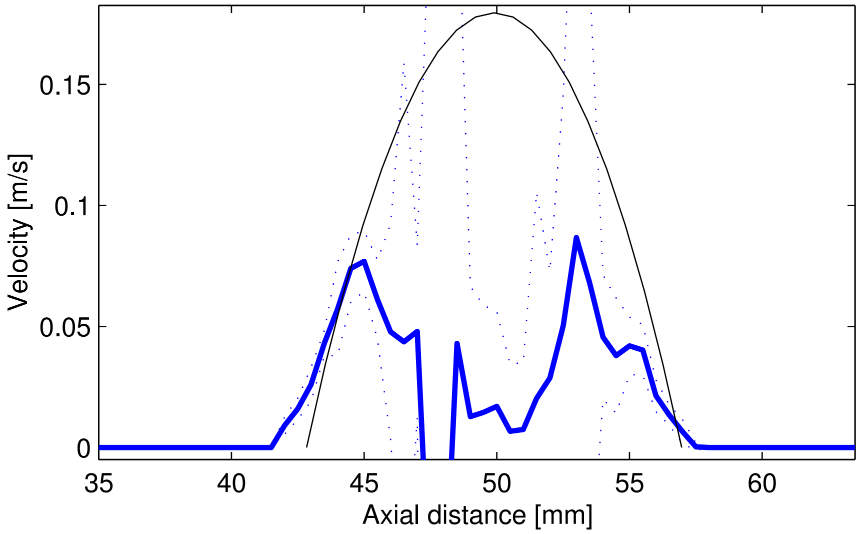


Figure 13.11: Velocity profile at $\gamma = 45^\circ$ using the traditional cross-correlation velocity estimation. The number of lines per estimate is $N_c = 24$, and the segment length is 2.5λ.

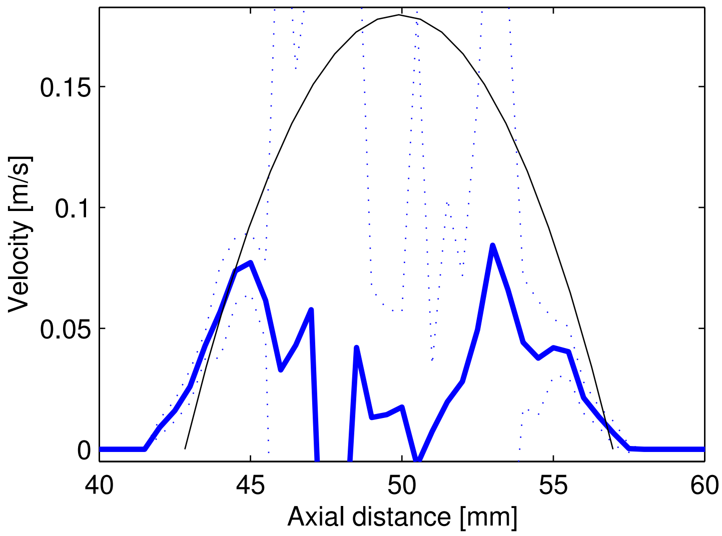


Figure 13.12: Velocity profile at $\gamma = 45^\circ$ using the traditional cross-correlation velocity estimation. The number of lines per estimate is $N_c = 4$, and the segment length is 5λ.

across the aperture. The modified cross-correlation velocity estimation was compared to the traditional approach, the results of which are given in Figures 13.11 and 13.12. The thin black lines outline the true velocity profile. The thick blue lines give the mean of the estimate, and the thin blue dotted lines show enclose the region of plus/minus one standard deviations.

Figures 13.11 and 13.12 show the cases in which the number of lines are $N_c = 24$ and $N_c = 4$, respectively. In order for the results to be comparable to the results when using the modified cross-correlation approach, the search length was 4 times smaller (the scatterers travel 4 times shorter distance). It can be seen that for small velocities, the motion artifacts do not distort the signal significantly, the correlation between successive high-resolution images is relatively high, and the velocity can be determined. For higher velocities, however, the effect of the mo-

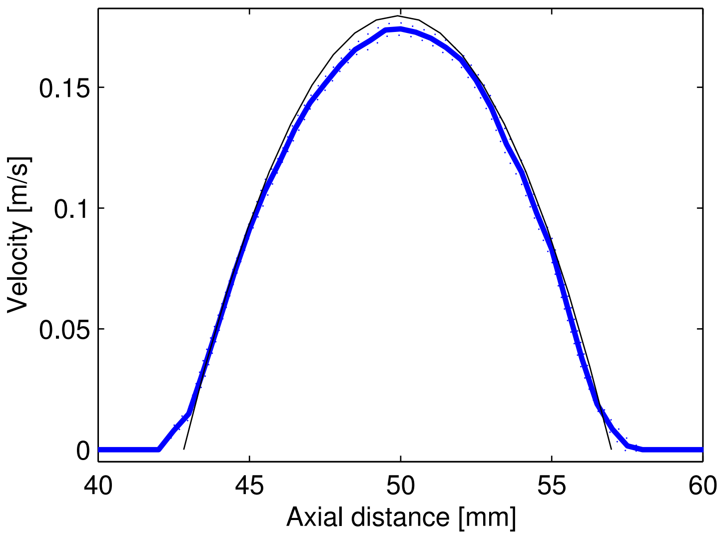


Figure 13.13: Velocity profile at $\gamma = 45°$ using the modified cross-correlation velocity estimation. The number of lines per estimate is $N_c = 24$, and the segment length is 2.5λ.

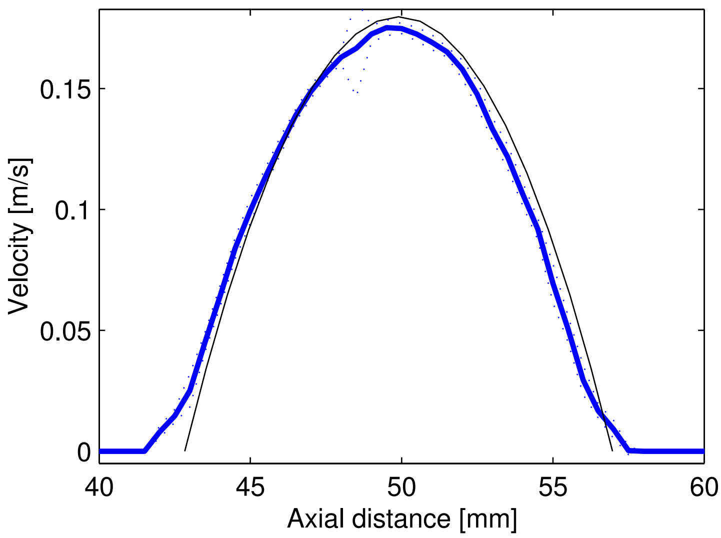


Figure 13.14: Velocity profile at $\gamma = 45°$ using the modified cross-correlation velocity estimation. The number of lines per estimate is $N_c = 4$ and the segment length is 5λ.

tion artifacts dominates the estimated velocity (like a negative profile around a certain pedestal). The standard deviation is also large. The mean values of the estimated velocity profiles are similar showing that the velocity estimation is influenced mainly by the motion artifacts which have deterministic character.

Figures 13.13 and 13.14 show the velocity profiles using the modified cross-correlation estimation procedure. The maximum bias (at the top of the profile) of the profile in Figure 13.13 is -3.05 % and the maximum standard deviation is 2.7 %. The exciting part is that velocity profile in Figure 13.14 has a similar performance, but is obtained only with 8 emissions (if $N_c = 4$ then RF lines from the images $H^{(1)}$ to $H^{(4)}$ are cross-correlated with RF lines in images $H^{(5)}$ to $H^{(8)}$). The bias of the maximum of the estimated profile with respect to the maximum of

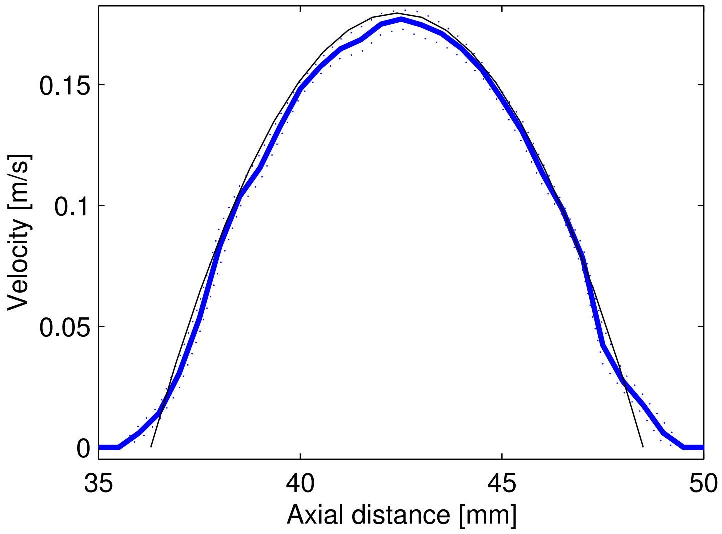


Figure 13.15: Velocity profile at $\gamma = 35°$ using the modified cross-correlation velocity estimation. The number of lines per estimate is $N_c = 4$, and the segment length is 5λ.

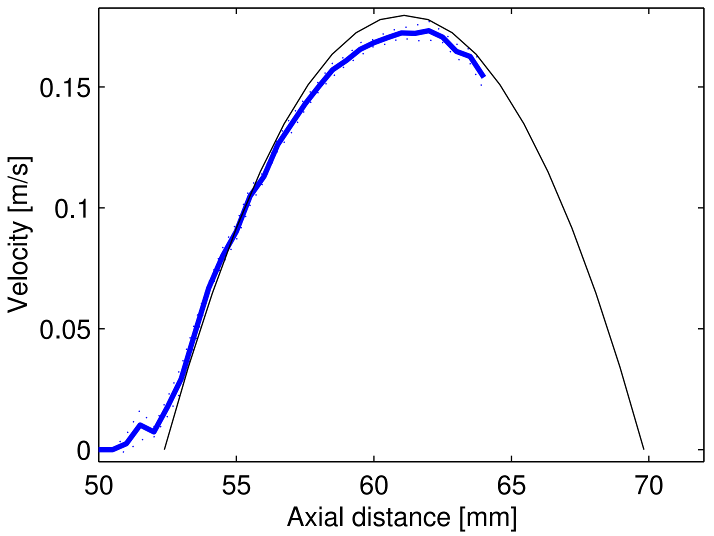


Figure 13.16: Velocity profile at $\gamma = 55°$ using the modified cross-correlation velocity estimation. The number of lines per estimate is $N_c = 4$, and the segment length is 5λ. The signal was range gated too-early and the estimates do not cover the whole vessel.

the true profile is -2.5 %. The standard deviation varies between 1.2 % at 50th mm and 5 % at 42d mm. The high value of 12.1 % of the standard deviation at depth of 48 mm is caused by a single false peak detection - magnitude of -0.17 m/s. If this peak is removed the standard deviation at this point becomes 2.15 %.

The fact that the velocity can be estimated from synthetic aperture imaging using only 8 emissions gives the foundations to develop velocity estimations for real-time 3D scanners. Scanning the volume plane by plane, one could use 8 emissions per plane. In such a way velocity estimates as well as B-mode data will be available for the whole volume. The performance of the velocity estimation using only 8 emissions for angles $\gamma = 35°$ and $gamma = 55°$ is shown in Figures 13.15 and 13.16, respectively.

In Figure 13.16 the estimated velocity profile is not complete because the received raw RF signals by the individual elements were range gated.

13.3 Discussion and prospectives of the velocity estimation with synthetic aperture imaging

In order to estimate the velocity in conventional ultrasound scanners, the pulse is transmitted several times in the same direction. This leads inevitably to a decrease in the frame rate. Because of its nature, synthetic transmit aperture ultrasound imaging, or more precisely its variation recursive ultrasound imaging generates high-resolution RF lines at every emission. This gives practically uninterrupted flow of data. The estimates can be based on larger data segments thus improving the precision of the estimates. It was also shown that the estimates can be made with only a few emissions as they are normally done in conventional systems. This opens up for the possibility to have real-time velocity estimates in real-time 3D systems, thus making these systems full-featured ultrasound scanners, not only baby face renderers as they are in the present moment.

This chapter showed that in order tow work with synthetic aperture imaging, the velocity estimation algorithms must be altered. The class of time-shift measurement systems is suitable for the purpose. In the following a brief overview of some of the most promising approaches will be treated.

13.3.1 Velocity estimation using cross-correlation along the blood vessel

The estimation of the velocity using cross-correlation of transverse beams was first suggested by Bonnefous [152]. It was further developed and perfected by Jensen and Lacasa [153, 154]. Figure 13.17 shows the measurement principle. The beam is focused along lines parallel to the axis of the vessel. Each of these lines from emission n is cross-correlated with the line from emission $n+1$. The offset of the peak of the cross-correlation function is a measure of the movement of the scatterers. The traditional cross-correlation approach traces the signals towards and away from the transducer. Therefore only the axial velocity components can be found. With this method the signal is traced along the blood vessel, and velocity components parallel to the transducer surface can be measured. In the initial work [153] a broad beam was sent in the tissue and in receive the delays were set to focus along the axis of the vessel. The method has two problems: (1) it is not equally sensitive in the axial and lateral directions, and (2) the beamprofile changes as a function of space, thus affecting the accuracy of the estimates. The latter of the problems was addressed by Jensen and Gori [46] by using matched filters. The filters have different impulse responses depending on the varying spatial impulse response (which can be numerically found), thus compensating for the fixed focus in transmit. This approach can be extended to synthetic aperture imaging in a straight-forward manner.

The beams are perfectly focused in transmit and receive. As seen from Figure 11.8 on page 123 a lateral translation of the scatterer causes a lateral translation of the low-resolution image. If one considers a laminar flow, then all the scatterers located at a radial distance r from the center of the vessel move with a constant velocity v. The distance which they travel for the time T_{prf}

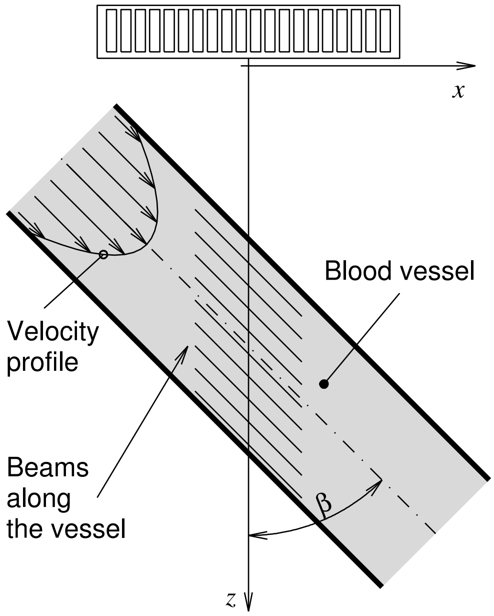


Figure 13.17: The data used to find the velocity is taken from lines along the vessel.

between every two emissions is

$$\Delta l = |\vec{v}| T_{prf}. \tag{13.22}$$

The traveled distance can be decomposed into an axial and lateral components (see: Figure 13.17):

$$\begin{aligned} \Delta x &= \Delta l \sin\beta \\ \Delta z &= \Delta l \cos\beta \end{aligned} \tag{13.23}$$

Using the arguments from Section 13.1.1 which are illustrated for the case of pure axial motion in Figure 13.3 one gets that:

$$\frac{1}{2}\mathbf{P}^{(n)}(x,z;(x_f^{(n)},z_f^{(n)})) = \mathcal{T}[N\Delta x, N\Delta z]\mathbf{P}^{(n-N)}(x,z;(x_f^{(n-N)},z_f^{(n-N)})) \tag{13.24}$$

In this case it is straight forward to show that a line $s^{(n)}(x,y)$ at emission n, defined along the beam:

$$\begin{aligned} x &= x_f + l\sin\beta \\ z &= z_f + l\cos\beta \end{aligned} \tag{13.25}$$

is a shifted version of the line $s^{(n-N)}(x,y)$:

$$s^{(n)}(x,y) = s^{(n)}(x - N\Delta x, y - N\Delta y). \tag{13.26}$$

It is convenient to express the line in terms of l, which in this case is distance from the focal point x_f, z_f along the direction defined by the angle β. The relation becomes:

$$s^{(n)}(l) = s^{(n-N)}(l - N\Delta l). \tag{13.27}$$

Cross-correlation the two lines gives:

$$R_{1N}(\tau) = R_{11}(\tau - \Delta l), \tag{13.28}$$

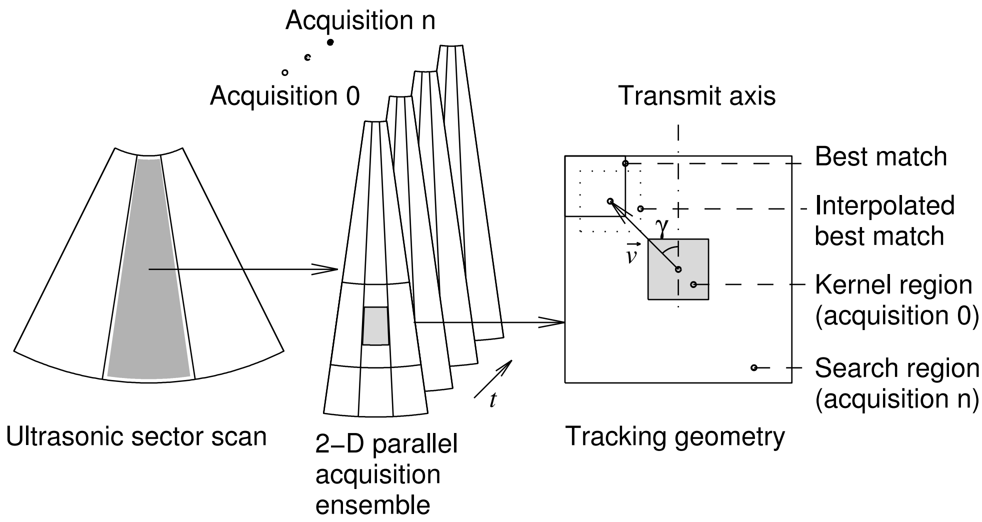


Figure 13.18: 2D vector velocity estimation using speckle tracking (ensemble tracking).

where R_{11} is the auto-correlation function of the signal, and τ is a lag in space. The velocity is then:

$$\hat{v} = \frac{\Delta l}{N T_{prf}}. \tag{13.29}$$

The minimum detectable velocity is dependent on the spacing between the samples comprising the scan line $s(t)$. This approach solves partly the problem with decorrelation due to migration of scatterers. More information on the subject can be found in [154].

13.3.2 Velocity estimation using speckle tracking

Estimating the blood velocities by tracking the speckle produced by moving blood has been suggested by Trahey and colleagues in 1987 [155, 156]. The idea is that the current random distribution of blood cells creates certain speckle patterns (two dimensional image). For the distance that the blood cells travel for the time of a few (usually 5 to 10) emissions these patterns remain intact and the change in their position can be traced by using a two dimensional convolution. The two-dimensional cross-correlation procedure has been simplified in subsequent papers [157], and made more robust by tracking the speckle pattern within smaller regions and using parallel beamforming [158, 159, 160].

The velocity estimation using speckle tracking (called in the recent papers "ensemble tracking" [160]) is illustrated in Figure 13.18. Several acquisitions are done in the same direction. In receive several lines within the boundaries of the transmit beam are formed in parallel. To improve the performance the RF data from the receive beams are rescaled to compensate for the lateral transmit beam profile. A small 2D kernel is chosen and a match for it is sought in a larger bounding 2D region. The pattern matching algorithm that was used in [160] is based on the so-called sum absolute difference (SAD) [157]:

$$\varepsilon(\eta,\xi,n) = \sum_{i=1}^{l}\sum_{j=1}^{k} |B_0(i,j)| - |B_n(i+\eta, j+\xi)|, \tag{13.30}$$

where ε is the SAD coefficient, n is the acquisition number, $B_n(i,j)$ is the brightness of the pixel at location (i, j) in the nth image, l is the lateral dimension of the kernel in pixels, k is the axial dimension of the kernel in pixels, and η and ξ are lateral and axial pixel offsets of a prospective matching region within the search region. The best match is found at the smallest value of the difference. The result of the process is interpolated in a fashion similar to the curve fitting shown in Figure 10.4, and the interpolated offsets $(\hat{\eta}_m, \hat{\xi}_m)$ at which the maximum occurs are found. The magnitude of the velocity is given by:

$$\begin{aligned} |\vec{v}| &= \frac{\sqrt{(\hat{\eta}_m \Delta x)^2 + (\hat{\xi}_m \Delta z)^2}}{n T_{prf}} \\ \gamma &= \arctan \frac{\hat{\eta}_m \Delta x}{\hat{\xi}_m \Delta z}, \end{aligned} \tag{13.31}$$

where γ is the angle between the beam axis and the velocity vector, Δx and Δz are the lateral and axial spatial sampling intervals respectively, and T_{prf} is the pulse repetition interval.

This approach can be modified for use with synthetic aperture imaging in the same way as the cross-correlation velocity estimators. The speckle patterns from emissions 0 and N, 1 and $N+1 \cdots N-1$ and $2N-1$ can be tracked and then the estimates can be averaged for higher accuracy.

In conclusion, this chapter showed that images obtained with the same sequence of transmissions over a short time period ($\approx$ 1 ms) exhibit high correlation. Many of the time-domain velocity estimation algorithms can be modified for use with synthetic aperture imaging. This chapter showed how to modify the velocity estimator using time shift measurement (aka cross-correlation velocity estimation) and through simulations showed that the amount of emissions per color flow map can be reduced to eight. If a 3D scan is performed plane by plane using 8 emissions per plane. a color flow map for each of the planes can be created and combined in a color flow map for the whole volume.

CHAPTER

FOURTEEN

Conclusion

The purpose of the Ph.D. project was to investigate and design beamforming procedures and ways of collecting data for real-time 3D imaging. Previous work in the field included the ”exploso-scan” - a method for wide beam transmission and parallel beams reception, and synthetic aperture focusing. After initial trials and assessment, it was concluded that the synthetic aperture imaging has the potential to produce images at the required frame rate and image quality. SAU has the advantage of producing images that are focused at every point both in transmit, and in receive. The acquisition time *per frame* can be significantly reduced by using sparse synthetic transmit aperture. This makes the real-time three-dimensional imaging using linear 1D and matrix 2D arrays possible.

Possible, though, there have not been previous reports of clinical trials using synthetic transmit aperture imaging because of various problems unsolved at the time when my Ph.D. study commenced. This dissertation has addressed several of the problems:

- The poor resolution in the elevation plane, when the 3D scanning is performed with a linear array.
- The high side and grating lobe levels when sparse arrays are used.
- The low data rate, which is not adequate for velocity estimation.
- The motion artifacts present in the image.
- The inability of the current algorithms to estimate the velocity in the presence of motion artifacts.

Some of the first *in-vivo* images using synthetic transmit aperture focusing have been acquired in the course of the Ph.D. study. To make this multiple elements have been used in transmit as suggested by other researchers. The central of the elements emits first and the outermost elements emit last, thus creating a diverging wave. The approach was further developed in Chapter 7 by unifying the use of multiple transducer elements, focused and defocused in the common concept of the virtual ultrasound sources. This concept is extended also to the 3D case. A linear array focused in the elevation plane is used to scan a whole volume. The resolution in the elevation plane is improved more than 3.5 times using synthetic aperture post focusing based on virtual ultrasound sources.

In order to decrease the motion artifacts and demands on hardware, sparse synthetic transmit aperture is used. The main problem is the selection of active elements. Chapter 8 describes the work done in the field of sparse array design, and suggests two novel designs for 2D sparse arrays. The new designs show improved performance in terms of main lobe to side lobe energy ratio, while preserving the maximum obtainable resolution.

A new approach called ”recursive imaging”, which decreases the amount of calculations per frame is suggested in Chapter 6. Using this beamformation technique new RF lines are formed at every emission for *all* imaging directions, something not possible in conventional imaging. The data flux is practically uninterrupted and long data sequences are available for the whole region of investigation. This makes it possible to create color flow maps using synthetic aperture data, something regarded as impossible until recently.

The existing velocity estimators cannot estimate the velocity from the data beamformed using recursive ultrasound imaging, unless motion compensation is applied, or new estimators are developed. An original model of the motion artifacts is developed in Chapter 11, showing that the motion compensation can be performed during the summation of the low-resolution images. This approach was applied on experimental data, and proved to be a viable method for motion compensation. The velocity can be estimated from the compensated data as shown in Chapter 12. The precision of the velocity estimation depends on the quality of motion compensation and vice versa. The presence of noise worsens the velocity estimations and potentially could lead to the failure of the whole scheme. This problem is solved in Chapter 13 where a modified velocity estimator is developed on the basis of the cross-correlation velocity estimator. It removes the necessity of having the motion compensation prior to the velocity estimation. Chapter 13 ends with a discussion of how other velocity estimation algorithms can be modified and used with synthetic aperture imaging.

Apart from the problems of acquisition and beamformation some other practical aspects such as the display of 3D ultrasound data are also considered in the dissertation. While some previous works use OpenGL for the final stage of the visualization, the approach given in Chapter 2 uses OpenGL for *scan-conversion*. All the work on the scan conversion is done by a 3D-accelerated video controller thus reducing the load on the main processor to a minimum. This gives the possibility to use software, rather than hardware solutions, thus decreasing the development time and cost of ultrasound scanners.

In short, the contributions of dissertation are:

- New designs of sparse 2D arrays are suggested.
- An unified concept of virtual ultrasound sources is outlined and applied to 3D ultrasound imaging with linear array.
- A new imaging procedure -“recursive ultrasound imaging” - making it possible to create a new high-resolution is developed and tested.
- A simple geometric model of the motion artifacts in synthetic aperture ultrasound is created.
- A new motion compensation scheme is successfully applied on measured ultrasound data
- A modified velocity estimator is suggested and tested with measurements and simulations.
- A practical approach of scan conversion of 2D and 3D images using off-the-shelf computer components is suggested and developed.

The thesis has covered most of the aspects associated with synthetic aperture imaging, as well as some aspects of the acquisition and the display of 3D volumes. The successful initial trials show the viability of implementing real-time 3D imaging using synthetic aperture focusing.

Part III

Appendix

APPENDIX

A

Fresnel diffraction between confocal spherical surfaces

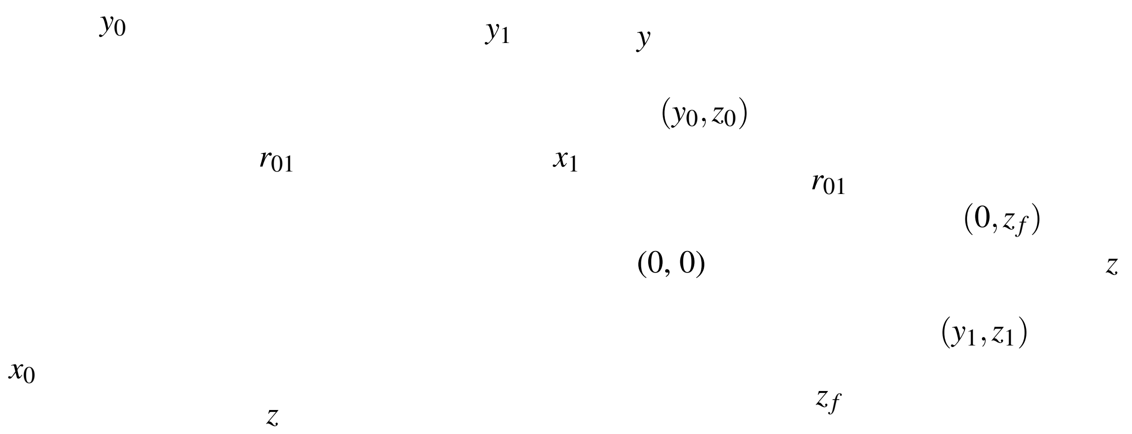


Figure A.1: Confocal spherical surfaces. On the left the 3-D case is shown. On the right a cut in the $(y - z)$ plane is shown.

The ultrasound imaging is performed in the near field using focused transducers. The focusing can be either mechanical or electronic. In the latter case the focus can be changed as a function of depth. This case corresponds to the case of diffraction between two confocal spherical surfaces such as the shown in Figure A.1. While in [23] Goodman describes the result of such a diffraction, here the full derivation is given.

The surface described by (x_0, y_0, z_0) radiates a continuous wave, and the field is studied along the spherical surface (x_1, y_1, z_1). The center of each of the spheres (x_0, y_0, z_0) and (x_1, y_1, z_1) lies on the surface of the other. In the depicted case the surface of the sphere (x_0, y_0, z_0) is tangent to the (x, y) plane at point $(0, 0, 0)$. The surface (x_1, y_1, z_1) is tangent to a plane parallel to the (x, y) plane lying at a distance z_f. The distance z_f is the focal depth. For simplicity consider only the two dimensional case depicted on the right side of Figure A.1. The two surfaces can be expressed as:

$$
\begin{aligned}
z_0 &= z_f - \sqrt{z_f^2 - y_0^2} \\
z_1 &= \sqrt{z_f^2 - y_1^2}
\end{aligned}
\tag{A.1}
$$

The distance r_{01} between the two points (y_0, z_0) and (y_1, z_1) is

$$
\begin{aligned}
r_{01} &= \sqrt{(y_1 - y_0)^2 + (z_1 - z_0)^2} \\
r_{01} &= \sqrt{(y_1 - y_0)^2 + \left(z_f - \sqrt{z_f^2 - y_0^2} - \sqrt{z_f^2 - y_1^2}\right)^2}
\end{aligned}
\tag{A.2}
$$

A simplification of the terms yields:

$$
r_{01} = \sqrt{3z_f^2 - 2y_0y_1 - 2z_f\sqrt{z_f^2 - y_0^2} - 2z_f\sqrt{z_f^2 - y_1^2} + 2\sqrt{z_f^2 - y_0^2}\sqrt{z_f^2 - y_1^2}} \tag{A.3}
$$

Assuming that $y_0^2/z_f^2 \ll 1$ and $y_1^2/z_f^2 \ll 1$ the radicals are expanded into Tailor series giving:

$$
r_{01} \approx \sqrt{3z_f^2 - 2y_0y_1 - 2z_f^2\left(1 - \frac{y_0^2}{2z_f^2}\right) - 2z_f^2\left(1 - \frac{y_1^2}{2z_f^2}\right) + 2z_f^2\left(1 - \frac{y_0^2}{2z_f^2}\right)\left(1 - \frac{y_1^2}{2z_f^2}\right)} \tag{A.4}
$$

$$
r_{01} \approx \sqrt{z_f^2 - 2y_0y_1 + \frac{y_0^2y_1^2}{2z_f^2}} \tag{A.5}
$$

$$
r_{01} \approx z_f - \frac{y_0y_1}{z_f} + \underbrace{\frac{y_0^2y_1^2}{4z_f^3}}_{\approx 0} \tag{A.6}
$$

$$
r_{01} \approx z_f - \frac{y_0y_1}{z_f} \tag{A.7}
$$

In order for (A.7) to be fulfilled either the transducer must be weakly focused ($\left(\frac{\max(y_0)}{z_f} \ll 1\right)$ or the observation interval must be small $\left(\frac{\max(y_1)}{z_f} \ll 1\right)$. Thus, for weakly focused transducers (A.7) is valid for a bigger off-axis distance, and for strongly focused transducers this approximation is valid over a narrow region.

Provided that these conditions are met, and going back to the 3-D case one gets:

$$
r_{01} \approx z_f - \frac{x_0x_1}{z_f} - \frac{y_0y_1}{z_f} \tag{A.8}
$$

Substituting (A.8) in the equation of the Huygens-Fresnel principle:

$$
\Phi(x_1, y_1; z) = \frac{z}{j\lambda} \iint_{\Sigma} \Phi(x_0, y_0; 0) \frac{\exp\left(j\frac{2\pi}{\lambda} r_{01}\right)}{r_{01}^2}\, dx_0\, dy_0
$$

one gets the following Fresnel approximation of the field distribution for the focal zone:

$$
\boxed{\Phi(x_1, y_1; z_f) = \frac{e^{jkz_f}}{j\lambda z_f} \iint_{-\infty}^{\infty} \Phi(x_0, y_0; 0) e^{-j\frac{2\pi}{\lambda z_f}(x_1x_0 + y_1y_0)}\, dx_0\, dy_0} \tag{A.9}
$$

This equation aside from the multipliers and scale factors, expresses the field observed on the right-hand spherical cap as the Fourier transform of the field on the left-hand spherical cap.

Since the image coordinates (x_1, y_1) lie on a spherical surface, it is often more convenient to express the result in spherical coordinates, and then the radiation pattern becomes a function of the elevation and azimuth angles.

APPENDIX

B

Application of different spatial sampling patterns for sparse array transducer design.

Authors : Svetoslav Ivanov Nikolov and Jørgen Arendt Jensen

Presented : Paper presented at the Ultrasonics International '99, Copenhagen, Denmark, 1999. Published in Ultrasonics, February, 2000.

Abstract

In the last years the efforts of many researchers have been focused on developing 3D real-time scanners.

The use of 2D phased-array transducers makes it possible to steer the ultrasonic beam in all directions in the scanned volume. An unacceptably large amount of transducer channels (more than 4000) must be used, if the conventional phased array transducers are extrapolated to the two-dimensional case. To decrease the number of channels, sparse arrays with different aperture apodization functions in transmit and receive have to be designed.

The design usually is carried out in 1D, and then transfered to a 2D rectangular grid. In this paper 5 different 2D array transducers have been considered and their performance was compared with respect to spatial and contrast resolution. An optimization of the element placement along the diagonals is suggested. The simulation results of the ultrasound fields show a decrease of the grating-lobe level of 10 dB for the diagonally optimized 2D array transducers.

1 Introduction

To obtain a three-dimensional scan of the body, an ultrasound scanner must be able to focus in any direction of the interrogated volume. This can be obtained either by mechanically rocking a focused transducer or by electronically steering the ultrasound beam as shown in Fig. B.1. The latter makes it possible also to implement parallel receive beamforming and to do the scanning

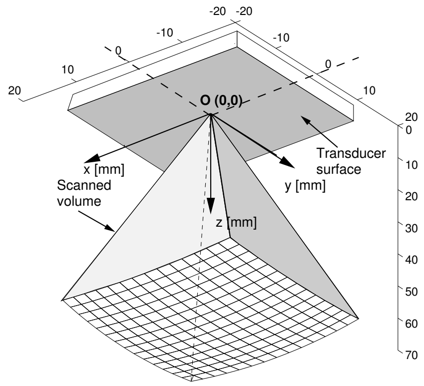


Figure B.1: Volumetric scanning. The center of the coordinate system is in the middle of the transducer surface.

in real time [92]. The system is considered to be linear, and can thus be characterized by its point-spread-function (PSF). Ideally the PSF is a spatial δ function. Side-lobes are present in the radiation pattern due to the finite size of the transducers . The periodic nature of the linear arrays introduces grating lobes. The grating lobes are outside the scanned region if the spacing between the elements is $\lambda/2$. To obtain high resolution with a small number of channels, arrays with a pitch of $p > \lambda/2$ must be used and the grating lobes enter the viewed field.

Randomly sparsed arrays do not have a periodical sampling pattern and thus they do not have prominent grating lobes in the radiation pattern. However, a pedestal of side-lobe energy at level of approximately -30 dB from the peak value is present. Although some optimization can be applied on the weighting coefficients [105], the performance can not be increased to a level comparable to the dense arrays.

The ultrasound system must be evaluated in terms of the two-way (transmit and receive) radiation pattern. A formal approach is introduced by the use of the *coarray* [52] also known as *effective aperture* [50].

The effective aperture is briefly introduced in Section 2 and Section 3 shows how it can be used to design linear transmit and receive apertures. In Section 4 the design is extended to 2D. Section 5 gives the radiation patterns of these apertures obtained by computer simulations of ultrasound fields.

2 Effective aperture concept

The effective aperture of an array is the aperture that has a radiation pattern identical to the two-way (transmit and receive) radiation pattern of the array. The connection between the array aperture function $a(x/\lambda)$ and the radiation pattern in the far field and in the focal region $P(s)$ is given by the Fourier transform [92]:

$$P(s) = \int_{-\infty}^{+\infty} a\left(\frac{x}{\lambda}\right) e^{j2\pi(x/\lambda)} d\left(\frac{x}{\lambda}\right) \tag{B.1}$$

where the aperture function describes the element weighting as a function of the element position, $s = \sin\phi$, ϕ is the angle measured from the perpendicular to the array, and x/λ is the

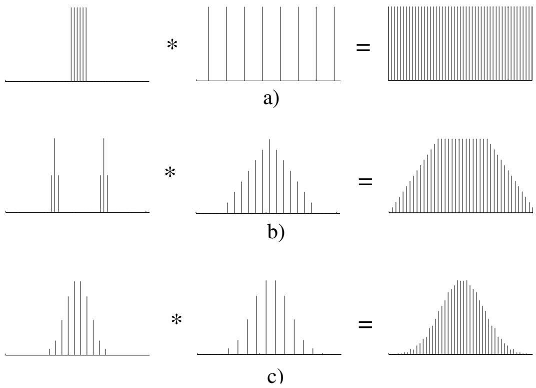


Figure B.2: Transmit, receive and effective apertures. The resulting effective aperture, from top to bottom, has rectangular, triangular and cosine^2 apodizations

element location in wavelengths. The two way radiation pattern is

$$P_{TR}(s) = P_T(s)P_R(s) \tag{B.2}$$

The radiation pattern of the effective aperture can be expressed as a spatial convolution of the transmit and receive apertures.

$$E(x/\lambda) = a_T(x/\lambda) * a_R(x/\lambda) \tag{B.3}$$

The design of the transmit and receive apertures is thus reduced to the problem of finding a suitable effective aperture with a desired Fourier transform. The elements in the effective aperture must be spaced at $\lambda/2$, and the weighting function shouldn't have discontinuities to avoid the side and grating-lobes. Since the radiation pattern is the Fourier transform of effective aperture, it is convenient to exploit the properties of the classical windowing functions : rectangular, triangular and hamming. These functions are separable, and the design can be carried out in 1D and then extended to 2D.

3 Aperture design strategies in 1D

Fig. B.2 shows three different design strategies leading to an effective aperture with $\lambda/2$ spaced elements.

A simple approach, shown in Fig. B.2 a) is to select a dense transmit aperture with N_{xmt} elements. Its width is $D_{xmt} = N_{xmt}\lambda/2$. The receive aperture must then have spacing between its elements $d_{rcv} = (N_{xmt} - 1)\lambda/2$. Hereby a fully populated effective aperture is obtained with the minimum number of elements in the transmit and receive apertures. Because of the rectangular shape of the apodization function, of the effective aperture, this design will be further referred to as "rectangular approximation".

Fig. B.2 b) shows how an apodized effective aperture can be obtained by introducing some redundancy in the number of transmit and receive elements. From the properties of the Fourier transform it is expected that this design has lower side-lobes than the design depicted in Fig. B.2 a). Further in the paper this kind of effective aperture will be referred to as "triangular", and the design as "triangular approximation".

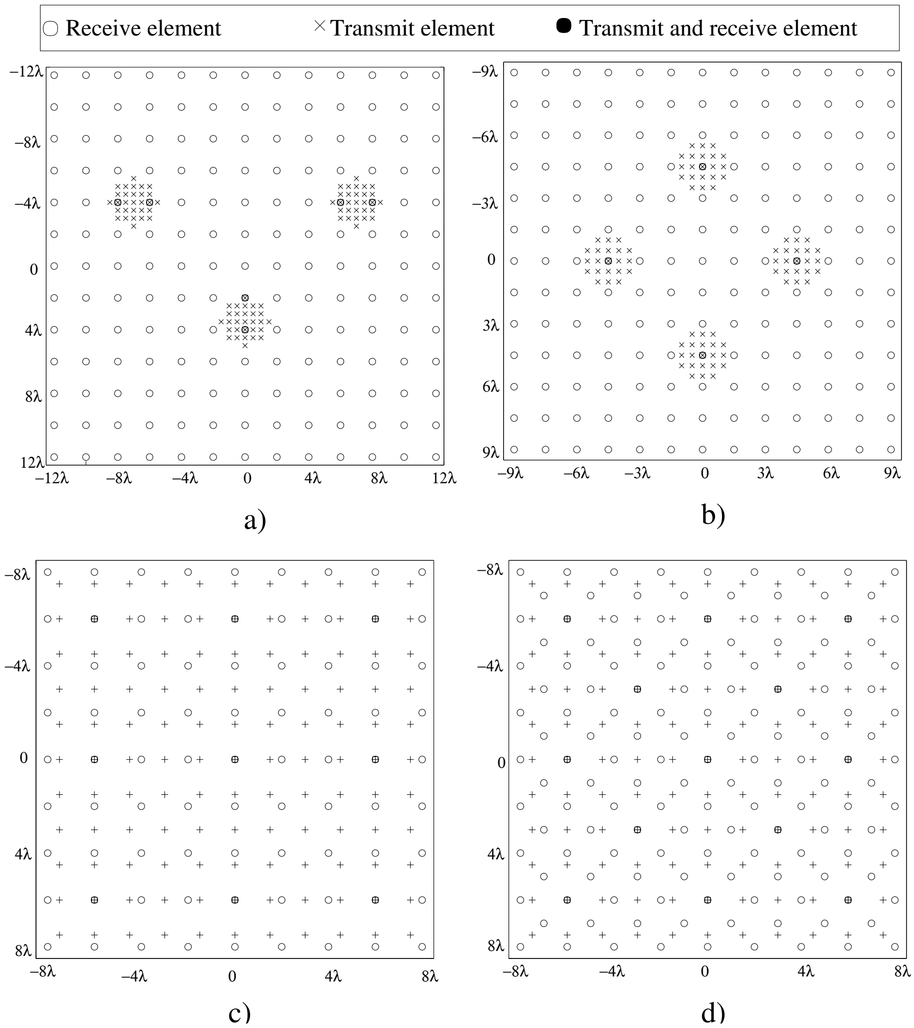


Figure B.3: The transmit and receive aperture geometries. From bottom down: apertures for triangular approximation; vernier arrays with triangular sampling pattern.

The transmit aperture has two active sub-apertures. The spacing between two active elements is $\lambda/2$. The number of active elements is $2N_{act}$. The width of the active sub-aperture is $D_{act} = N_{act}\lambda/2$. The spacing between the elements in the receive aperture is $d_{rcv} = D_{act}/2$. If it has N_{rcv} active elements, its size is $D_{rcv} = (N_{rcv}-1)D_{act}/2$. The spacing between the two active sub-apertures is $d_{sub} = D_{rcv}/2$.

Fig. B.2 c) shows how to obtain effective aperture, apodized with the coefficients of a Hamming window. In [50] these arrays are called "vernier arrays" and therefore this term will be used further in the article. The spacing of the arrays is chosen to be $n\lambda/2$ and $(n-1)\lambda/2$, where n is an integer number. This guarantees that the spacing between the samples of the effective aperture is $\lambda/2$. Additional apodization gives control over the radiation pattern. For the apertures in Fig. B.2 a value of $n = 3$ was used. From the properties of the hamming window it can be expected that this configuration would yield the lowest side and grating lobes, at the expense of decreased resolution.

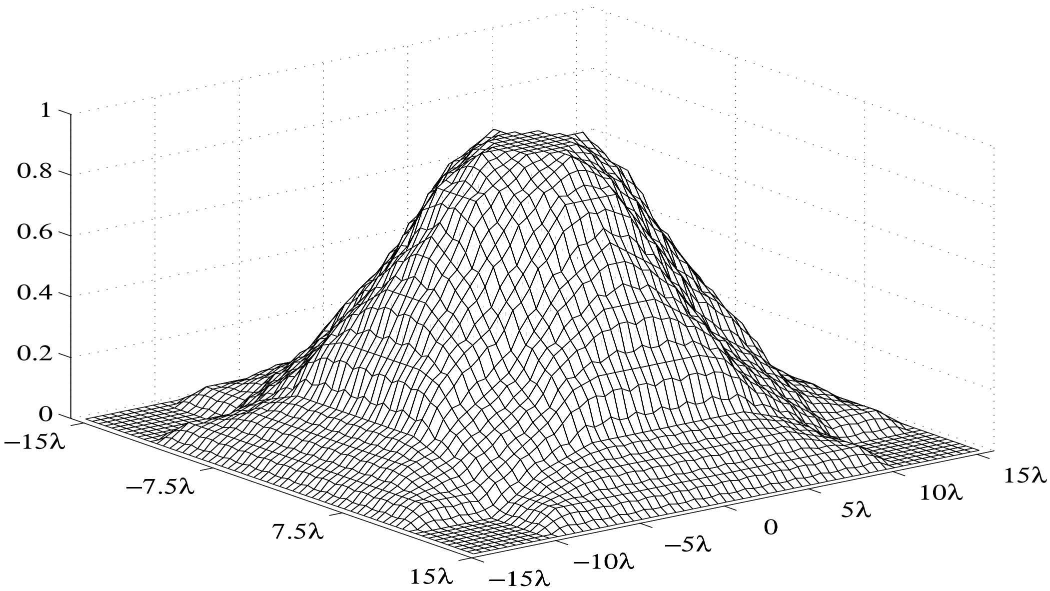


Figure B.4: Effective aperture obtained by a triangular approximation.

4 Transition to 2D

After the apertures are created in the one-dimensional case, their design must be extended to the two-dimensional space. Usually a rectangular grid is used for sampling of the transducer surface. The distance between two elements along the same row or column is $\lambda/2$.

The extension of the rectangular approximation to 2D is straight forward. Let the transmit aperture be a rectangular grid of size 11x11 elements, spaced at $\lambda/2$ distance. The horizontal and vertical spacing between the elements of the receive grid, in this case is 5λ.

Fig. B.3 a) and b) show two examples of triangular approximations. The resulting effective aperture is shown in Fig. B.4. Configuration a) has more transmit elements in a single active sub-aperture than configuration b). To maintain the same number of elements aperture a) has less active sub-apertures.

The vernier approximation can also be extended to the two-dimensional case by selecting the element spacing independently for the x and y axes, as it was previously done in [50]. Such configuration is shown in Fig.B.3 c) and will be further referred to as ”rectangular vernier approximation”. The spacing between the elements in the receive aperture is $4\lambda/2$ and in the transmit aperture it is $3\lambda/2$. The vernier nature of the sampling grid is preserved along the rows and the columns but is broken along the diagonals. This results in high grating lobes in the $(x-z)$ and $(y-z)$ planes. In Fig. B.4 d) a new, diagonally optimized pattern is suggested. A new element is inserted in the diagonal direction between every two receive elements. In this way the element spacing along the diagonals in the receive aperture becomes $2\lambda/\sqrt{2}$. The diagonal spacing in the transmit aperture is $\lambda/2$, thus keeping the vernier nature of the sampling pattern along the diagonals. According to the Fourier transform, this design decreases the grating-lobe energy with more than 5 dB.

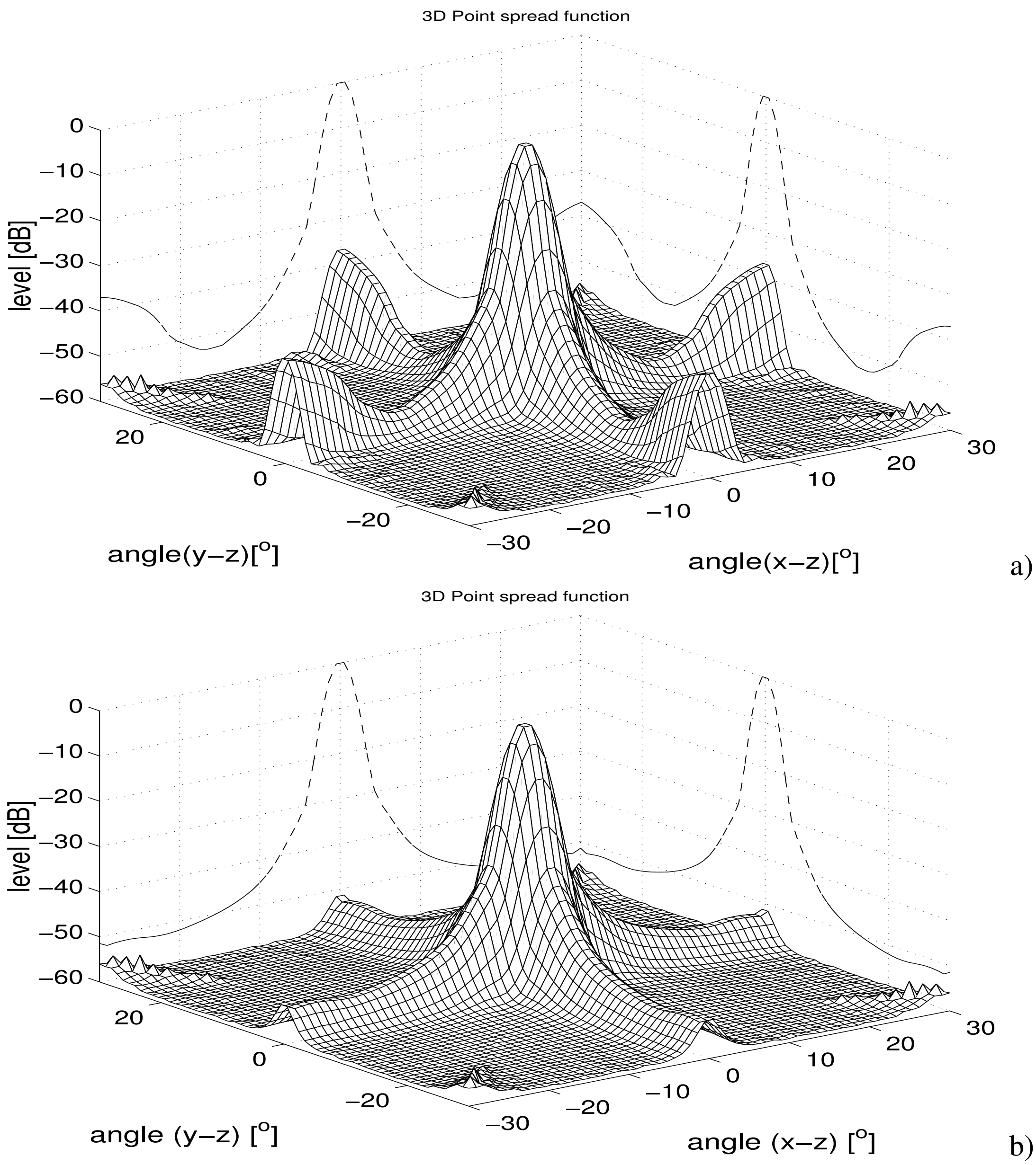


Figure B.5: Point spread functions. The arrays are designed using: a) Vernier interpolation b) Diagonally optimized Vernier approximation.

5 Simulation results

All simulations are made by the program Field II [19]. A 60°x60° volume containing a single point scatterer is scanned. The point-spread function is obtained by taking the maximum in a given direction. The simulation parameters are listed in Table B.1.

The size of the transmit and receive apertures are dependent on the design method. The apertures were designed to contain a total of 256 transmit and receive channels. Fig. B.5 shows two point-spread functions: a) a Vernier approximation, extended to 2D on a rectangular grid [50], and b) a Vernier approximation optimized along the diagonals. The diagonal optimization results in almost 10 dB decrease of the highest grating lobe level. From Table B.2 it can be seen that the loss of resolution at -3 dB is 9%.

Parameter	Notation	Value	Unit
Speed of sound	c	1540	m/s
Central Frequency	f_0	3	MHz
Sampling Frequency	f_s	105	MHz
Pitch	p	0.5	mm
Bandwidth	BW	70	%

Table B.1: Simulation parameters.

Design Method	-3dB deg.	-6dB deg.	MSR dB	Peak dB
Rectangular	1.21	1.67	30.6	-32.1
3 Clusters	2.15	2.81	47.7	-41.3
4 Clusters	3.09	4.04	49.8	-42.2
Rect Vernier	2.37	3.09	51.3	-38.0
Trian Vernier	2.61	3.57	51.8	-49.1

Table B.2: Simulation results. The table shows the -3dB and -6dB beam-widths, the mainlobe-to-sidelobe energy ratio (MSR), and the highest level of the grating lobes.

The point spread function $PSF(\theta,\phi)$ is a function of the azimuth and elevation angles θ and ϕ. The main lobe can be defined as:

$$ML(\theta,\phi) = \begin{cases} 1, & 20\lg\left(\frac{PSF(\theta,\phi)}{PSF(0,0)}\right) > -30 \\ 0, & otherwise \end{cases}$$

The mainlobe-to-sidelobe energy ratio is calculated by :

$$MSR = \frac{\sum\sum |PSF(\theta,\phi)|^2 ML(\theta,\phi)}{\sum\sum |PSF(\theta,\phi)|^2 (1 - ML(\theta,\phi))}$$

The contrast of the image is dependent on this ratio. Table. B.1 shows that the aperture obtained by the rectangular approximation has the highest resolution. Unfortunately it forms a side-lobe pedestal at -33 dB and performs poorly in viewing low-contrast objects like cysts. On the other hand, the apertures formed using triangular approximation have lower resolution, but the level of side-lobe and grating-lobe energy decreases with the angle. Thus distant high-radiation regions do not influence the dark regions in the image and images with higher contrast can be obtained. The diagonally optimized vernier array has the highest MSR and has the potential of yielding images with the highest contrast at the expense of a 9% reduced resolution.

6 Conclusion

The periodic arrays have lower side-lobe energy than the sparse arrays and they have the potential of giving images with higher contrast. The *effective aperture concept* is a fast and easy way to design sparse arrays with desired properties. The design can be first implemented in 1D and then extended to 2D. The results from simulations on Vernier arrays show that radiation patterns with higher MSR are obtained when the design includes the diagonals of the rectangular grid.

7 Acknowledgements

This work was supported by grant 9700883 and 9700563 from the Danish Science Foundation, and by a grant from B-K Medical A/S.

APPENDIX

C

Recursive ultrasound imaging

Authors : Svetoslav Ivanov Nikolov, Kim Gammelmark and Jørgen Arendt Jensen

Published : Proceedings of the IEEE Ultrasonics Symposium, Lake Tahoe, 1999.

Abstract

This paper presents a new imaging method, applicable for both 2D and 3D imaging. It is based on Synthetic Transmit Aperture Focusing, but unlike previous approaches a new frame is created after every pulse emission. The elements from a linear transducer array emit pulses one after another. The same transducer element is used after N_{xmt} emissions. For each emission the signals from the individual elements are beam-formed in parallel for all directions in the image. A new frame is created by adding the new RF lines to the RF lines from the previous frame. The RF data recorded at the previous emission with the same element are subtracted. This yields a new image after each pulse emission and can give a frame rate of e.g. 5000 images/sec.

The paper gives a derivation of the recursive imaging technique and compares simulations for fast B-mode imaging with measurements .

A low value of N_{xmt} is necessary to decrease the motion artifacts and to make flow estimation possible. The simulations show that for $N_{xmt} = 13$ the level of grating lobes is less than -50 dB from the peak, which is sufficient for B-mode imaging and flow estimation.

The measurements made with an off-line experimental system having 64 transmitting channels and 1 receiving channel, confirmed the simulation results. A linear array with a pitch of 208.5 μm, central frequency $f_{0tr} = 7.5$ MHz and bandwidth $BW = 70\%$ was used. The signals from 64 elements were recorded, beam-formed and displayed as a sequence of B-mode frames, using the recursive algorithm. An excitation with a central frequency $f_0 = 5$ MHz ($\lambda = 297$ μm in water) was used to obtain the point spread function of the system. The -6 dB width of the PSF is 1.056 mm at axial distance of 39 mm. For a sparse synthetic transmit array with $N_{xmt} = 22$ the expected grating lobes from the simulations are -53 dB down from the peak value at, positioned at $\pm 28^\circ$. The measured level was -51 dB at $\pm 27^\circ$ from the peak.

Images obtained with the experimental system are compared to the simulation results for different sparse arrays. The application of the method for 3D real-time imaging and blood-velocity estimations is discussed.

1 Introduction

Advances in DSP technology [36] make the real-time 3-D volumetric scanning a feasible imaging modality in medical ultrasound. Extending the traditional 2-D cross-sectional scanning to 3-D does not yield real-time imaging and new imaging methods have to be developed. The synthetic aperture techniques are attractive alternatives. They make it possible to increase the frame rate of B-mode ultrasound imaging, and obtain a dynamically focused image in both transmit and receive. The synthetic beam-formation approaches can be divided into three classes [36]:

- synthetic receive aperture [70]
- synthetic receive and transmit apertures [65], [61]
- synthetic transmit aperture [1],[94]

In synthetic aperture imaging the time needed to acquire a single high-resolution image (HRI) T_{HRI} is proportional to the number of emissions N_{xmt}, the time necessary to record the reflected ultrasound wave from a single emission T_{rec}, and the number of scan-lines N_l. It is inversely proportional to the number of the parallel receive beam-formers N_{prb}:

$$T_{HRI} = T_{rec} \cdot N_{xmt} \cdot N_l / N_{prb} \tag{C.1}$$

Synthetic receive processing involves transmitting with a full aperture and receiving with two or more sub-arrays. Several emissions are needed for every scan-line, thus, increasing T_{HRI}[36].

In contrast to synthetic receive aperture processing, synthetic transmit aperture imaging involves transmission from two or more sub-apertures and receiving with a full array. In receive the RF signals from the transducer elements can be beam-formed simultaneously for all directions in the image, *i.e.* $N_{prb} = N_l$ [1], [94]. For every emission a low-resolution image (LRI) is created. After acquiring N_{xmt} low-resolution images, the RF lines of these images are added together to form a single high-resolution image. The acquisition time for a LRI with a typical depth of 15 cm, assuming a speed of sound $c = 1540$ m/s, is $T_{rec} = 200$ μs. If $N_{xmt} = 64$, then $T_{HRI} = 12.8$ ms, and the frame rate is 78 frames/sec. This paper suggests further development of the transmit synthetic processing. Instead of discarding the already created high-resolution image and starting the imaging procedure all over again, the next high-resolution image is recursively built from the previous one by adding the beam-formed RF-lines from the next low-resolution image to the existing high-resolution RF-lines. The RF-lines from the low-resolution image obtained at the previous emission of the current transmit sub-aperture are subtracted from the result.

The suggested recursive calculation procedure makes it possible to create a new high-resolution image at every pulse emission, i.e. $N_{xmt} = 1$ and, thus, increase the frame rate up to 5000 frames/sec.

2 Recursive ultrasound imaging

Phased linear arrays are used to create sector B-mode images. The image consists of N_l scan-lines with common origin. Each scan-line l is defined by the angle with the normal vector to the transducer surface θ_l.

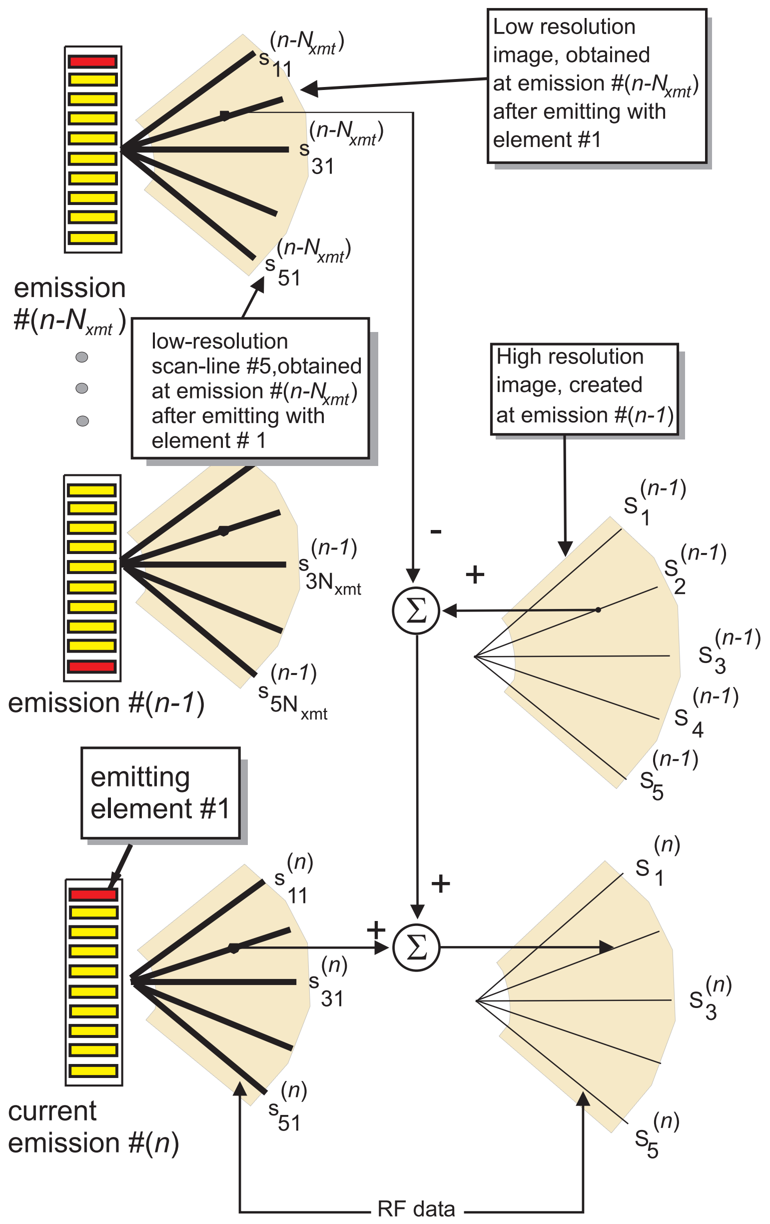


Figure C.1: Recursive ultrasound imaging. In transmit only one element is excited. Multiple receive beams are formed simultaneously for each transmit pulse. Each element is excited again after N_{xmt} emissions ($N_{xmt} = N_{xdc} = 10$ in this example).

A pulse emitted by only one transducer element propagates as a spherical wave, when the element is small, and the received echo signal carries information from the whole region of interest. By applying different delays in receive, *any* of the scan-lines $l \in [1 \dots N_l]$ can be formed. The data from one emission is used to beam-form *all* of the scan-lines creating one image as shown in Fig. C.1. The created image has a low resolution, since only one element is used for emission. A high-resolution image is created by summing the RF lines from N_{xmt} low resolution images, each of them created after emitting with a different transducer element. Let the number of the current emission be n, the number of the transducer elements be N_{xdc}, the recorded signal by the element j after emitting with element i be $r_{ij}^{(n)}$, and let the necessary delay and the weighting coefficient for beam-forming scan-line l be respectively d_{lij} and a_{lij}.

The beam-forming of a scan-line for a low-resolution image can then be expressed as (see Fig. C.1):

$$s_{li}^{(n)}(t) = \sum_{j=1}^{N_{xdc}} a_{lij} \cdot r_{ij}^{(n)}(t - d_{lij}), \tag{C.2}$$

where t is the time relative to the start of pulse emission. The number of skipped elements between two consecutive transmissions $n-1$ and n is:

$$N_{skip} = \text{floor}[(N_{xdc} - N_{xmt})/(N_{xmt} - 1)] \tag{C.3}$$

If $N_{xdc} = 64$ and $N_{xmt} = 4$ then $N_{skip} = 20$. The values for N_{xdc} should be a multiple of N_{xmt}, so that N_{skip} is an integer number.

The relation between the index i of the emitting element and the number n of the emission is given by:

$$i = [((n-1) \cdot (N_{skip} + 1)) \bmod N_{xdc}] + 1 \tag{C.4}$$

If $N_{skip} = 20$ then $i = 1, 22, 43, 64, 1, \cdots$. It can be seen that emissions n and $n \pm N_{xmt}$, are done by the same transducer element.

The forming of the final scan-lines for the high-resolution image can be expressed as:

$$S_l^{(n)}(t) = \sum_{k=n-N_{xmt}+1}^{n} s_{li}^{(k)}(t) \tag{C.5}$$

Equation (C.5) implies that a high-resolution image can be formed at any emission n, provided that N_{xmt} low-resolution images already exist. The images that will be formed at emissions n and $n-1$ can be expressed as:

$$S_l^{(n)}(t) = \sum_{k=n-N_{xmt}+1}^{n} s_{li}^{(k)}(t) \tag{C.6}$$

$$S_l^{(n-1)}(t) = \sum_{k=n-N_{xmt}}^{n-1} s_{li}^{(k)}(t) \tag{C.7}$$

Subtracting $S_l^{(n-1)}(t)$ from $S_l^{(n)}(t)$ gives:

$$S_l^{(n)}(t) = S_l^{(n-1)}(t) + s_{li}^{(n)}(t) - s_{li}^{(n-N_{xmt})}(t) \tag{C.8}$$

In (E.2) the new high-resolution scan-line depends on the low-resolution scan-lines at emissions n and $n - N_{xmt}$, and on the high-resolution scan-line at emission $n-1$. This dependence can be extended over a number of low- and high- resolution scan-lines obtained at previous emissions, and Equation (E.2) can be generalized as:

$$S_l^{(n)}(t) = \sum_{k=1}^{B} c_k \cdot S_l^{(n-k)}(t) + \sum_{q=0}^{Q} b_q \cdot s_{li}^{(n-q)}(t), \tag{C.9}$$

where B and Q are respectively the number of high- and low-resolution scan-lines on which $S_l^{(n)}$ depends, and c_k and b_q are weighting coefficients. The recursive imaging procedure uses information from previous emissions and therefore suffers from motion artifacts. They can be reduced by decreasing Q or/and B and in this way trading resolution for motion artifacts. If $B = 1$ and $Q = 0$, the imaging procedure becomes *add-only recursive imaging.*

3 Add-only recursive imaging

Let B and Q in (C.9) be respectively 1 and 0. The calculation procedure becomes:

$$S_l^{(n)}(t) = c_1 \cdot S_l^{(n-1)}(t) + b_0 \cdot s_{li}^{(n)}(t) \tag{C.10}$$

The difference between equations (E.2) and (C.10) is that instead of being subtracted, the information obtained after the emission with element i decays exponentially with time. In this way the information from the past is less prone to introduce motion artifacts in the image. The other benefit is that less memory is needed, since only two frames are stored. The high-resolution image is created by only adding weighted low-resolution images. This process starts at emission $n = 1$. Let all of the transducer elements participate in creating the synthetic transmit aperture ($N_{xmt} = N_{xdc}$). At emission n the high-resolution image is a weighted sum of all the low-resolution images obtained at the emissions with the single elements. Consider only the low-resolution images obtained after emissions with element i. The first emission with element i is $n = i$. The second emission with the same element is $n = i + N_{xmt}$. Element i is used after every N_{xmt} emissions. The sum C_{li} of the low-resolution scan-lines $s_{li}^{(n)}$ obtained at these emissions will be called *partially beam-formed scan-line* C_{li}. The high-resolution scan-lines are a sum of the partially beam-formed scan lines :

$$S_l(t) = \sum_{i=1}^{N_{xmt}} C_{li}(t) \tag{C.11}$$

If $b_0 = 1$, then the partially beam-formed scan-line for element i, $C_{li}^{(n)}$ at emission n is:

$$C_{li}^{(n)}(t) = s_{li}^{(n)}(t) + c_1^{N_{xmt}} \cdot s_{li}^{(n-N_{xmt})}(t) + c_1^{2N_{xmt}} \cdot s_{li}^{(n-2N_{xmt})}(t) + \cdots \tag{C.12}$$

This is a geometric series. If the tissue is motionless then:

$$s_{li}^{(n)}(t) = s_{li}^{(n-N_{xmt})}(t) = \cdots = s_{li}(t) \tag{C.13}$$

$$C_{li}^{(n)}(t) = [1 + c_1^{N_{xmt}} + c_1^{2N_{xmt}} + \cdots] s_{li}(t) \tag{C.14}$$

$$C_{li}^{(n)}(t) = s_{li}(t) \cdot \frac{1}{1 - c_1^{N_{xmt}}} \tag{C.15}$$

If $c_1 = 0.9$ and $N_{xmt} = 64$ then $1/(1 - c^{N_{xmt}}) \approx 1$ and $C_{li}^{(n)}(t) \approx s_{li}^{(n)}(t)$. Substituting the result in (C.11) gives the following result for the high-resolution scan-line:

$$S_l^{(n)}(t) \approx \sum_{i=0}^{N_{xmt}-1} c_1^i \cdot s_{li}^{(n-i)}(t) \tag{C.16}$$

Using (C.16) for imaging, instead of (E.2) gives images with lower resolution due to the weighting in the sum. In this case the resolution is traded for motion artifacts and less memory storage requirements, which is beneficial for flow estimation.

System parameter	Notation	Value	Unit
Speed of sound	c	1540	m/s
Central frequency	f_0	3	MHz
Sampling frequency	f_s	105	MHz
Oscillation periods	N_{osc}	3	
Pitch	$pitch$	0.257	mm
Number of elements	N_{xdc}	64	
Relative two-sided -6 dB bandwidth	B	70	%

Table C.1: Simulation parameters for a 3 MHz phased array system.

	$N_{act} = 1$		$N_{act} = 11$	
N_{xmt}	Position	Level	Position	Level
64	NA	NA	NA	NA
22	$\pm 40°$	-58 dB	$\pm 40°$	-58 dB
13	$\pm 21°$	-54 dB	$\pm 21°$	-53 dB
8	$\pm 13°$	-48 dB	$\pm 13°$	-47 dB

Table C.2: The position and level of the first grating lobe as a function of the number of emissions N_{xmt}.

4 Simulation results

Simulations were done to evaluate the performance of the imaging system as a function of the number of emissions N_{xmt}. Equation (E.2) was used to create high-resolution images of the point spread function.

All the simulations were done with the program Field II [19]. The parameters are listed in Table C.1.

The beam-formed signal was decimated 10 times and then envelope detected by a Hilbert trans-

System parameter	Notation	Value	Unit
Speed of sound	c	1485	m/s
Central frequency	f_0	5	MHz
Sampling frequency	f_s	40	MHz
Oscillation periods	N_{osc}	3	
Pitch	$pitch$	0.2085	mm
Number of elements	N_{xdc}	64	
Relative two-sided -6 dB bandwidth	BW	70	%

Table C.3: Parameters of the EXTRA measurement system [161]. The same parameters were used in simulations to obtain the expected levels and positions of the grating lobes

	Expected		Measured	
N_{xmt}	Position	Level	Position	Level
64	NA	NA	NA	NA
13	$\pm 15^\circ$	-53 dB	± 17	-51 dB
8	$\pm 8^\circ$	-47 dB	± 10	-44.5 dB

Table C.4: Measured versus expected grating lobe position and level for $N_{act} = 11$.

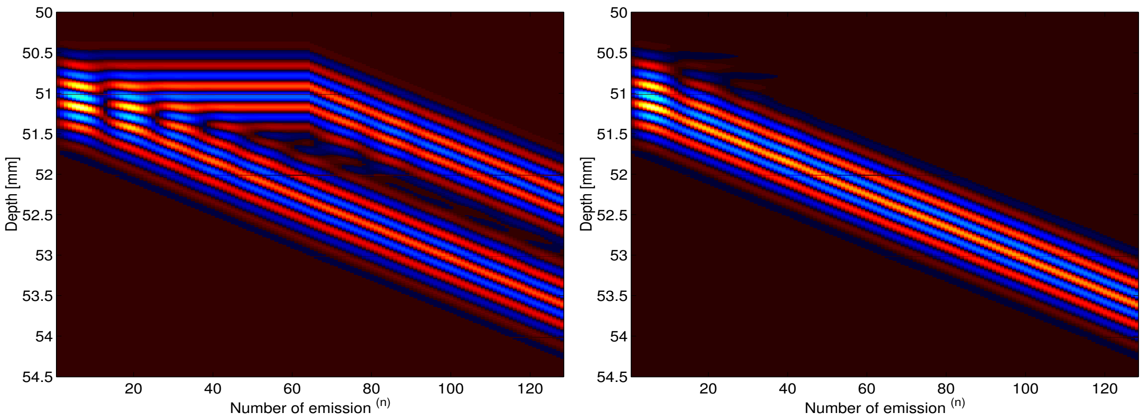


Figure C.2: The development of a single high-resolution scan-line as a function of the number of emissions n for normal recursive imaging (top), and for add-only recursive imaging (bottom).

formation. The envelope of the signal was logarithmically compressed with a dynamic range of 60 dB. Since the B-mode image is a sector image, the point spread function was obtained by taking the maximum of the reflected signal from every direction. In the first simulation only one element ($N_{act} = 1$) was used for a single emission. The -6 dB width of the acquired point-spread-function was 1.01° and the -40 dB width was 5.03°. The levels and positions of the grating lobes as a function of the number of emissions N_{xmt} are shown in Table C.2. The simulations, however, did not account for the attenuation of the signal and the presence of noise. For real applications the energy sent into the body at one emission must be increased. One way is to use multiple elements whose delays are set to create a spherical wave [1]. To verify the method, simulations with 11 active elements forming a spherical wave at every transmission, were done. The width of the point-spread-function was identical to the one obtained with $N_{act} = 1$. The levels and positions of the grating lobes are given in Table C.2. These results show that the radiation pattern of a single element can be successfully approximated by using several elements to increase the signal-to-noise ratio.

One of the problems, accompanying all synthetic aperture techniques are motion artifacts, and simulations with moving scatterers were therefore done. The signal from a single point scatterer moving at a constant speed $v = 0.1$ m/s away from the transducer was simulated. The simulation parameters were the same as those in Table C.1 except for the number of oscillation which in this case were $N_{osc} = 5$. The pulse repetition frequency was $f_{prf} = 5000$ Hz. Figure C.2 shows one RF-line of the high-resolution image as a function of the number of emissions n. From Fig. C.2, top it can be seen that the recursive imaging procedure suffers from motion artifacts as the other synthetic focusing algorithms. However, it can be seen from Fig. C.2, bottom that these artifacts are reduced for the add-only recursive imaging and it can be used for velocity

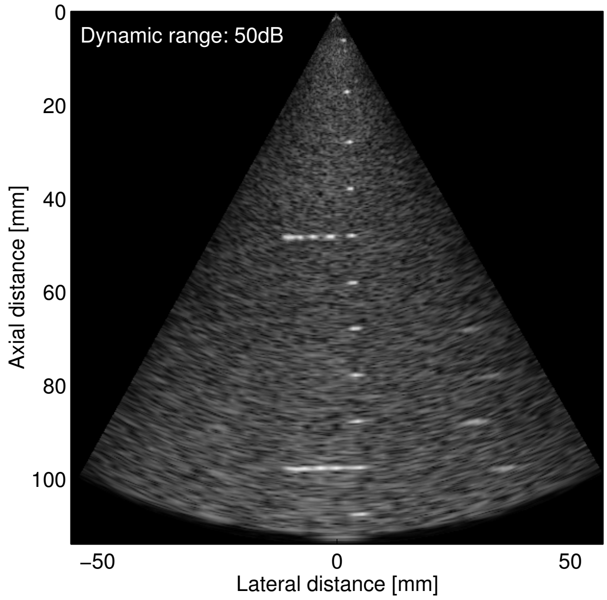


Figure C.3: Synthetic image of a wire phantom.

estimations.

5 Experimental results

The measurements were done with the off-line experimental system EXTRA [161] in a water tank and on a wire phantom with attenuation. The parameters of the system are listed in Table C.3. The transducer was a linear array with a pitch $p = \lambda$. These parameters differ from the desired ($p = \lambda/2$), and new simulations were made in order to determine the expected point-spread-function, and to compare it to the measured one. The expected -6 dB width point-spread-function was $1.38°$ and the -40 dB width was $-4.4°$. The expected positions and levels of the grating lobes are given in Table C.4. The results of the measurements are in good agreement with the simulation results. The result of scanning a wire phantom is given in Fig. C.3. This image was obtained using 11 elements at every emission. The phantom has a frequency dependent attenuation of $0.25\ \mathrm{dB} \cdot (\mathrm{cm} \cdot \mathrm{MHz})^{-1}$. The focusing delays are calculated for every scan-line sample, and the two-way propagation time of the acoustic wave is taken into consideration. Therefore the image has the quality of a dynamically focused in transmit and receive image.

6 Conclusions

A new fast imaging method has been presented. The created images have the quality of dynamically focused image in transmit and receive. The time necessary to create one frame is equal to the time of acquisition of a single scan line in the conventional scanners. The signal-to-noise ratio can be increased by using multiple elements in transmit. The motion artifacts are decreased by add-only recursive imaging and the acquired information can be used for velocity estimations.

Some of the possible applications of the method are in the real-time three-dimensional imaging and blood velocity vector estimation.

The method can be further optimized by the use of coded excitations to increase the penetration depth and the signal-to-noise ratio.

7 Acknowledgement

This work was supported by grant 9700883 and 9700563 from the Danish Science Foundation and by B-K Medical A/S.

The XTRA system was developed by S. K. Jespersen and provided by CADUS, Center for Arteriosclerosis Detection with Ultrasound, Technical University of Denmark.

APPENDIX

D

3D synthetic aperture imaging using a virtual source element in the elevation plane

Authors : Svetoslav Ivanov Nikolov and Jørgen Arendt Jensen

Published : Proceedings of the IEEE International Ultrasonics Symposium, Puerto Rico, 2000.

Abstract

The conventional scanning techniques are not directly extendable for 3D real-time imaging because of the time necessary to acquire one volume. Using a linear array and synthetic transmit aperture, the volume can be scanned plane by plane. Up to 1000 planes per second can be scanned for a typical scan depth of 15 cm and speed of sound of 1540 m/s. Only 70 to 90 planes must be acquired per volume, making this method suitable for real-time 3D imaging without compromising the image quality. The resolution in the azimuthal plane has the quality of a dynamically focused image in transmit and receive. However, the resolution in the elevation plane is determined by the fixed mechanical elevation focus.

This paper suggests to post-focus the RF lines from several adjacent planes in the elevation direction using the elevation focal point of the transducer as a virtual source element, in order to obtain dynamic focusing in the elevation plane.

A 0.1 mm point scatterer was mounted in an agar block and scanned in a water bath. The transducer is a 64 elements linear array with a pitch of 209 μm. The transducer height is 4 mm in the elevation plane and it is focused at 20 mm giving a F-number of 5. The point scatterer was positioned 96 mm from the transducer surface. The transducer was translated in the elevation direction from -13 to +13 mm over the scatterer at steps of 0.375 mm. Each of the 70 planes is scanned using synthetic transmit aperture with 8 emissions. The beamformed RF lines from the planes are passed through a second beamformer, in which the fixed focal points in the elevation plane are treated as virtual sources of spherical waves. Synthetic aperture focusing is applied on them. The -6 dB resolution in the elevation plane is increased from 7 mm to 2 mm. This gives a uniform point spread function, since the resolution in the azimuthal plane is also 2 mm.

1 Introduction

In the last years the interest in 3-D ultrasound imaging has been constantly increasing. However, due to technological limitations, there is only one real-time 3-D scanner [6], which uses 2-D matrix transducer arrays. Most other scanners employ conventional linear arrays to scan the volume of interest plane-by-plane, and then the information is reconstructed in a workstation. For a typical scan-depth of 15 cm and speed of sound 1500 m/s, the time for scanning a single plane consisting of 100 scan lines is 20 ms. Because of the long acquisition time for a single plane, this method has a low frame rate. Another draw-back is the non-uniform resolution in the elevation and azimuth planes. The latter can be solved by using 1.5-D arrays, but the frame rate remains low.

The frame rate can be increased by employing a sparse transmit synthetic aperture as suggested in [1]. In this approach only a few emissions are used per plane. If only 5 emissions were used, the time for scanning the plane is reduced from 20 ms to 1 ms, increasing the frame rate 20 times.

Previously a method for increasing the resolution of ultrasound images obtained by a fixed-focus transducer was suggested in [84]. In this approach the fixed focal point is treated as a virtual source of ultrasound, and the recorded RF lines are post focused to increase the resolution.

This paper suggests the combination of the two methods to improve both the frame rate and the resolution, since the linear array transducers are usually focused in the elevation plane. The planes are scanned one-by-one using synthetic transmit aperture focusing, and then the beamformed scan lines from the planes are refocused in the elevation plane to increase the resolution.

The paper is organized as follows. Section 2 gives the theory behind the methods and how the two methods are combined. The results from simulations and measurements are given in Sections 3 and 4, respectively. Finally the conclusions are drawn in Section 4.

2 Theoretical background

The following sections give the theoretical background for obtaining images using a synthetic aperture imaging and for performing post focusing.

2.1 Synthetic transmit aperture

When a single element of a linear array ultrasound transducer is excited, a spherical acoustic wave is created, provided that the element is small enough. The back scattered signal carries information from the whole region of investigation. In receive the RF lines in all directions are beamformed in parallel. Then another transmit element is excited and the process is repeated. The beamformed RF lines are summed after N_{xmt} elements have been used in transmit. The beamforming process is shown in Fig. D.1 and can be described as follows:

$$\begin{aligned} s_l(t) &= \sum_{i=1}^{N_{xmt}} \sum_{j=1}^{N_{xdc}} a_{lkj}(t) r_{kj}(t-\tau_{lkj}(t)), \quad l \in [1 \ldots N_l], \\ k &= f(i) \end{aligned} \tag{D.1}$$

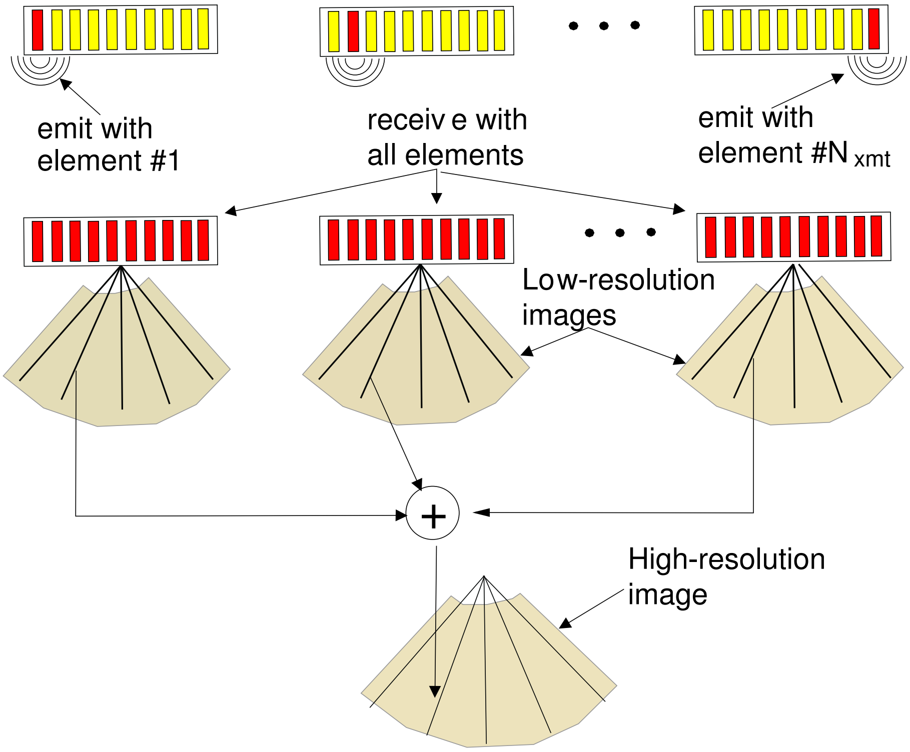


Figure D.1: Synthetic transmit aperture focusing

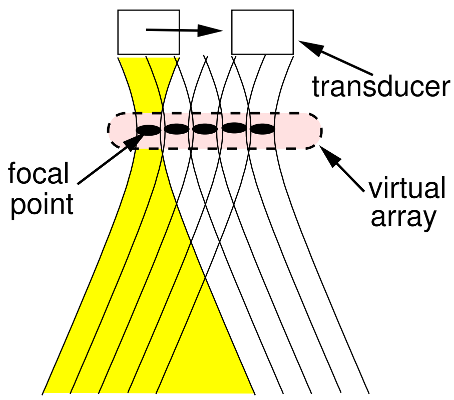


Figure D.2: Forming a virtual array from the focal points of the scanned lines

where l is the number of the scan line, t is time relative to the trigger of the current transmit, $r_{kj}(t)$ is the signal received by the jth element after transmitting with the kth element. $a_{lkj}(t)$ and $\tau_{lkj}(t)$ are the applied apodization factor and delay, respectively. N_{xdc} is the number of transducer elements, and $N_{xmt} \leq N_{xdc}$ is the number of emissions. The index of the transmitting element k is related to the number of the current emission i by a simple relation f. k is equal to i, when $N_{xmt} = N_{xdc}$. Only some of the transducer elements are used in transmit if $N_{xmt} < N_{xdc}$, and a sparse transmit aperture is formed [1]. Because the delay $\tau_{lkj}(t)$ is applied only in receive, and is a function of time, the image is dynamically focused in transmit and receive assuming a linear and stationary propagation medium.

2.2 Focusing using virtual source element

In the elevation plane the transducers are either unfocused or have a fixed focus, and hereby the scanned image has a poor resolution in this plane.

Figure D.2 shows a transducer in the elevation plane at several successive positions. The wavefront below the focal point can be considered as a spherical wave within a certain angle of divergence [84], and the focal point can be treated as a virtual source of ultrasound energy. The 3D volume is scanned by translating the transducer in the elevation direction in steps of Δy. The focal points lie on a line parallel to the transducer surface. The data can be considered as acquired by using one virtual element in transmit and receive.

Thus, synthetic aperture focusing can be applied on the beamformed RF lines from several emissions in order to increase the resolution in the elevation direction.

Let n, $1 \leq n \leq N_p$, denote the position and N_p be the number of several successive positions used for refocusing the data. The scan line $s_l(t)$ beamformed at position n will be denoted as $s_l^{(n)}(t)$. The final lines in the volume $S_l(t)$ are beamformed according to:

$$S_l(t) = \sum_{n=1}^{N_p} w_n(t) s_l^{(n)}(t - d_n(t)), \tag{D.2}$$

where $w_n(t)$ is a weighting coefficient and $d_n(t)$ is the applied delay. The delay necessary to focus at a given distance z ($z \geq f_{ez}$) is given by:

$$\begin{aligned} d_n(t) &= \frac{2}{c}\left(z - f_{ez} - \sqrt{(z - f_{ez})^2 + \left((n - 1 - \frac{N_p - 1}{2})\Delta y\right)^2}\right) \\ t &= \frac{2z}{c} \end{aligned} \tag{D.3}$$

where c is the speed of sound and f_{ez} is the distance to the elevation focus.

The best obtainable resolution in the elevation plane after post focusing is expected to be [83]:

$$\delta y_{6dB} \approx k \frac{0.41\lambda}{\tan\frac{\theta}{2}}, \tag{D.4}$$

where λ is the wavelength, and θ is the angle of divergence after the focal point. The variable k ($k \geq 1$), is a coefficient depending on the apodization. For a rectangular apodization k equals 1 and for Hanning apodization it equals 1.64.

The angle of divergence can be approximated by [84]:

$$\frac{\theta}{2} \approx \tan^{-1}\frac{h}{2f_{ez}}, \tag{D.5}$$

where h is the size of the transducer in the elevation plane, and f_{ez} is the distance to the fixed focus in the elevation plane. Substituting (D.5) in (D.4) gives:

$$\delta y_{6dB} \approx 0.82\lambda k \frac{f_{ez}}{h} \tag{D.6}$$

Equation (D.6) shows that the resolution is depth independent. However, this is true only if the number of the transducer positions is large enough to maintain the same F-number for the virtual array as a function of depth. For real-life applications the achievable resolution can be substantially smaller.

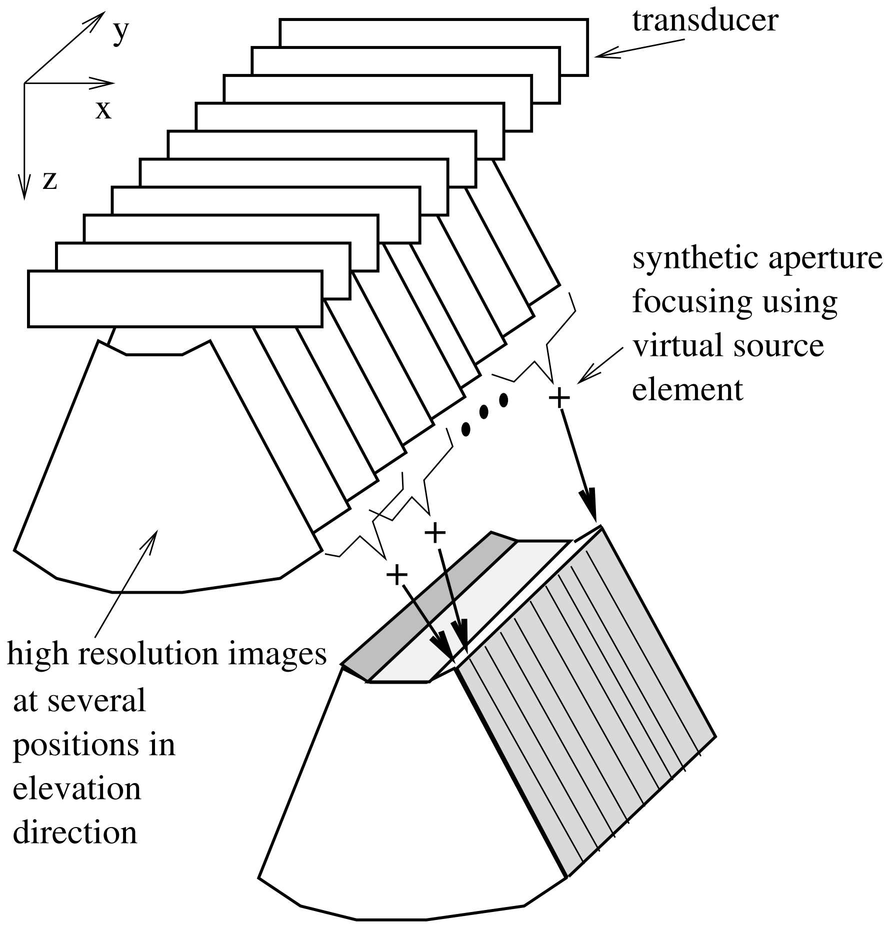


Figure D.3: The beamforming stages for 3D focusing.

2.3 Combining the two methods

The whole process can be divided into two stages and is summarized in Fig. D.3. In the first stage a high resolution image is created using only a few emissions, say $N_{xmt} = 5$ as given in Fig. D.1. This is repeated for several positions. Then the beamformed RF lines from these images are delayed, weighted, and summed a second time using (D.2) to form the final 3D volume. The considerations so far have been only for transducers translated in the y direction. The synthetic aperture focusing is applicable for any kind of transducer motion (translation, rotation, or a free-hand scan), as long as the exact positions of the focal points are known.

3 Simulations

The simulations were done using the program Field II [19]. The simulation parameters, given in Table D.1, were chosen to match the parameters of the system used for the measurements.

Seven point scatterers lying at depths from 70 to 100 mm were simulated at 70 positions. The distance between every two positions in the elevation direction was 0.7 mm. Figure D.4 on the left shows the -10 dB isosurfaces of the measured point-spread-functions (PSF). Then the beamformed scan lines were post-focused using $N_p = 30$ planes to create one new plane. If a dynamic apodization is to be used, then the number of usable positions N_p for depth z from the

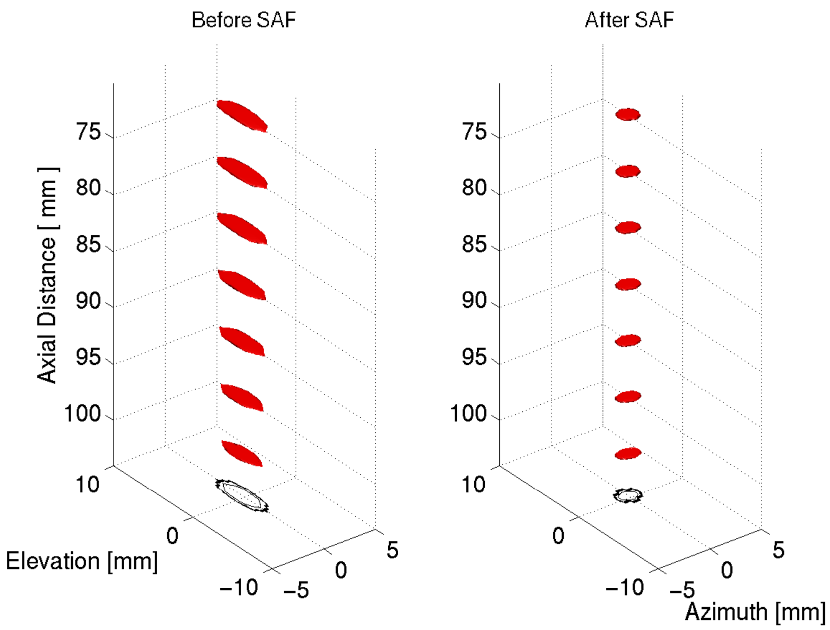


Figure D.4: The 3-D point-spread function outlined at -10 dB.

Parameter name	Notation	Value	Unit
Speed of sound	c	1480	m/s
Sampling freq.	f_s	40	MHz
Excitation freq.	f_0	5	MHz
Wavelength	λ	296	μm
-6 dB band-width	BW	4.875 - 10.125	MHz
Transducer pitch	p	209	μm
Transducer kerf	$kerf$	30	μm
Number of elements	N_{xdc}	64	-
Transducer height	h	4	mm
Elevation focus	f_{ez}	20	mm

Table D.1: Simulation parameters

real transducer can be determined by:

$$N_p = \left\lfloor 2h \frac{z - f_{ez}}{f_{ez}} \frac{1}{\Delta y} \right\rfloor , \tag{D.7}$$

Figure D.4 on the right shows the PSF after the post focusing was applied. Table D.2 shows the -6 dB resolution in the azimuth and the elevation planes. The lateral size of the PSF in the azimuth plane increases linearly with depth:

$$\delta x_{6db} = z \sin \phi_{6dB}, \tag{D.8}$$

where ϕ_{6dB} is the angular size of the PSF in polar coordinates.

The δy_{6dB} prior to the synthetic aperture focusing also increases almost linearly with depth, which shows that the beam is diverging with a certain angle as shown in Fig. D.2. After applying the synthetic aperture focusing δy_{6dB} becomes almost constant as predicted by (D.6).

Depth [mm]	δx_{6dB} [mm]	Before SAF δy_{6dB} [mm]	After SAF δy_{6dB} [mm]
70	1.44	4.78	1.72
75	1.54	5.16	1.72
80	1.65	5.48	1.72
85	1.75	5.80	1.85
90	1.85	6.18	1.85
95	1.96	6.56	1.85
100	2.06	6.75	1.97

Table D.2: The resolution at -6 dB as a function of depth.

A Hann window was used for w_n, and this gives $k \approx 1.6$. Substituting $h = 4$ mm, $f_{ez} = 20$ mm, and $\lambda = 0.296$ mm, gives $\delta y_{6dB} \approx 1.87$.

4 Measurements

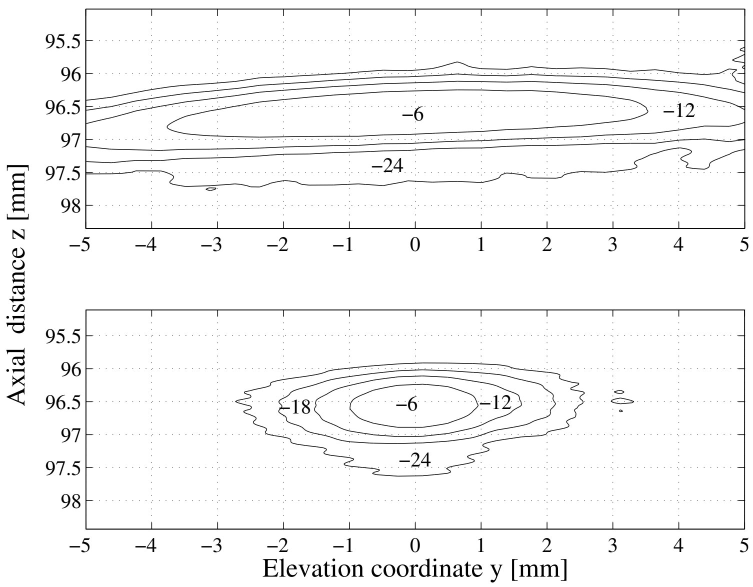


Figure D.5: PSF in the elevation plane: (top) before and (bottom) after synthetic aperture focusing. The innermost contour is at level of -6 dB, and the difference between the contours is also 6 dB.

The measurements were done using the department's off-line experimental system XTRA [161]. The parameters of the system are the same as the ones used in the simulations and are given in Table D.1.

In [84] it is argued that due to the narrow angle of divergence after the focal point, the grating lobes are greatly suppressed. Therefore it is possible to traverse the elevation direction at steps Δy bigger than one wavelength λ. Two experiments were conducted:

1. A point scatterer mounted in an agar block 96 mm away from the transducer was scanned, at step Δy = 375 μm. The diameter of the point scatterer was 100 μm.

2. A wire phantom was scanned at steps of Δy = 700 μm. The wires were positioned at depths from 45 to 105 mm, 20 mm apart. At every depth there were two wires, perpendicular to each other. The diameter of the wires was 0.5 mm.

The first experiment was conducted to verify the resolution achieved in the simulations. The goal of the second experiment was to verify that the resolutions in the elevation and azimuthal planes are comparable in size.

Using only a few emissions per plane corresponds to using a sparse transmit aperture. The use of wires as a phantom gives a good signal-to-noise ratio (compared to the 0.1 mm point scatterer) necessary to evaluate the level of the associated grating lobes. The SNR was further increased by using 11 elements in transmit to create a spherical wave instead of 1 as described in [1].

Figure D.5 shows the PSF of the point scatterer in the elevation plane. The contours are drawn at levels 6 dB apart. Sixty planes ($N_p = 60$) were used in the post-focusing. This maintains the same size of the virtual array as the one in the simulations. The achieved resolution at -6 dB is 2 mm and is comparable with the resolution obtained in the simulations.

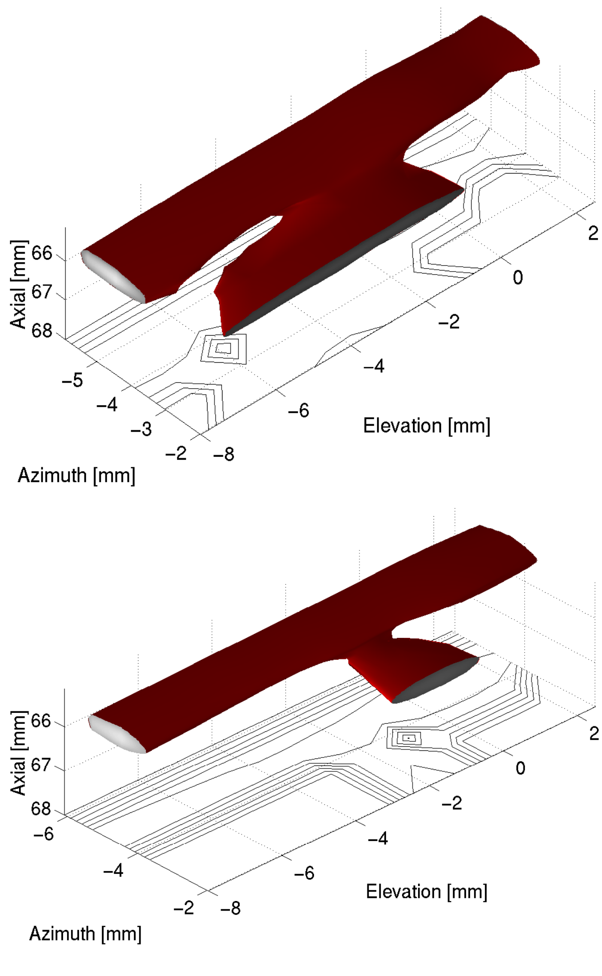


Figure D.6: Outline at -10 dB of the wire phantom: (top) before, and (bottom) after post focusing.

Figure D.6 shows a -10 dB outline of the wire phantom. The image was beamformed using $N_{xmt} = 8$ emissions per plane. The post focusing was performed using $N_p = 21$ planes for each new. The image shows that the resolution in the elevation and azimuth planes are of comparable size.

One of the problems of the approach are the grating lobes. They can be caused by two factors: (a) using only a few emissions per plane and (b) using a large step Δy between two planes.

The step Δy must not exceed the size of the PSF at the elevation focus, in order not to get image with discontinuities. For a larger step Δy, a transducer with a higher F-number must be used. Such transducers have a smaller angle of divergence θ and therefore the level of the grating lobes in the elevation direction is greatly suppressed. However, this is not the case in the azimuth plane. Table D.3 shows the compromise between the level of the grating lobes and

N_{xmt}	Position	Level	$\max(N_p)^1$
64	NA	NA	7
13	± 17	-51.0 dB	38
8	± 10	-44.5 dB	62
5	± 7	-41.3 dB	100

Table D.3: Grating lobes in the azimuth plane.

the maximum available number of planes for the post focusing. The figures are derived for speed of sound $c = 1540$ m/s and scan depth $\max(z) = 15$ cm. The table shows, that the larger the number of emissions per position N_{xmt} is, the larger the step δy must be, in order to maintain the number of volumes per second constant. This requires the use of a transducer which is not strongly focused in the elevation plane.

5 Conclusions

An approach for 3-D scanning using synthetic transmit aperture in the azimuth plane followed by synthetic aperture post focusing in the elevation plane was presented. The acquisition is done plane by plane using a conventional linear array transducer. The obtained resolution in the elevation plane is ≈ 2 mm, and is comparable to the one in the azimuth plane. Up to 10 volumes per second can be scanned if only 8 emissions per plane are used, and each volume contains 62 planes. This makes the approach a feasible alternative for real-time 3-D scanning.

6 Acknowledgements

This work was supported by grant 9700883 and 9700563 from the Danish Science Foundation and by B-K Medical A/S.

Some of the measured data were supplied by Søren Pihl Rybro, Peter Foged Christensen, and Madalina Breten.

The measurement system was built by Søren Kragh Jespersen as part of his Ph.D. study.

APPENDIX

E

Velocity estimation using recursive ultrasound imaging and spatially encoded signals

Authors : Svetoslav Nikolov, Kim Gammelmark and Jørgen Jensen

Published: Proceedings of the IEEE Ultrasonics Symposium, Puerto Rico, October 2000.

Abstract

Previously we have presented a recursive beamforming algorithm for synthetic transmit aperture focusing. At every emission a beamformed low-resolution image is added to an existing high-resolution one, and the low-resolution image from the previous emission with the current active element is subtracted yielding a new frame at every pulse emission.

In this paper the method is extended to blood velocity estimation, where a new Color Flow Mapping (CFM) image is created after every pulse emission. The underlying assumption is that the velocity is constant between two pulse emissions and the current estimates can therefore be used for compensation of the motion artifacts in the data acquired in the next emission.

Two different transmit strategies are investigated in this paper: (a) using a single defocused active aperture in transmit, and (b) emitting with all active transmit sub-apertures at the same time using orthogonal spatial encoding signals.

The method was applied on data recorded by an experimental system. The estimates of the blood velocity for both methods had a bias less than 3 % and a standard deviation around 2 % making them a feasible approach for blood velocity estimations.

1 Introduction

Modern scanners estimate the blood velocity by sending ultrasound pulses in the same direction and processing the signal returned from a given depth. To create a map of the velocity distribution in the area of investigation, the signal must be sent several consecutive times in each of several different directions. The precision of the estimates increases, if the estimates are based on a larger number of acquisitions in one direction. This, however, decreases the frame rate and the choice is based on a compromise between frame rate and precision.

This compromise can be avoided if a new frame is created after every emission and its data used for velocity estimation. The continuous flow of data allows the use of stationary echo canceling filters with longer impulse responses, and estimates based on a larger number of emissions, which both improve the estimates' precision.

One approach to create a new frame at every pulse emission is to use Recursive Ultrasound Imaging [80]. The beamformed data as proposed in [80] is suitable for B-mode imaging but not for blood velocity estimation, because of the present motion artifacts.

In this article the CFM is calculated after every emission, and the velocity estimates from the current frame are used for correcting the motion artifacts in the next one. Since the estimates are based on longer sample sequences, they have a high precision, and the motion artifacts can thereby be compensated fully.

Since each emission is performed only by one element, and the blood is moving, the performance of the above mentioned procedure depends on the shot sequence. This dependency can be avoided by using the same elements at every emission with a spatial encoding scheme as suggested in [3].

2 Theoretical background

The following sections give the theoretical background for recursive ultrasound imaging and the use of spatially encoded transmits to increase the signal-to-noise ratio.

2.1 Recursive imaging

A pulse emitted by only one transducer element propagates as a spherical wave, when the element is small, and the received echo signal carries information from the whole region of interest. By applying different delays in receive, *any* of the scan-lines $m \in [1 \dots N_m]$ can be formed. The data from one emission is used to beam-form *all* of the scan-lines creating one image as shown in Fig. E.1. The created image has a low resolution, since only one element is used for emission. A high-resolution image is created by summing the RF lines from N_{xmt} low resolution images, each of them created after emitting with a different transducer element. Let the number of the current emission be k, the number of the transducer elements be N_{xdc}, the recorded signal by the element j after emitting with element i be $r_{ij}^{(k)}$, and let the necessary delay and the weighting coefficient for beam-forming of scan-line m be d_{mij} and a_{mij}, respectively. The beam-forming of a scan-line for a low-resolution image can then be expressed as (see Fig. E.1):

$$s_{mi}^{(k)}(t) = \sum_{j=1}^{N_{xdc}} a_{mij} \cdot r_{ij}^{(k)}(t - d_{mij}), \tag{E.1}$$

where t is time relative to start of pulse emission. Provided that the tissue below the transducer is motionless, the forming of the final scan-lines for the high-resolution image can be expressed as [80]:

$$S_m^{(k)}(t) = S_m^{(k-1)}(t) + s_{mi}^{(k)}(t) - s_{mi}^{(m-N_{xmt})}(t) \tag{E.2}$$

This method, however, suffers from a low signal-to-noise (SNR) ratio and from motion artifacts. Using multiple elements in transmit to send "defocused" ultrasound wave improves [1] the

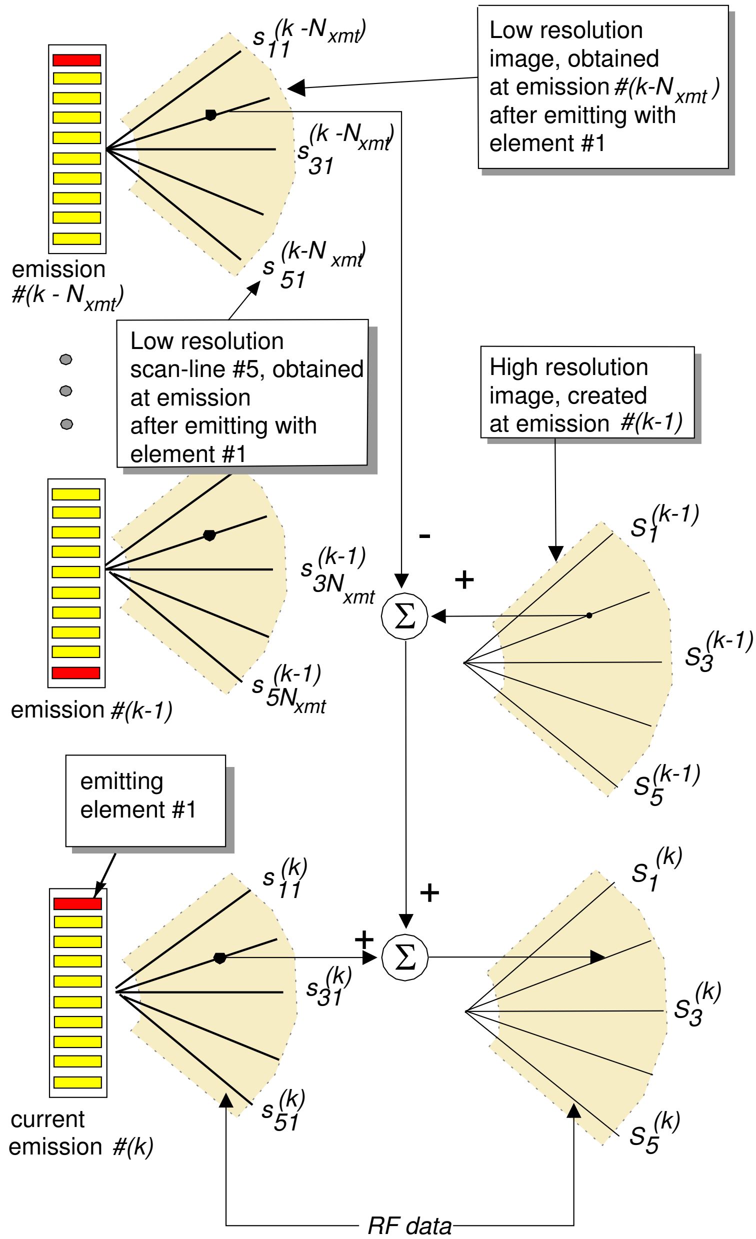


Figure E.1: Recursive ultrasound imaging. In transmit only one element is excited. Multiple receive beams are formed simultaneously for each transmit pulse. Each element is excited again after N_{xmt} emissions ($N_{xmt} = N_{xdc} = 10$ in this example).

situation. Further, the SNR can be increased by using encoded signals. The encoding can be temporal (for example using linear frequency modulated excitation) or spatial as described in the next section.

2.2 Spatial encoding

The idea behind the spatial encoding is to send with all of the N_{xmt} elements as shown in Fig. E.2, instead of sending with only one element i, $1 \le i \le N_{xmt}$ at a time [3]. The signal sent into

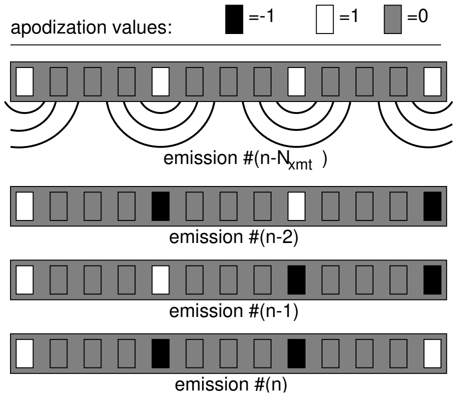


Figure E.2: Spatially encoded transmits using 4 transmit elements.

the tissue by each of the transmit elements i is

$$e_i(t) = q_i \cdot e(t), \quad 1 \le i \le N_{xmt}, \tag{E.3}$$

where $e(t)$ is a basic waveform and q_i is an encoding coefficient.

Assuming a linear propagation medium, the signal $r_j(t)$ received by the jth element can be expressed as:

$$r_j(t) = \sum_{i=1}^{N_{xmt}} q_i \cdot r_{ij}(t), \tag{E.4}$$

where $r_{ij}(t)$ would be the signal received by element j, if the emission was done only by element i.

From Eq.(E.1) it can be seen that the components $r_{ij}(t)$ must be found in order to beamform the signal. The received signals can be expressed in a matrix form:

$$\begin{bmatrix} r_j^{(1)} \\ r_j^{(2)} \\ \vdots \\ r_j^{(N_{xmt})} \end{bmatrix} = \begin{bmatrix} q_1^{(1)} & q_2^{(1)} & \cdots & q_{N_{xmt}}^{(1)} \\ q_1^{(2)} & q_2^{(2)} & \cdots & q_{N_{xmt}}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ q_1^{(N_{xmt})} & q_2^{(N_{xmt})} & \cdots & q_{N_{xmt}}^{(N_{xmt})} \end{bmatrix} \begin{bmatrix} r_{1j} \\ r_{2j} \\ \vdots \\ r_{N_{xmt}j} \end{bmatrix} \tag{E.5}$$

where the superscript $^{(k)}$, $1 \le k \le N_{xmt}$ is the number of the emission, $q_i^{(k)}$ is the encoding coefficient applied in transmit on the transmitting element i, and $r_j^{(k)}$ is the signal received by the jth element. In the above system of equations the time is skipped for notational simplicity. Also stationary tissue is assumed so that:

$$r_{ij}^{(1)} = r_{ij}^{(2)} = \cdots = r_{ij}^{(N_{xmt})} = r_{ij} \tag{E.6}$$

More compactly, the equation can be written as:

$$\vec{r}_j = \mathbf{Q}\vec{r}_{ij}, \tag{E.7}$$

where $\mathbf{Q}$ is the encoding matrix. Obviously the responses $r_{ij}(t)$ are:

$$\vec{r}_{ij} = \mathbf{Q}^{-1}\vec{r}_j \tag{E.8}$$

A suitable encoding matrix $\mathbf{Q}$ is the Hadamard matrix $\mathbf{H}$ [3]. The inverse Hadamard matrix is a scaled version of itself, i.e. for a matrix $\mathbf{H}_{N_{xmt}}$ with $N_{xmt} \times N_{xmt}$ elements, the inverse is $\mathbf{H}_{N_{xmt}}^{-1}$ = $1/N_{xmt}\mathbf{H}_{N_{xmt}}$.

The above derivation strongly relies on the assumption in (E.6). In the case of abdominal scanning, and for low values of N_{xmt}, this assumption is "almost fulfilled". However, in cardiac imaging and blood velocity estimation, this assumption is severely violated and the movement of the blood and heart must be compensated for.

2.3 Motion compensation

The motion compensation is considered for two cases: (a) recursive imaging without spatial encoding and (b) recursive imaging with spatial encoding.

Without spatial encoding

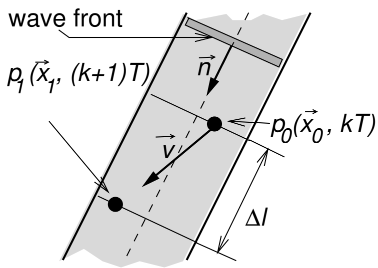


Figure E.3: Motion compensation for recursive imaging without spatial encoding.

During the first stage of the beamforming process, low-resolution images are created, using dynamic receive focusing. The assumption is that within one scan line $s_{mi}(t)$, the wavefront propagates as a plane wave, as shown in Fig. E.3. Figure E.3 shows the movement of one point scatterer within the limits of one scan line. The scatterer moves with velocity $\vec{v}$, from position $p_0(\vec{x}_0, kT)$ to a new position $p_1(\vec{x}_1, (k+1)T)$ for the time T between two pulse emissions. The movement across the beam (perpendicular to the the direction $\vec{n}$) determines the strength of the backscattered energy, while the movement along the beam determines the time instance when the backscattering occurs.

For the case depicted in Fig. E.3 the difference in time when the backscattering occurs for the positions p_0 and p_1 is:

$$\tau = \frac{2 \cdot \Delta l}{c}, \tag{E.9}$$

where Δl is the distance traveled from one pulse emission to the next:

$$\Delta l = \langle \vec{v}, \vec{n} \rangle T, \tag{E.10}$$

where $\langle \vec{v}, \vec{n} \rangle$ is the inner product between the velocity vector $\vec{v}$ and the directional vector $\vec{n}$.

The velocity at emission k as a function of time t from the emission of the pulse along the line m is $v_m^{(k)}(t)$. The delay τ is also a function of t, $\tau^{(k)} = \tau_m^{(k)}(t)$. The beamformation process with the velocity incorporated in it becomes:

$$\begin{aligned}
&\text{for } m = 1 \text{ to } N_m \\
&s_{mi}^{(k)}(t) = \sum_{j=1}^{N_{xdc}} a_{mij} \cdot r_{ij}^{(k)}(t - d_{mij}) \\
&M_m^{(k)}(t) = S_m^{(k-1)}(t) - s_{mi}^{(k-N_{xmt})}(t) \\
&\tau_m^{(k+1)}(t) = \frac{2|\vec{v}_m^{(k)}(t)|T}{c} \\
&\Delta_m^{(k)}(t) = \Delta_m^{(k-1)}(t) + \tau_m^{(k)} - \tau_m^{(k-N_{xmt}+1)}(t) \\
&S_m^{(k)}(t) = M_m^{(k)}[t + \tau_m^{(k-N_{xmt}+1)}(t)] + s_{mi}^{(k)}[t - \Delta_m^{(k)}(t)]
\end{aligned} \tag{E.11}$$

where Δ is the delay between the first and the last of the low-resolution images, currently comprised in the high-resolution one.

With spatial encoding

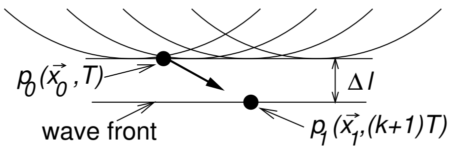


Figure E.4: Motion compensation for recursive imaging with spatial encoding.

Figure E.4 shows the model adopted for motion compensation in the presence of spatially encoded signals. Several transducer elements across the whole span of the transducer aperture are used in transmit. The sum of the emitted waves creates a planar wave propagating in a direction perpendicular to the transducer surface. A point scatterer moves for one pulse-repetition period from positions $p_0(\vec{x}_0, kT)$ to a new position $p_1(\vec{x}_1,(k+1)T)$. The difference between the time instances, when the scattering occurs, is :

$$\tau = \frac{2\Delta l}{c}, \tag{E.12}$$

where Δl is the distance traveled by the point scatterer:

$$\Delta l = v_z T \tag{E.13}$$

In the above equation v_z is the component of the velocity normal to the transducer surface. The delay τ is a function of the time t, and the emission number k, $\tau = \tau^{(k)}(t)$

Thus, the signals received by element j for emission number $k \in [1, N_{xmt}]$ are:

$$\begin{aligned}
r_j^{(2)}(t) &= r_j^{(1)}(t - \tau^{(1)}(t)) \\
r_j^{(3)}(t) &= r_j^{(2)}(t - \tau^{(2)}(t)) \\
&\vdots \\
r_j^{(N_{xmt})}(t) &= r_j^{(N_{xmt}-1)}(t - \tau^{(N_{xmt}-1)}(t))
\end{aligned} \tag{E.14}$$

The reconstruction must be performed prior to beamforming the signal at a given point. First the received signals $r^{(k)}(t)$ are appropriately delayed, and then the system of equations (E.5) is solved.

3 Experimental results

3.1 Measurement setup

The measurements were done, using the department's off-line experimental system XTRA [161]. The most important parameters are listed in Table E.1.

Parameter name	Notation	Value	Unit
Speed of sound	c	1540	m/s
Sampling freq.	f_s	40	MHz
Excitation freq.	f_0	5	MHz
Pulse duration	T_p	1.5	cycles
-6 dB band-width	BW	4.875 - 10.125	MHz
Transducer pitch	p	209	μm
Transducer kerf	$kerf$	30	μm
Number of elements	N_{xdc}	64	-
Transducer height	h	4	mm
Elevation focus	f_{ez}	20	mm

Table E.1: Measurement parameters

A tissue mimicking phantom with frequency dependent attenuation of 0.25 dB/[cm.MHz] and speed of sound $c = 1540$ m/s was scanned at 65 positions in a water bath. From position to position the phantom was moved 70 μm at an angle of 45° to the transducer surface. Assuming a pulse repetition frequency $f_{prf} = 1/T = 7000$, this movement corresponds to a plug-flow with velocity $|\vec{v}| = 49.5$ cm/s.

A precision translation system was used for the movement of the phantom. The precision of the motuon in the axial and lateral directions were: $\Delta z = 1/200$ mm, and $\Delta x = 1/80$ mm, respectively.

3.2 Velocity estimation

In the theoretical considerations, it was assumed that the blood velocity was estimated, without any considerations about the velocity estimator.

The cross-correlation estimator suggested in [133] is suitable for the broad band pulses used by this method. In the implementation it is assumed, that the two consecutive high-resolution lines $S^{(k)}(t_2)$ and $S^{(k-1)}(t_1)$ are related by:

$$S^{(k)}(t_2) = S^{(k-1)}(t_1 - t_s), \tag{E.15}$$

where t_s is a time lag due to the movement of the scatterers and is related to the axial component

	Reference	Spatially encoded	Non encoded
$\overline{\|\vec{v}\|}$ [m/s]	0.496	0.486	0.479
$\sigma/\overline{\|\vec{v}\|}$ %	2.3	2.2	1.8

Table E.2: Results from the velocity estimation at angle $(\vec{v}, \vec{n}) = 45°$

of the velocity v_z by:

$$t_s = \frac{2v_z}{c}T \tag{E.16}$$

The peak of the cross-correlation between segments of $S^{(k)}(t)$ and $S^{(k-1)}(t)$ would be found at time $\hat{t}_s$. Estimating $\hat{t}_s$ leads to the estimation of $\vec{v}$.

3.3 Reference velocity estimation

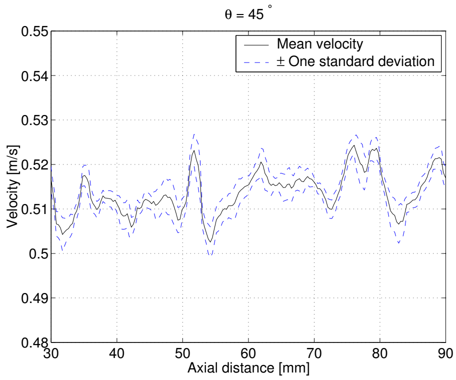


Figure E.5: Mean reference velocity.

In order to obtain a reference estimate of the velocity, at each position of the phantom a high resolution image was created using 13 emissions per image. The velocity was estimated using a cross-correlation estimator. The correlation length was equal to the length of the transmitted pulse. The number of lines, over which the calculated correlation function was averaged was 8. The search length was $\pm\lambda/4$ to avoid aliasing problems. Figure E.5 shows the mean velocity $\overline{|\vec{v}|}$ for the central line as a function of depth. The mean was calculated over 55 estimates.

In the axial direction the translation system has a precision of $\Delta z = 5$ μm, which is 10 % of the desired step. The results are, thus, within the precision of the system.

3.4 Recursive velocity estimation

The mean velocity $\overline{|\vec{v}|}$ and the normalized standard deviation $\sigma/\overline{|\vec{v}|}$ estimated using recursive ultrasound imaging are shown in Table E.2. The angle between the velocity vector $\vec{v}$ and the directional vector $\vec{n}$ of the scan line is $\angle(\vec{v}, \vec{n}) = 45°$. The number of frames is 36. In this table $\overline{|\vec{v}|}$ is the average of the mean velocity in the range from 30 to 80 mm. σ is also averaged in the same range. The angle dependence of the estimates is shown in Figure E.6. The dashed lines show the velocity at $\pm\sigma$.

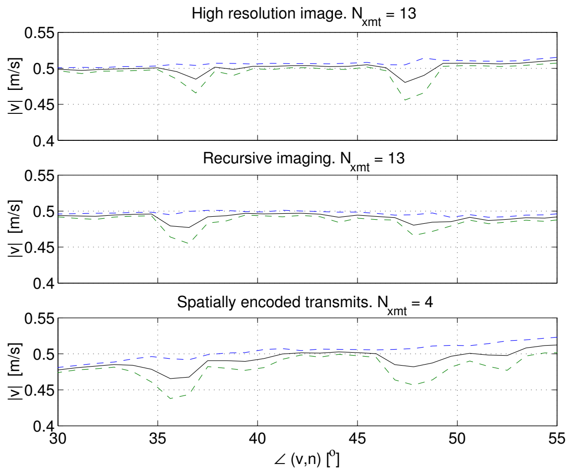


Figure E.6: The mean velocity and the velocity at $\pm\sigma$ as a function of angle.

It can be seen that the reference velocity estimation exhibits a smaller bias than the velocity estimations using recursive imaging.

The recursive imaging using spatially encoded transmits exhibits angular dependence. At angles of $42^\circ - 45^\circ$ it has a low bias and standard deviation, comparable to that of the reference velocity estimates. One of the possible reasons for the angular dependency is the low number of emissions ($N_{xmt} = 4$), resulting in higher side and grating lobes in the image.

4 Conclusions

In this paper a method for motion compensation and velocity estimation using recursive ultrasound imaging was presented. The method provides the blood velocity estimator with as much as several thousand measurements per second for every sample in the investigated region.

It has been experimentally verified that the method works for a speckle generating phantom with frequency dependent attenuation. One limitation is that no noise was present in the experiment and the velocity was constant.

Future work will include velocity profiles and mixture of moving and stationary tissue.

5 Acknowledgements

This work was supported by grant 9700883 and 9700563 from the Danish Science Foundation and by B-K Medical A/S.

The measurement system XTRA was built by Søren Kragh Jespersen, as part of his Ph.D. study.

APPENDIX

F

Fast simulation of ultrasound images

Authors : Jørgen Arendt Jensen and Svetoslav Ivanov Nikolov

Published: Proceedings of the IEEE Ultrasonics Symposium, Puerto Rico, 2000.

Abstract

Realistic B-mode and flow images can be simulated with scattering maps based on optical, CT, or MR images or parametric flow models. The image simulation often includes using 200,000 to 1 million point scatterers. One image line typically takes 1800 seconds to compute on a state-of-the-art PC, and a whole image can take a full day. Simulating 3D images and 3D flow takes even more time. A 3D image of 64 by 64 lines can take 21 days, which is not practical for iterative work. This paper presents a new fast simulation method based on the Field II program. In imaging the same spatial impulse response is calculated for each of the image lines, and making 100 lines, thus, gives 100 calculations of the same impulse response delayed differently for the different lines. Doing the focusing after this point in the simulation can make the calculation faster. This corresponds to full synthetic aperture imaging. The received response from each element is calculated, when emitting with each of the elements in the aperture, and then the responses are subsequently focused. This is the approach taken in this paper using a modified version of the Field II program. A 64 element array, thus, gives 4096 responses. For a 7 MHz 64 element linear array the simulation time for one image line is 471 seconds for 200,000 scatterers on a 800 MHz AMD Athlon PC, corresponding to 17 hours for one image with 128 lines. Using the new approach, the computation time is 10,963 seconds, and the beamforming time is 9 seconds, which makes the approach 5.5 times faster. For 3D images with 64 by 64 lines, the total conventional simulation time for one volume is 517 hours, whereas the new approach makes the simulation in 6,810 seconds. The time for beamforming is 288 seconds, and the new approach is, thus, 262 times faster. The simulation can also be split among a number of PCs for speeding up the simulation. A full 3D one second volume simulation then takes 7,500 seconds on a 32 CPU 600 MHz Pentium III PC cluster.

1 Introduction

The simulation of ultrasound imaging using linear acoustics has been extensively used for studying focusing, image formation, and flow estimation, and it has become a standard tool in ultrasound research. Simulation, however, still takes a considerable amount of time, when

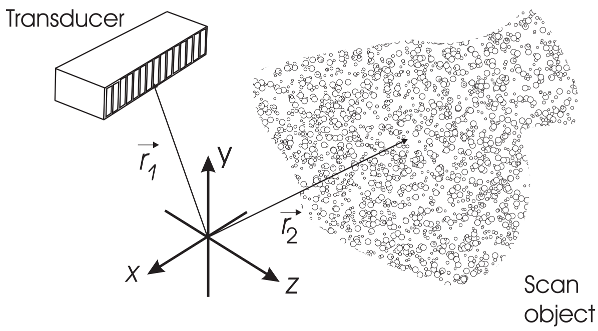


Figure F.1: Set-up for simulation of ultrasound imaging.

realistic imaging, flow, or 3D imaging are studied. New techniques for reducing the simulation time are, thus, desirable.

The main part of an ultrasound image consists of a speckle pattern, which emanates from the signal generated by tissue cells, connective tissue, and in general all small perturbations in speed of sound, density, and attenuation. The generation of this can be modeled as the signal from a large collection of randomly placed point scatterers with a Gaussian amplitude. Larger structures as vessel or organ boundaries can be modeled as a deterministicly placed set of point scatterers with a deterministic amplitude. The relative amplitude between the different scatterers is then determined by a scatterer map of the structures to be scanned. Such maps can be based on either optical, CT or MR images, or on parametric models of the organs. Currently the most realistic images are based on optical images of the anatomy [162]. Blood flow can also be modeled by this method. The red blood cells, mainly responsible for the scattering, can be modeled as point scatterers and the flow of the blood can be simulated using either a parametric flow model [163] or through finite element modeling [164]. The received signal is then calculated, and the scatterers are propagated between flow emissions. The simulation of all linear ultrasound systems can, thus, be done by finding the summed signal from a collection of point scatterers as shown in Fig. F.1. The random selection of point scatterers should consist of at least 10 scatterers per resolution cell to generate fully developed speckle, and for a normal ultrasound image this results in 200,000 to 1 million scatterers. The simulation of the responses from these scatterers must then be done for each line in the resulting image, and the simulation for the whole collection is typically done 100 times with different delay focusing and apodization. This makes the simulation take several days even on a fast workstation.

A second possibility is to do fully synthetic aperture imaging in which the received response by all elements are found, when transmitting with each of the elements in the array. The response of each element is then only calculated once, and the simulation time can be significantly reduced. This is the approach suggested in this paper. A second advantage of such an approach is that the image is focused after the field simulation. The same data can, thus, be used for testing a number of focusing strategies without redoing the simulation. This makes is easier to find optimized focusing strategies.

2 Theory

The field simulation must find the received signal from a collection of point scatterers. Using linear acoustics the received voltage signal is [30]:

$$v_r(t) = v_{pe}(t) \underset{t}{\star} f_m(\vec{r}_1) \underset{r}{\star} h_{pe}(\vec{r}_1, t), \tag{F.1}$$

where $\underset{r}{\star}$ denotes spatial convolution, $\underset{t}{\star}$ temporal convolution, and $\vec{r}_1$ the position of the point scatterer. $v_{pe}(t)$ is the pulse-echo wavelet, which includes both the transducer excitation and the electro-mechanical impulse response during emission and reception of the pulse. f_m accounts for the inhomogeneities in the tissue due to density and speed of sound perturbations that generates the scattering, and h_{pe} is the pulse-echo spatial impulse response that relates the transducer geometry to the spatial extent of the scattered field. Explicitly written out the latter term is:

$$h_{pe}(\vec{r}_1, t) = h_t(\vec{r}_1, t) \underset{t}{\star} h_r(\vec{r}_1, t) \tag{F.2}$$

where $h_t(\vec{r}_1, t)$ is the spatial impulse response for the transmitting aperture and $h_r(\vec{r}_1, t)$ is the spatial impulse response for the receiving aperture. Both impulse responses are a superposition of spatial impulse responses from the individual elements of a multi-element aperture properly delayed and apodized. Each impulse response is:

$$h(\vec{r}, t) = \sum_{i=1}^{N_e} a_i(t) h_i(\vec{r}_1, t - \Delta_i(t)), \tag{F.3}$$

where $a_i(t)$ denotes the apodization and $\Delta_i(t)$ focusing delay, which both are a function of position in tissue and thereby time. N_e is the number of transducer elements.

The received signal from each scatterer must be calculated for each new focusing scheme corresponding to the different lines in an image. The resulting *rf* signal is then found by summing the responses from the individual scatterers using (F.1). The number of evaluations of spatial impulse responses for individual transducer elements is:

$$N_h = 2 N_e N_s N_i, \tag{F.4}$$

where N_s is the number of point scatterers and N_i is the number of imaging directions. It is assumed that the number of elements in both transmitting and receiving aperture are the same, and that the apodization and focusing are included in the calculation. A convolution between $h_t(\vec{r}_1, t)$, $h_r(\vec{r}_1, t)$ and $v_{pe}(t)$ must be done for each scatterer and each imaging direction. This amounts to

$$N_c = 2 N_s N_i \tag{F.5}$$

convolutions for simulating one image.

The same spatial impulse response for the individual elements are, thus, being evaluated N_i times for making an image, and an obvious reduction in calculation time can be gained by just evaluating the response once. This can be done by making a synthetic aperture simulation approach. Here the response on each of the receiving elements from excitation of each of the transmitting elements are calculated. The received responses from the individual elements are beamformed afterwards. Hereby the number of evaluations of the spatial impulse responses is

$$N_{hs} = N_e N_s. \tag{F.6}$$

The number of convolutions is increased to

$$N_{cs} = N_s N_e^2 + N_s N_e, \tag{F.7}$$

since all emissions must be convolved with the response from all receiving elements and $v_{pe}(t)$ must be convolved with the responses. This can be reduced to

$$N_{cs} = N_s (N_e + \sum_{i=1}^{N_e} i) = 0.5 N_s (N_e^2 + 3N_e), \tag{F.8}$$

if the transmitting and receiving elements are the same, whereby the signal received is the same due to acoustic reciprocity [165], when the transmitting and receiving elements are interchanged. The beamforming is done after the calculation, but this can be done very efficiently as demonstrated in Section 3. The improvement in calculation of responses is given by

$$I_h = \frac{2N_e N_s N_i}{N_e N_s} = 2N_i \tag{F.9}$$

and for the convolutions

$$I_c = \frac{2N_s N_i}{0.5 N_s (N_e^2 + 3N_e)} = \frac{4N_i}{(N_e^2 + 3N_e)} \tag{F.10}$$

For a 64 element array and an image with 100 directions, the theoretical improvements are $I_h = 200$ and $I_c = 0.0933 = 1/10.7$. The efficiency of the approach is, thus, very dependent on the actual balance between evaluating the spatial impulse responses and performing convolutions. A significant speed-up is attained for few elements and many imaging directions, since few convolutions are performed. The balance is affected by the method for calculating the spatial impulse responses. The simulation program Field II [166, 19] offers three different possibilities, which are all based on dividing the transducer into smaller elements. The program uses a far-field rectangle solution [166], a full solution for triangles [26], or a bounding line solution [28]. The first solution is very fast, whereas the last two solutions are highly accurate but significantly slower. The choice of method will, thus, affect the balance.

A second aspect, in the implementation of the approach, is the use of memory. The number of bytes, when using double precision data, is

$$B_y = 8N_e^2 N_r \tag{F.11}$$

where N_r is the number of samples in the response. For at 64 elements array covering a depth of 15 cm, this gives 625 MBytes at a sampling frequency of 100 MHz. This is clearly too much for current standard PCs and even for some workstations. The simulation must be made at a high sampling frequency to yield precise results, but the data can, however, be reduced by decimating the signals after simulation of individual responses. A factor of 4 can *e.g.* be used for a 3 or 5 MHz transducer. The memory requirement is then 156 MBytes, which is more acceptable. It is, however, still large, and much larger than the cache in the computer. It is therefore necessary to reduce the number of cache misses. This is sought achieved in the program by sorting the scatterers according to the distance to the array, which gives results that are placed close in the memory. The memory interface of the computer is, however, very important in obtaining a fast simulation.

A significant reduction can in general be attained with the approach as will be shown later, and the method makes it very easy and fast to try out new focusing schemes once the basic data has been simulated. This would demand a full recalculation in the old approach.

f_0	7 MHz	Transducer center frequency
f_s	120 MHz	Sampling frequency
D	3	Decimation factor for RF data
h_e	5 mm	Height of element
pitch	$\lambda/2$	Distance between elements
w	$0.9\lambda/2$	Width of element
k_e	$0.1\lambda/2$	Kerf between elements
N_e	64	Number of elements

Table F.1: Simulation parameters for phased array imaging.

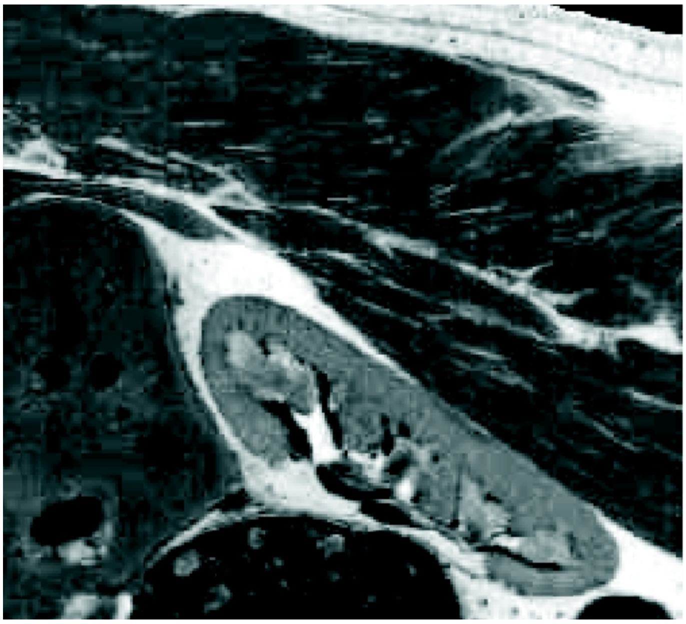


Figure F.2: Optical image from the visual human project of a right kidney and a liver lobe.

3 Examples

All the examples in the following section have been made by a modified version of the Field II simulation system. The parameters used in the simulation are shown in Table F.1.

An artificial kidney phantom based on data from the Visible Human Project[1] has been used as the simulation object. The phantom consists of 200,000 point scatterers within a box of 100 $\times$ 100 $\times$ 35 mm (lateral, axial, elevation dimension), which gives a realistic size for a full computer simulation of a clinical image. The optical image in Fig. F.2 is used for scaling the standard deviation of the Gaussian random amplitudes of the scatterers. The relation between the gray level value in the image and the scatterer amplitude scaling is given by:

$$a = 10 \cdot \exp(\mathrm{img}(\vec{r}_k)/100) \qquad \text{(F.12)}$$

where img is the gray-level image data with values from 0 to 127, and $\vec{r}_k$ is the discrete position of the scatterer. This scaling ensures a proper dynamic range in the scatterering from the various structures. The resulting image using a synthetic aperture simulation is shown in Fig. F.3

[1] Optical, CT and MR images from this project can be found at: http://www.nlm.nih.gov/research/visible/visible_human.html

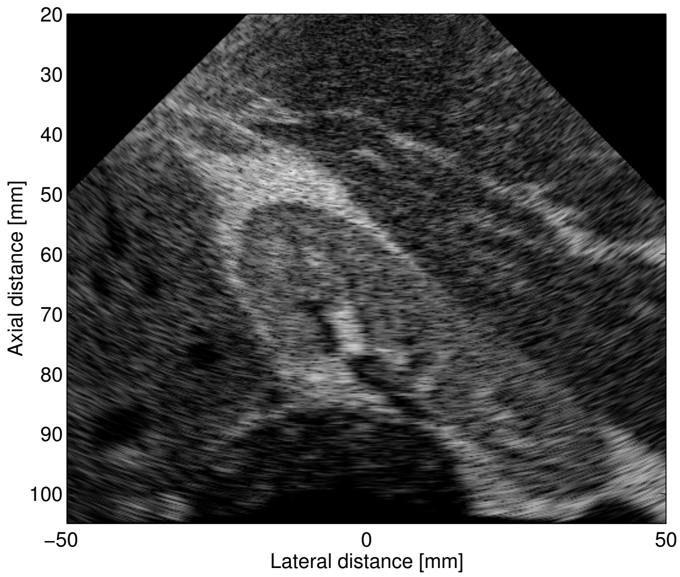


Figure F.3: Synthetic ultrasound image of right kidney and liver based on an optical image from the visual human project.

N_s	N_e	Method	Time [s]	Improvement
20,000	32	Line	2944	-
20,000	32	Synthetic	494	5.96
20,000	64	Line	6528	-
20,000	64	Synthetic	1108	5.65
200,000	64	Line	60288	-
200,000	64	Synthetic	10972	5.49

Table F.2: Simulation times for scanning the human-liver phantom.

The simulations have been carried out using a standard PC with 512 MBytes RAM and an Athlon 800 MHz CPU running Linux RedHat 6.2. The various simulation times are shown in Table F.2. These data also include the beam focusing for the synthetic aperture simulation, which took 9 seconds in all cases. It can be seen that the simulation times increase nearly linearly with the number of elements, and linearly with the number of scatterers. The improvement for a phased array image with 128 lines is roughly a factor 5.5 to 6, which lies between the two boundaries given earlier. The actual improvement is dependent on the object size, transducer, sampling frequency, CPU, and memory interface, and the numbers will be different for other scan situations.

A three-dimensional scanning has also been implemented. A two-dimensional sparse array ultrasound transducer was used, and a volume consisting of 64 $\times$ 64 lines with 200,000 scatterers was made. Simulating one line takes 455 seconds, which gives a full simulation time of 1,863,680 seconds (21 days and 13 hours). Using the new approach the whole volume can be simulated in one pass. This takes 6,810 seconds and the beamforming 288 seconds, which in total gives an improvement in simulation time by a factor of 262. A further benefit is that different focusing strategies also can be tested without a new simulation, and a new volume image can then be made in 288 seconds.

A parallel simulation has also been performed using a Linux cluster consisting of 16 PCs with dual 600 Pentium III processors and 256 MBytes of RAM for every 2 CPUs. The scatterers are then divided into 32 files and the simulation is performed in parallel on all machines. The total simulation time is 935 seconds for the 2D simulation and beamforming takes 9 seconds, when using 200,000 scatterers for the phantom. A full simulation of a clinical image, thus, takes 15 minutes and 44 seconds, which is acceptable for iterative work. This should be compared with 16 hours and 45 minutes on one CPU using the line based simulation. This approach can also be employed for the three-dimensional scanning and can reduce the time for one volume to roughly 212+288 = 500 seconds. Simulating 15 volumes of data corresponding to one second of volumes for a 3D scanner can then be done in 7,500 seconds or roughly 2 hours.

Acknowledgment

This work was supported by grant 9700883 and 9700563 from the Danish Science Foundation and by B-K Medical A/S, Gentofte, Denmark.

APPENDIX

G

Experimental ultrasound system for real-time synthetic imaging

Authors : Jørgen Arendt Jensen (1), Ole Holm (2), Lars Joost Jensen (2), Henrik Bendsen (2), Henrik Møller Pedersen (1,2), Kent Salomonsen (2), Johnny Hansen (2) and Svetoslav Nikolov (1)

Published : Proceedings of the IEEE Ultrasonics Symposium, Lake Tahoe, 1999.

Abstract

Digital signal processing is being employed more and more in modern ultrasound scanners. This has made it possible to do dynamic receive focusing for each sample and implement other advanced imaging methods. The processing, however, has to be very fast and cost-effective at the same time. Dedicated chips are used in order to do real time processing. This often makes it difficult to implement radically different imaging strategies on one platform and makes the scanners less accessible for research purposes. Here flexibility is the prime concern, and the storage of data from all transducer elements over 5 to 10 seconds is needed to perform clinical evaluation of synthetic and 3D imaging. This paper describes a real-time system specifically designed for research purposes.

The purpose of the system is to make it possible to acquire multi-channel data in real-time from clinical multi-element ultrasound transducers, and to enable real-time or near real-time processing of the acquired data. The system will be capable of performing the processing for the currently available imaging methods, and will make it possible to perform initial trials in a clinical environment with new imaging modalities for synthetic aperture imaging, 2D and 3D B-mode and velocity imaging.

The system can be used with 128 element transducers and can excite 128 channels and receive and sample data from 64 channels simultaneously at 40 MHz with 12 bits precision. Data can be processed in real time using the system's 80 signal processing units or it can be stored directly in RAM. The system has 24 GBytes RAM and can thus store 8 seconds of multi-channel data. It is fully software programmable and its signal processing units can also be reconfigured under software control. The control of the system is done over an Ethernet using C and Matlab. Programs for doing *e.g.* B-mode imaging can directly be written in Matlab and executed on the system over the net from any workstation running Matlab. The overall system concept is presented and an example of a 20 lines script for doing phased array B-mode

imaging is presented.

1 Introduction

New imaging techniques based on synthetic imaging are currently being suggested and investigated [61, 75]. The methods can potentially increase both resolution and frame rate, since the images are reconstructed from RF data from the individual transducer elements. Hereby a perfectly focused image in both transmit and receive can be made. Research in real time 3D imaging is also underway [92, 1]. The purpose is to make systems that in real time can display a pyramidal volume of the heart, where different slices hereafter can be visualized. These images have a poor signal-to-noise ratio, and several groups are working on employing coded signals to enhance the signal-to-noise ratio [167].

All of the above techniques require digital signal processing on the signals from the individual transducer elements, and in some instances it is also necessary to send out coded signals on the individual elements. For research purposes this can be difficult to attain with commercial scanners, since they are often highly integrated and it is difficult to access individual signals. Programming commercial scanners for new imaging techniques is often also either cumbersome or impossible. It is, thus, beneficial to develop a dedicated research system, that can acquire, store, process, and display ultrasound images from multi-element transducers.

2 System specification

The purpose of the system is to make possible the acquisition of multi-channel data in real-time from clinical multi-element ultrasound transducers, and to enable real-time or near real-time processing of the acquired data. The system will be capable of performing the processing for all currently available imaging methods, and will make it possible to carry out initial trials with new imaging modalities for synthetic aperture imaging, 3D imaging, and 2D and 3D velocity estimation. It is capable of working in a clinical environment to evaluate the performance of various algorithms. The system is specifically intended for research purposes, and is not intended for commercial use.

The function of the system is defined by the different imaging methods for which it can be used. Each of the imaging types will be described and the consequence for the system then given.

Linear array imaging: A linear array image is generated by a multi-element transducer with 128 to 256 elements. The beam is moved by selecting *e.g.* 64 adjacent elements and emitting a focused beam from these. The focusing in receive is also done by a number of elements, and multiple foci are used. Apodization in both transmit and receive are often applied. The focusing delay in both transmit and receive are both less than 40 μs. The number of active elements is usually 32 to 64. The transducer frequency is from 2 MHz to 10 MHz. Imaging is done down to a depth of 30 cm.

The demands on the system is, thus, for 64 channels simultaneous sampling at 40 MHz. The maximum delay in both transmit and receive is 40 μs. The maximum time to sample one line is $2 \times 0.3/1540 + 40 \cdot 10^{-6} = 430$ μs corresponding to 17,200 samples at 40 MHz.

Phased array imaging: The beam is here electronically swept over the imaging area by using a 128 to 256 element array. All the elements are used at the same time, and focusing time delays used are less than 50 μs. The transducer frequency is from 2 MHz to 10 MHz. Investigations are done to a depth of 20 cm.

The demands on the system is, thus, for 128 channels sampling at 40 MHz. The demands on delay, sampling time and storage are the same as for linear array imaging.

Flow estimation, spectrum: Beamforming is done in one direction with either a linear or phased array. The flow signal from blood has 40 dB less power than that from stationary tissue. The dynamic range of the flow signal is 30 dB. The effective number of bits must be 12 or more, when the signals from all channels have been combined. The pulse emitted can have from 4 to 16 periods of the center frequency of the transducer or a coded signal can be employed.

Flow imaging: Imaging is done by pulsing repeatedly in one direction and then change the direction to generate an image. An image can therefore be assembled from up to a 1000 pulse emissions.

Three-dimensional imaging: A matrix element transducer is used with up to 40 $\times$ 40 elements. Only some of the elements are used for transmit and receive. The area of the elements is small and pulsing should be done with 100 to 300 volts. Coded signals should be used. Coded pulses with up to 64 cycle periods must be possible with a high amplitude accuracy. This corresponds to emission over a period of 32 μs with a sampling frequency of 40 MHz and an accuracy of 12 bits.

Phasing is done with a delay up to 50 μs, and parallel lines are generated by using parallel beam formers and reusing data from one pulse emission. The system must be capable of reading the data sampled from one elements a number of times, and use different phasing schemes for each cycle through the data.

Synthetic aperture imaging: A standard linear or phased array multi-element transducer is used. Pulsing is done on a few elements and the received response is acquired for all elements. The image is then reconstructed from only a small number of pulse emissions by using the data from all the elements.

This type of imaging needs large amounts of storage and the ability to reuse the data for the different imaging directions. This should be solved by having a multi-processor system connected to the sampling system for storage and image reconstruction.

It must be possible to acquire several seconds of data. Assuming sampling in 80 % of the time at 40 MHz, and 8 seconds of sampling gives a storage need of 256 Mbytes per channel.

3 System realization

The multi-channel sampling and processing system consists of four distinct modules: The transmit unit, the receive/transmit (Rx/Tx) amplifiers, receiver sampling unit, and the sync/master unit. The main blocks are depicted on the drawing in Fig. G.1. The connection to the transducer is through a 128 wire coaxial cable through the Rx/Tx amplifiers. The transmitter sends the signals through the transmit amplifier, and the receiver unit samples the

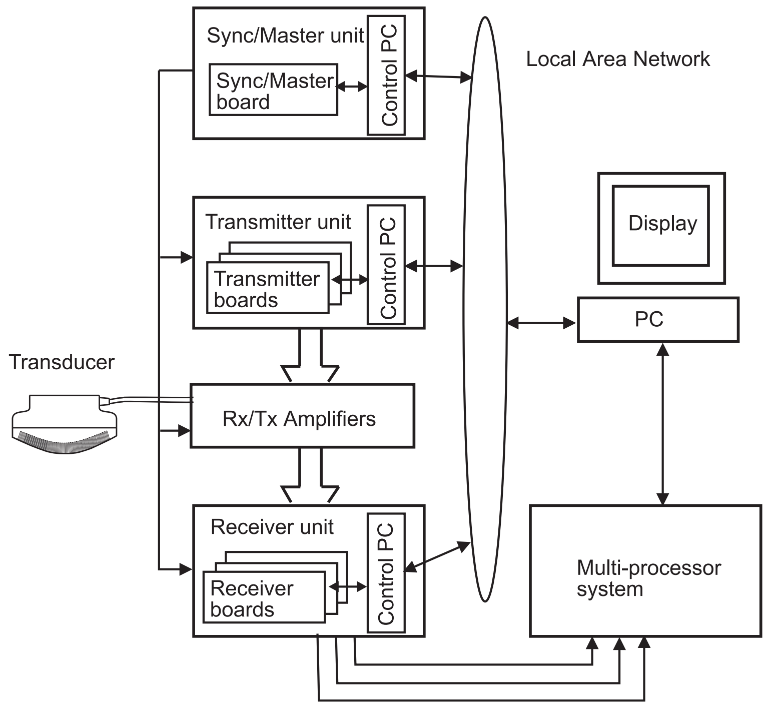


Figure G.1: Overall diagram of system.

amplified and buffered signals from the Rx/Tx amplifiers. The sync/master unit holds a crystal oscillator and controls the central timing of the scanning process. The overall operation of the system is controlled through a number of single board PCs in the individual units interconnected through a standard 100 Mbit Ethernet. The waveforms and phasing data are transmitted from the controlling PC to the transmitters and receiver boards. The data from the sampling is processed by FPGAs (field programmable gate arrays), that can be configured for specific signal processing tasks over the net. One Focus FPGA is used for each element and a Sum FPGA is placed for each eight elements. The processed and summed signal can then be routed from Sum FPGA to Sum FPGA. The resulting signal is read by one or more signal processors, that can be connected through serial interfaces capable of transmitting 40 Mbytes per second. Each processor has 6 such links. Data and programs are transferred through these links. The beamformed and envelope detected signal is send via the link channels to the PC for display.

The following paragraphs detail the overall design of the individual boards.

3.1 Transmitter

The transmitter is capable of generating an arbitrary transmitted pulse with a bandwidth below 20 MHz. The transmitter consists of 8 transmitter boards each equipped with 16 channels. In total the transmitter controls 128 channels.

A transmitter board consists of two control FPGA's, 16 pulse RAM's, two delay RAM's and sixteen 12 bit digital to analog converters (DAC).

Each of the 16 channels has a pulse RAM memory implemented as a 128 k $\times$ 12 bit SRAM, which allows the user to store for instance 32 different pulse emissions of 100 μs duration.

The delay RAM holds the start address of the pulse emission in the pulse RAM and the corresponding delay for each line. The delay RAM is implemented as 32 k $\times$ 32 bit SRAM. At the start of each line the pulse emission is delayed according to the delay value for each channel.

3.2 Receiver

The Receiver board is illustrated in Fig. G.2. The board samples and processes 8 analog signals selected from 16 inputs.

The receiver, transmitter and sync/master boards are accessible from a general purpose LINUX based compact PCI single board PC, which controls the different boards.

A 12 bit analog to digital converter (ADC) samples one input channel and the data is temporarily stored in a SRAM buffer. The data in the SRAM buffers are processed by the Focus FPGA using the Focus RAM. The Sum FPGA processes the focused data from the 8 Focus FPGA's and transfers it to the Analog Devices signal processor (ADSP) or stores it in the storage RAM. The functionality of the individual blocks of the receiver board is explained in further detail below.

Focus FPGA

The Focus FPGA controls the initial storing and processing of the sampled data. The Focus FPGA fetches the sampled data from the SRAM and the corresponding focusing parameters from the Focusing RAM and processes the data before transferring the result to the Sum FPGA.

Two independent memory burst SRAM banks are used to bank switch between the sampled data and processed data. While the sampled data is being written to one of the two banks, the other bank can be read by the Focus FPGA. Each SRAM is implemented as 256 kbytes, which is equivalent to a line length of 3.3 ms sampled at 40 MHz.

The basic focusing algorithm uses a combination of coarse and fine delays. The coarse delay is in steps of the 25 ns sampling interval, and it is implemented as a general table look up address generator. For each sample a 16 bit address index is read from the SDRAM. In this way a random sorting algorithm can be implemented. The fine delay is implemented as a linear interpolation with two signed 10 bit apodization coefficients, which are read from the focusing RAM for each sample.

The Focus FPGA is implemented using a XILINX device from the Virtex family: XCV300 in a 352 pin BGA package speed grade -4. The simple B-mode beamformer described above uses less than 10% of the logical resources of the chip. This makes it possible to investigate hardware implementations of more advanced beamformers including pulse compression, synthetic aperture, and parallel beamforming.

Sum FPGA

The Sum FPGA is used to perform digital signal processing on the 8 channels. The most basic operation is to sum the 8 focused channels. Further, it is used as the gateway between the eight independent sampling channels and the ADSP. The Sum FPGA controls the 2 Gbyte storage SDRAM.

When the focusing is done in the Focus FPGA, the 8 channels are added to the accumulated

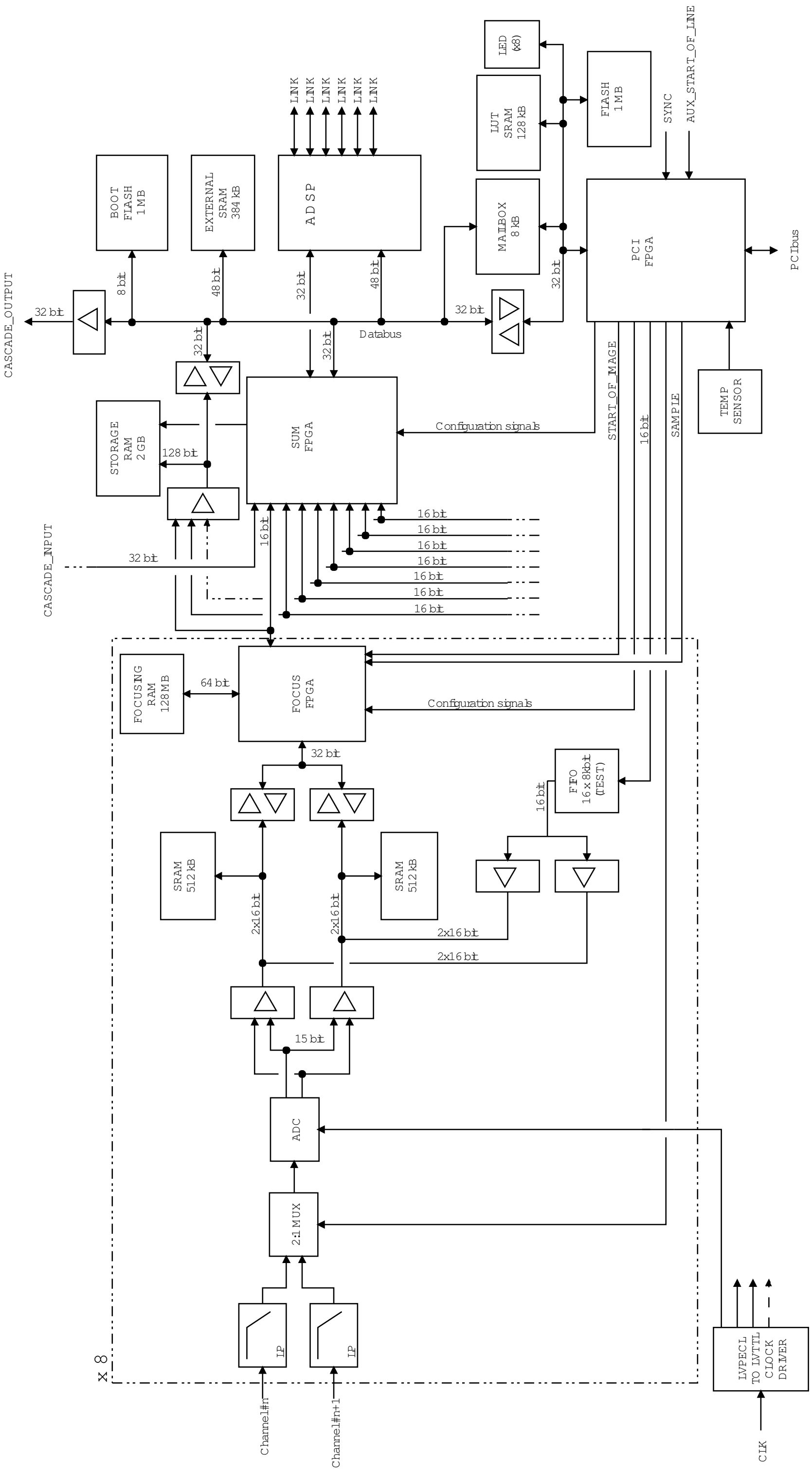


Figure G.2: Main diagram of Receiver board.

sum that is passed to the next Receiver board using a high speed cascade bus connecting the Sum FPGA's directly with each other. The last Sum FPGA in the chain uses the ADSP link ports to transfer the final result to the multiprocessor cluster.

The Sum FPGA is implemented using a XILINX device from the Virtex family: XCV1000 in a 560 pin BGA package speed grade -4. The simple design described above uses less than 5% of the logical resources of the chip.

3.3 Sync/master unit

The Sync/master unit controls the timing of the system. A highly stable oven controlled crystal oscillator generates the 40 MHz clock frequency. The clock jitter is below 5 ps. The clock is distributed using coax cables and emitter coupled logic (ECL) in order to minimize jitter. The timing source also transmits a synchronization signal. The receiver and transmitter uses the SYNC signal to start and stop the sampling cycles. An image consists of a number of lines, each with a transmission of a pulse and reception of the echoes. The transmitter and receiver generates an internal LINE signal from the SYNC signal to control the sampling process for the received signal.

4 Programming of the system

From the software point of view, the system consists of several computers that are directly connected to a number of transmitter and receiver boards. The computers are linked by a LAN, and uses Linux as operating system and TCP/IP as the underlying communication protocol.

The software was designed to meet the following requirements:

- Flexibility and ease of use. It is of prime importance that new imaging methods can be quickly implemented with a minimal amount of programming also for new users.
- Data encapsulation. All data for the imaging algorithm are stored in the boards of the scanner.
- Distributed computations. The computers work independently one from another, and each calculates only those imaging parameters concerning the boards plugged in it.
- Portability. The software is written in ANSI C and is platform independent.

The client/server communication model was adopted for the software. The computers controlling the boards run a *server* program. The server waits for requests coming from the LAN and processes them. The requests can be sent by any *client* program running on a computer connected to the LAN using the TCP/IP communication protocol.

Figure G.3 shows the client-server model of the software. At start-up the server detects which boards are in the PCI enclosures. The computers can handle any combination of transmitter and receiver boards, plugged into the same PCI back-plane. The server is in idle mode until an event occurs. In the case of a hardware malfunction, the server sends an emergency message to a program, called *monitor daemon*. Another event is a request by the *client*. The request can be for transferring parameters to the boards, for performing some calculations, or for retrieving data from the boards to the computer with the user.

The interface to the client program is implemented as a MATLAB tool-box. The function calls are implemented to be as close to the functions in the simulation program Field II [19] as possible. Algorithms created using Field II can thereby easily be tested on the scanner with

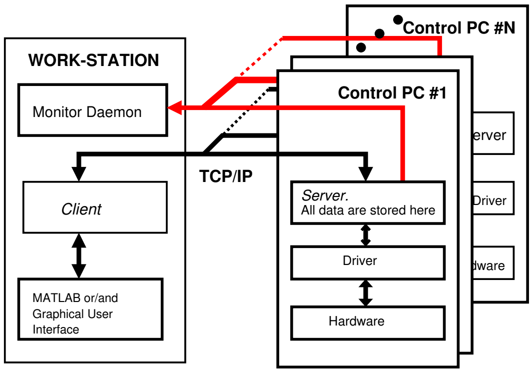


Figure G.3: Client-server model of software.

only minor changes in the Matlab program. An example for setting the system to perform phased array B-mode imaging is shown below:

```
% Auto-detect and initialize the system
sys_init('auto');

% Set the pulse repetition frequency
sys_set_fprf(fprf);

% Set the sampling range gate in receive
sys_set_sampling_interval(start_depth, end_depth);

% Set the number of scan-lines per frame
sys_set_no_lines(no_lines);

% Define the transducer. Necessary for the delay calculations
tr_linear_array(no_elements, width, height, kerf);
% Do for all lines in the image:
for line_no = 1 : no_lines

  % Set the pulse and the apodization for the current line
  xmt_excitation(waveform(line_no));
  xmt_apodization(line_no,xmt_apo(line_no, : ));
  rcv_apodization(line_no,times, rcv_apo(line_no, : , : ));
  % Set the focus, defined in 3D coordinates
  xmt_focus(line_no,focus(line_no));
  rcv_dynamic_focus(line_no, theta(line_no), fi(line_no));
end

% Set the time-gain compensation curve
tmg_tgc(tgc_vector);
% Start the continuous imaging process
tmg_start
```

In order to make the system perform linear array imaging only one line needs to be added, which changes the origin of the individual scan-lines.

5 Conclusion

A system for doing research into new imaging methods in medical ultrasound has been described. The system is capable of emitting complex, arbitrary waveforms, and can sample and store data for all transducer channels in real time. Sufficient storage is included in the system for 8 seconds of data at a sampling frequency of 40 MHz for a 64 element transducer. The system is easily programmable through Matlab and a network interface. It can perform real time processing and image display for all commercially available ultrasound systems and can function in a clinical environment.

Acknowledgment

This work was supported by grant 9700883 and 9700563 from the Danish Science Foundation, by B-K Medical A/S, and by grant EF-782 from the Danish Academy of Technical Sciences.

APPENDIX

H

Real time 3D visualization of ultrasonic data using a standard PC

Authors : Svetoslav Ivanov Nikolov, Juan Pablo Gómez Gonzaléz, Jørgen Arendt Jensen

Published : Paper presented at Ultrasonics International 2001, Delft. To be published in Ultrasonics.

Abstract

This paper describes a flexible, software-based scan converter capable of rendering 3D volumetric data in real-time on a standard PC. The display system is part of a Remotely Accessible and Software-Configurable Multichannel Ultrasound Sampling system (RASMUS system) developed at the Center for Fast Ultrasound Imaging.

The RASMUS system is connected to a PC via the link channels of an ADSP 21060 signal processor. A DMA channel transfers the data from the ADSP to a memory buffer. A software library based on OpenGL uses this memory buffer as a texture map that is passed to the graphics board.

The scan conversion, image interpolation, and logarithmic compression are performed by the graphics board, thus reducing the load on the main processor to a minimum.

The scan conversion is done by mapping the ultrasonic data to polygons. The format of the image is determined only by the coordinates of the polygons allowing for any kind of geometry to be displayed on the screen. Data from color flow mapping is added by alpha-blending. The 3D data are displayed either as cross-sectional planes, or as a fully rendered 3D volume displayed as a pyramid. All sides of the pyramid can be changed to reveal B-mode or C-mode scans, and the pyramid can be rotated in all directions in real time.

The PC used in RASMUS has a 800 MHz AMD Athlon processor and an NVIDIA GeForce2 video card. The resolution is 1280x1024 pixels, 32 bits per pixel. The system can display a B-mode video at 360 frames per second (fps), or it can simultaneously display up to 4 fps. A 3D volume is rendered at 41 fps.

1 Introduction

Modern ultrasound scanners are able to display B-mode (brightness mode) data, combined B-mode data and a sonogram, and/or B-mode data with a superimposed color flow map. The visualization of ultrasonic data involves the following operations:

1. Logarithmic compression.
2. Scan conversion.
3. Display and user interface.

The most computationally extensive operation is the scan conversion. The number of lines is not sufficient and interpolation of the data must be done [13, 14]. Usually specialized hardware for the scanner is designed to speed up the process. Because the implementation is hardware dependent, only a fixed number of imaging modalities are supported. Every new modality requires the introduction of changes in the implementation.

The purpose of this work was to develop a flexible software-based display of ultrasound data.

The advantages of using a software solution on a standard PC platform are:

1. Low cost. The ever growing PC market provides inexpensive components. The development is confined only to the parts specific to the visualization of ultrasound data, not to the problem of developing video controllers.
2. Scalability. The actual performance of the system can be scaled according to the needs of the scanner.
3. Reusability. The use of open standards ensures the continuance and reusability of the work already done. Using higher-level, platform-independent solutions ensures that the manufacturer is not bound to specific hardware.
4. Flexibility. A software solution can be easily reconfigured to accommodate different image geometries. The user interface can be based on a standard graphical user interface of a PC.

The idea of using a software solution and standard microprocessor equipment is not new [15]. A pure software solution is computationally extensive and requires a lot of processing power and memory bandwidth. The previous work in scan conversion has been focused on the algorithmic optimization [14, 16, 168]. No matter how efficient one algorithm is, it still requires a lot of processing power of the central processor. This problem has been addressed by the vendors of video cards for the needs of the PC gaming and computer aided design (CAD) industries. Specialized graphics processing units with 256 bit cores and 128 bit wide memory interface easily outperform most of the general-purpose processors in displaying graphics. Hence, a software display using hardware acceleration is capable of not only displaying all of the imaging modalities, but it can render a 3D volume in real time. This can be done by defining the mapping of the coordinates not as look-up tables but as geometric transformation of polygons. A number of libraries and application programmer interfaces (API) are readily available for the purpose. We have chosen to implement the display using OpenGL because many of the video cards have

OpenGL hardware acceleration, and it is available for most operating systems and computer languages.

In the next section the requirements to the program are given. Then the implementation is described in Section 3, and finally the performance is measured and discussed in Section 5.

2 System specification

This section describes the requirements on the software display so it can be used with the RASMUS system [41] for all image types. The demands are:

1. The display must be integrated with RASMUS and show images in real time at more than 30 frames per second. This imposes certain requirements on the transfer rate of RF data to the program. Assuming a speed of sound of $c = 1540$ m/s, a pulse repetition frequency of $f_{prf} = 5000$ Hz, a sampling frequency $f_s = 40$ MHz, and an image depth from 0.005 to 0.15 m, the maximum amount of data transfered is 71.8 MB/sec.

2. It should be capable of displaying images saved on a disk for off-line examination. Hence the display part should be separate from the data transferring part. Thus the source of data can be supplied either from the scanner, from a hard disk, or any other source. The implementation must therefore include inter-process communications.

3. Ability to display more than one images simultaneously. RASMUS is an experimental system and the result of different algorithms can then be compared.

4. Both B-mode images and color flow maps must be displayed. Both types of images can be of any geometry - phased array sector images, linear array recti-linear images, linear array tilted images, convex array sector images, 3D images with a pyramidal scan, or 3D images obtained as a rectilinear scan, etc.

5. The implementation must reduce the load on the main processor to a minimum. Thus all the processing related to the image display must be carried out by the video controller.

6. The program must be platform independent and portable. The development must be based on open standards and standard software tools.

3 Realization

The whole implementation is software based. The programs are written in ANSI C using the OpenGL library. The reason for using OpenGL is that the interface is a de facto standard, that is stable and fast. A number of implementations exist such as those by SGI and Microsoft as well as the Open Source version represented by the library ”Mesa” (http://www.mesa3d.org). The library is supported under most of the operating systems such as Linux, the various flavors of UNIX and MS Windows. The sources are compiled for most of the hardware platforms and OpenGL oriented drivers are available from the vendors of video controllers. The whole development of the image display was done under Linux, using open source tools.

There are three implementation issues:

1. How to feed data in real-time to the display system PC.
2. How to define the geometry of the image and and how to map the data.
3. How to display 3D data in a way that is both familiar to the clinician, and makes it intuitive to manipulate in 3D and relate it to the anatomy.

3.1 Data transfer

The data is fed to the PC display system through a Blacktip PCI board ver 2.0 by BittWare Research Systems from the RASMUS system. The PCI board features an Analog Devices ADSP 21060 processor with 2 MB external RAM. The RASMUS system [41] consists of a number of custom made transmitter and receiver boards. Each of the receiver boards contains 8 receive channels. The sampled data from the individual channels is appropriately delayed, summed and passed to the next board for cascade summing. The last of the receiver boards hosts an ADSP 21060 which is connected via two of link channels to the ADSP on the Blacktip board in the PC display system. The transfer rate is 80 MB/sec. The PLX PCI 9060 PCI controller of the board transfers the data from the RAM memory on the board to the memory of the computer. Working as a bus master it is capable of a theoretical peak transfer rate of 128 MB/sec. The control of the Blacktip board is performed by custom-made software and drivers. The driver reserves the last 4 MB of the physical memory of the computer and uses them for two buffers for the DMA transfer. The buffers are swapped - while one is being filled in with new data, the other is being used as a texture map by the visualization program, which maps the buffers in its memory space.

3.2 Display of phased array images

Fig. H.1 illustrates the scan conversion geometry. The phased-array data are acquired as a function of angle θ and depth r. It is mapped to the screen Cartesian coordinates x, y, along the horizontal and vertical directions respectively as shown in Fig. H.1 (a). Traditionally this is done by using either look-up tables or some modified algorithm for line drawing. Modern video controllers, such as the graphical processing unit GeForce 256 by NVIDIA (http://www.nvidia.com) used in the display system, can render up to 15 million polygons per second. The gaps in the formed image are filled by interpolation. It has been found that bilinear interpolation [14] is a good trade-off between quality and computational cost. Modern accelerated video-cards support bilinear and trilinear interpolation as a standard hardware accelerated feature. Thus, the problems associated with the interpolation are taken care of by the video controller without any software overhead. Having this in mind it is quite natural to let the specialized hardware take care of the video display. The only problem that the developer must consider is how to express the mapping of the data.

The approach used in this work is to split the the input data and the output image into convex polygons. The envelope detected lines are stacked together as shown in Fig. H.1(a) with dots. Each column corresponds to one scan line in the image. The arcs of the output image are approximated by a number of segments, and the whole image is represented by a number of polygons shown in the figure with dashed lines. The segments approximating the arcs have length equal to $L \approx r \cdot (\theta_{\max} - \theta_{\min})/N$, where N is the number of segments and r is the radius of the respective arc. The bigger the number of polygons the better the approximation of the

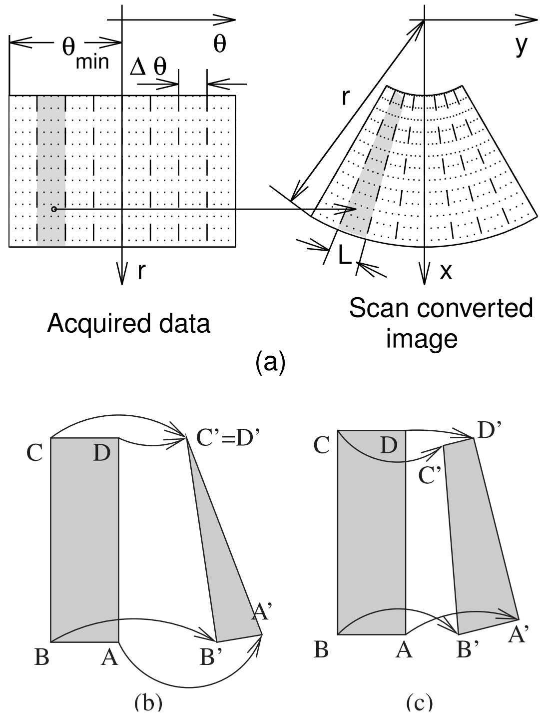


Figure H.1: Mapping of ultrasonic data from polar to Cartesian coordinate system.

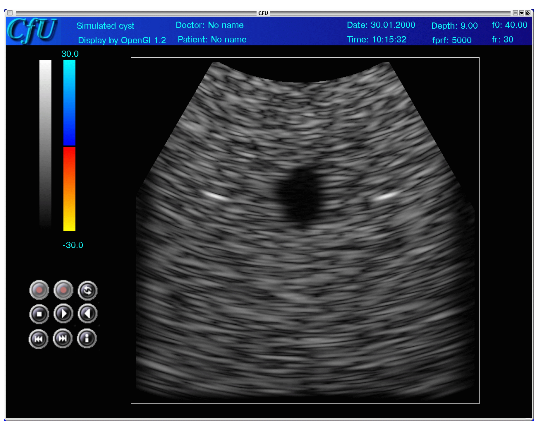


Figure H.2: Phased array image.

output image is. Such division corresponds to equal in size rectangles in the (r,θ) space of the acquired data (The lateral size of these rectangles, $\Delta\theta$, can be different than the actual angle step used for the scanning). The only task for the software is to specify the mapping from the normalized coordinates in the data space to the coordinates of the polygons comprising the scan-converted image as shown in Fig. H.1(b) and (c). The normalized coordinates of the four vertexes A,B,C,D of each of the rectangles from the data space are mapped to the coordinates of the four vertexes A',B',C',D' of the corresponding tetragon in the image space. Fig. H.1(b) and (c) show the cases when the start scan depth is 0 and bigger than 0, respectively. In OpenGL the data are treated as a textures mapped onto the polygons. The choice of the number of tetragons forming the image has a major impact on the on the appearance of the image. As a rule of a thumb the number of polygons should be at least twice the number of acquired scan

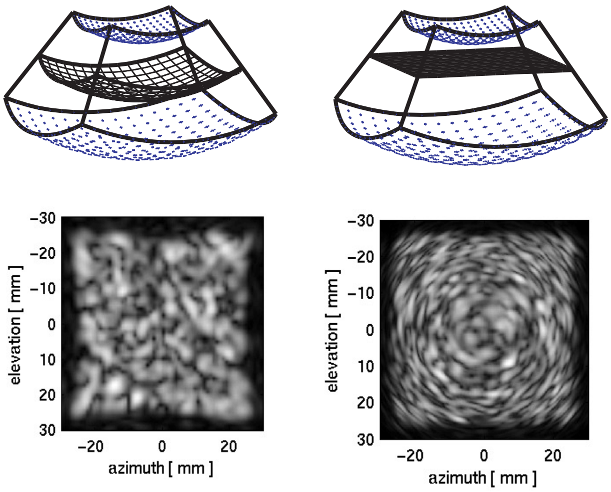


Figure H.3: Displaying data parallel to the transducer surface: c-scan (left) and a cross-section with a plane (right).

lines. This ensures that a sample from one scan line is used in several polygons resulting in a gradual transition. A video card like the one used in the display system can render more than 15 million triangles per second so the number of used polygons can be rather high. The size of the image shown in Fig. H.2 is 512×512 pixels (this is the size of the active area - without the user interface), the width of the sector ($\theta_{max} - \theta_{min}$) is $\pi/3$, the number of scan lines is 64, and the number of polygons N used to create it, is $N = 256$.

The user has the ability to zoom in the image. For every new zoom, the coordinates of the displayed polygons are recalculated in order to maintain the quality of the display. The number of *displayed* polygons is kept constant.

For linear scan images there is no need for going from one coordinate system to another (for example from polar to Cartesian coordinates). In this case only a single rectangle should be displayed on the screen. The only concern of the developer is to ensure a fixed aspect ratio.

The color flow mapping is done by defining a second set of polygons, corresponding to the coordinates of the velocity estimates to be displayed. The display is done by using alpha blending, which is setting a transparency coefficient.

3.3 Display of 3D images

The display software is targeted at real-time 3D display for volumetric data acquired at a rate higher than 10 volumes/sec [6]. The acquisition of volumetric data in real-time can be done using a 2D matrix arrays. The data in this case is in polar coordinates and must be converted to Cartesian ones. The display must at the same time be both familiar to the medical doctor and giving orientation in the 3D volume. It was decided that the best display for the purpose would be the visualization of a solid with the geometric shape of the scanned volume as the one shown in Fig. H.3. The faces of the solid are rendered with the envelope detected data. The clinician is able to view the data inside the volume by moving in and out the bounding planes. If the top and the bottom (the faces parallel to the transducer surface) are chosen to be planar, then the speckle exhibits radial symmetry as seen from the right image in Fig. H.3. This

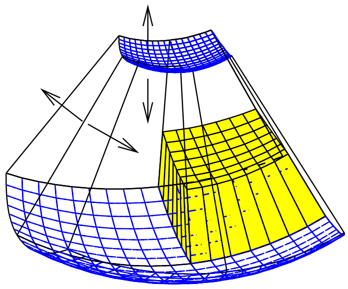


Figure H.4: Display of and navigation in 3D ultrasonic data.

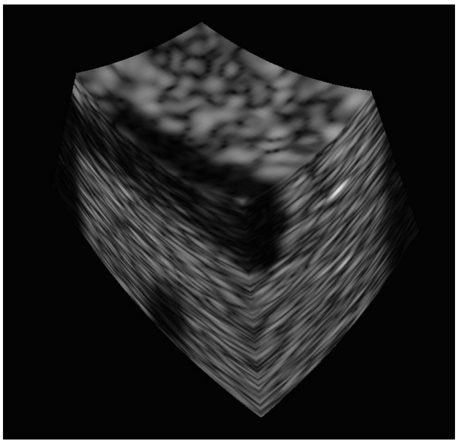


Figure H.5: 3D volume.

radial symmetry is an artifact from the interpolation. Instead the program allows the clinician to change the start and end depths of the displayed volume. The browsing of data in the elevation and azimuth directions is done by changing the start and end angles. These operations are illustrated in Fig. H.4. An example of the rendered volume is given in Fig. H.5. The same display method as presented in Section 3.2 used with a 3D orientation of the polygons. The zooming in the volume is implemented by changing the position of the view camera. Because not all of the sides of the volume are facing the user, only those that are actually visible are passed for rendering, thus reducing the overall data stream.

4 Performance

Two aspects of the system are of interest: (1) the ability of interfacing it to a real-time acquisition hardware, (2) the ability to display data in real time. The tests are based on measured performance rather than estimates of the computational cost.

The PC on which the development and tests were done comprises an ASUS K7V motherboard with VIA KX133 chip set and 800 MHz AMD Athlon processor. The video card is an ASUS V6800 with an NVIDIA GeForce 256 GPU and 32 MB DDR. The software was tested under a standard Red Hat 6.2 distribution of Linux. The X server is XFree86 (http://www.xfree86.org), version 4.0.1 with drivers from NVIDIA.

The sustained transfer rate of data from the memory of the Blacktip board to a shared memory buffer in the PC (including the latencies of interrupt handling, task switching, display of information, etc.) is 76 MB/s, sufficient for the real-time connectivity of the system.

The performance of the visualization software was tested for the display of a phased array sector image and a linear array rectangular image. The size of the input data set was the same (1024×64). The display mode was set to 1280×1024 pixels, 32 bits per pixel. Table **??** shows the speed of display for a window with size 800×600 and 1280×1024 pixels (full screen). The size of the displayed image is 512×512 and 860×860 pixels, respectively. For the cases of more than one views, the images had the same type of geometry either sector (pie) or rectangle, since these exhibit the extremities in performance. The cases of mixed geometries have performance ranging between these extreme cases. The speed of display scales with the size of the image that must be rendered and the size of the data that must be transfered to the memory of the graphics board. This is quite clearly manifested in the speed for the 3D display, in which the change in the number of displayed polygons almost does not affect the display.

5 Conclusion

This paper showed that it is possible to use a software display program based on standard software tools and standard PC hardware to scan convert, interpolate and display 2D and 3D ultrasound images in real time. Taking advantage of the hardware acceleration of OpenGL up to 360 phased array B-mode images and 41 3D volumes per second can be displayed which is sufficient for all clinical applications in medical ultrasound. All the processing is done by the video controller, thus, reducing the load on the main processor, which is used for other tasks such as envelope detection, user interface and control.

6 Acknowledgements

This work was supported by grant 9700883 and 9700563 from the Danish Science Foundation, and by B-K Medical A/S, Gentofte, Denmark.

APPENDIX

I

Velocity estimation using synthetic aperture imaging

Authors : Svetoslav Ivanov Nikolov and Jørgen Arendt Jensen

Published : This paper will be presented at the IEEE Ultrasonics Symposium Atlanta, 2001, and published in the proceedings of the conference. The work on the article is still in progress. Only the abstract is given.

Abstract

In a previous paper we have demonstrated that the velocity can be estimated for a plug flow using recursive ultrasound imaging. The approach involved the estimation of the velocity at every emission and using the estimates for motion compensation. An error in the estimates, however, would lead to an error in the compensation further increasing the error in the estimates.

In this paper the approach is further developed such that no motion compensation is necessary. In recursive ultrasound imaging a new high resolution image is created after every emission. The velocity was estimated by cross correlating RF lines from two successive emissions n and n+1 and then average over a number of lines. In the new approach images n and n+N, n+1 and n+N+1 are cross correlated, where N is the number of emissions for one image. These images experience the same phase distortion due to motion and therefore have a high correlation without motion compensation. The advantage of the approach is that a color flow map can be created for all directions in the image simultaneously at every emission, which makes it possible to average over a large number of lines. This makes stationary echo canceling easier and significantly improves the velocity estimates. Simulations using the Field II program yielded a bias of the estimate of -3.5 a mean standard deviation less than 2.0 velocity profile.

The method was investigated using our real-time experimental system RASMUS using a B-K Medical 8804, 7.5 MHz linear array. The pulse repetition frequency was 7 kHz. The method was applied on a flow phantom made using 100 um polymer spheres dissolved in water, driven by a SMADEGARD Type EcoWatt 1 pump generating a peak velocity of 0.42 m/s. The mean velocity was determined by a MAG 1100 flow meter by Danfoss. The length of the tube was 1 m, allowing for a parabolic profile to develop. The number of color flow lines per frame was 90, the number of color flow maps per second was 7000, the bias of the estimate was -7 % and the mean standard deviation was 1.9 %.

The system is capable of acquiring *in-vivo* images and these will be presented.

APPENDIX

J

Description of the XTRA system

The XTRA system shown in Figure J.1 was specified by Søren Kragh Jespersen. The sole purpose of the system is to be able to acquire sampled data from individual channels. It works on the principle of synthetic receive aperture imaging. It has 64 transmit channels that can be multiplexed to 256 transducer elements. In receive only a single channel is accessible. The channel can be multiplexed to 256 transducer elements. The delays in transmit can be set individually for up to 64 channels. The transmit clock frequency is 60 MHz. The largest realizable delay in transmit is 2048 clock cycles.

In receive different sampling frequencies are available, by dividing the master clock. The A/D converter is 12 bit. The system can perform also time gain compensation (TGC). A new TGC value is set at a frequency of 1.875 MHz.

The transmit pulses are RF pulses at a center frequency specified by the user. The duration of the pulse is specified in a number of half cycles.

The system parameters are summarized in Table J.1.

The transducer connected to the system is a 192 elements linear array. The parameters of the

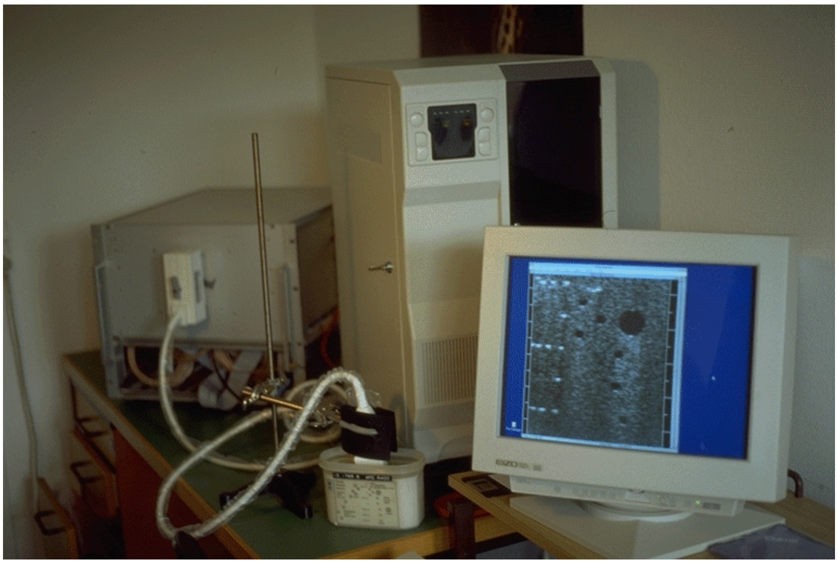


Figure J.1: The hardware of the XTRA system. The 19” rack is seen on the left, the control PC is in the middle, and the screen showing a conventional image of tissue mimicking cyst and a wire phantom is seen on the right. (Picture taken from [39])

Parameter	Value	Unit
A/D converter	1	-
Transmit channels	64	-
Clock frequency	60	MHz
Sampling frequency	60, 40, 30, 24, 20, $\frac{120}{7}$, 12	MHz
TGC output frequency	1.875	MHz
A/D word length	12	bits
Transmit sequence length	2048	samples
TGC sequence length	1024	samples
Delay resolution	≈ 16.67	ns

Table J.1: Selected specification of XTRA system.

Elements	192	-
Center frequency	7.5	MHz
Pitch	209	μm
Kerf	30	μm
Height	4	mm
Elevation focus	20	mm
Fractional bandwidth	70	%

Table J.2:

transducer are given in Table J.2

All the experiments with XTRA were done using pulses at a center frequency of 5 MHz, thus increasing the ratio wavelength/pitch. The latter effectively "pushes" the grating lobes outside the scanned region. The sampling frequency was always set to 40 MHz.

Bibliography

[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., Ferroelec., Freq. Contr.*, 45:980–988, 1998.

[2] H. Ş. Bilge, M. Karaman, and M. O'Donnell. Motion estimation using common spatial frequencies in synthetic aperture imaging. In *Proc. IEEE Ultrason. Symp.*, volume 2, pages 1551–1554, 1996.

[3] R. Y. Chiao, L. J. Thomas, and S. D. Silverstein. Sparse array imaging with spatially-encoded transmits. In *Proc. IEEE Ultrason. Symp.*, pages 1679–1682, 1997.

[4] J. Henderson, K. Willson, J. R. Jago, and T. A. Whittingham. A survey of the acoustic outputs of diagnostic ultrasound equipment in current clinical use. *Ultrasound Med. Biol.*, 21:699–705, 1995.

[5] J. J. Wild. The use of ultrasonic pulses for the measurement of biologic tissues and the detection of tissue density changes. *Surgery*, 27:183–188, 1950.

[6] O. T. von Ramm, S. W. Smith, and Henry G. Pavy. High-speed ultrasound volumetric imaging system – Part II: Parallel processing and image display. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 38:109–115, 1991.

[7] T. R. Nelson and T. D. Elvins. Visualization of 3D ultrasound data. *IEEE Comp. Graph. App.*, 13:50 –57, 1993.

[8] T. R. Nelson and D. H. Pretorius. Interactive acquisition analysis and visualization of sonographic voliume data. *Int. J. Imag. Sys. Tech.*, 8:26–37, 1997.

[9] T. R. Nelson and D. H. Pretorius. Three-dimensional ultrasound imaging. *Ultrasound Med. Biol.*, 24:1243–1270, 1998.

[10] J. Carr. *Surface reconstruction in 3D medical imaging*. PhD thesis, University of Canterbury, Christchurch, New Zealand, 1996.

[11] J. T. Yen, J. P. Steinberg, and S. W. Smith. Sparse 2-D array design for real time rectilinear volumetric imaging. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 47:93–110, 2000.

[12] Y. Kim, J. H. Kim, C. Basoglu, and T. C. Winter. Programmable ultrasound imaging using multimedia technologies: a next-generation ultrasound machine. *IEEE Trans. Inform. Technol. Biomed.*, 1:19 –29, 1997.

[13] S. C. Leavitt, B. F. Hunt, and H. G. Larsen. A scan coversion algorithm for displaying ultrasound images. *HP-Journal*, 34:30–34, 1983.

[14] A. P. Berkhoff, H. J. Huismann, J. M. Thijssen, E. M. G. P. Jacobs, and R. J. F. Homan. Fast scan conversion algorithms for displaying ultrasound sector images. *Ultrason. Imaging*, pages 87–100, 1994.

[15] C. Basoglu, Y. Kim, and V. Chalana. A real-time scan conversion algorithm on commercially available microprocessors. *Ultrason. Imaging*, 18:241–260, 1996.

[16] D. C. Chang, W. R. Wu, and J. H. Jeng. An efficient architecture for ultrasonic scan conversion for implementing the cubic convolution interpolation. In O. Nalcioglu, editor, *IEEE Nucl. Sci. Symp. Confer, Rec. (Cat. No. 97CH36135)*, pages 1546–1550. IEEE, 1997.

[17] J. P. G. Gonzaléz. Hardware accelerated ultrasound visualization system. Master's thesis, Department of Information Technology, Technical University of Denmark, 2000.

[18] S. I. Nikolov, J. P. G. Gonzaléz, and J. A. Jensen. Real time 3D visualization of ultrasonic data using a standard pc. *In press. To be published in Ultrasonics*, X:XXX–XXX, 2001.

[19] J. A. Jensen. Field: A program for simulating ultrasound systems. *Med. Biol. Eng. Comp.*, 10th Nordic-Baltic Conference on Biomedical Imaging, Vol. 4, Supplement 1, Part 1:351–353, 1996b.

[20] M. F. Insana and D. G. Brown. Acoustic scattering theory applied for soft biological tissues. In K. K. Shung and G. A. Thieme, editors, *Ultrasonic scattering in biological tissues*, pages 76–124. CRC Press, Boca Raton, 1993.

[21] A. D. Pierce. *Acoustics, An Introduction to Physical Principles and Applications*. Acoustical Society of America, New York, 1989.

[22] D. H. Johnson and D. E. Dudgeon. *Array signal processing. Concepts and techniques.* Prentice-Hall., Englewood Cliffs, New Jersey, 1993.

[23] J. W. Goodman. *Introduction to Fourier Optics. Second edition.* The McGraw-Hill Companies, Inc., New York, 1996.

[24] J. A. Jensen. Linear description of ultrasound imaging systems. Technical report, Dept. of Information Technology, Technical University of Denmark, 1999.

[25] M. Arditi, F. S. Forster, and J. Hunt. Transient fields of concave annular arrays. *Ultrason. Imaging*, 3:37–61, 1981.

[26] J. A. Jensen. Ultrasound fields from triangular apertures. *J. Acoust. Soc. Am.*, 100(4):2049–2056, 1996a.

[27] J. A. Jensen. Simulating arbitrary geometry ultrasound transducers using triangles. In *Proc. IEEE Ultrason. Symp.*, volume 2, pages 885–888, 1996c.

[28] J. A. Jensen. A new calculation procedure for spatial impulse responses in ultrasound. *J. Acoust. Soc. Am.*, pages 3266–3274, 1999.

[29] P. R. Stepanishen. Pulsed transmit/receive response of ultrasonic piezoelectric transducers. *J. Acoust. Soc. Am.*, 69:1815–1827, 1981.

[30] J. A. Jensen. A model for the propagation and scattering of ultrasound in tissue. *J. Acoust. Soc. Am.*, 89:182–191, 1991a.

[31] J. A. Jensen, D. Gandhi, and W. D. O'Brien. Ultrasound fields in an attenuating medium. In *Proc. IEEE Ultrason. Symp.*, pages 943–946, 1993.

[32] M. Soumekh. *Fourier array imaging*. Prentice-Hall., New Jersey, 1994.

[33] M. Soumekh. *Synthetic aperture radar. Signal processing with MATLAB algorithms*. John Wiley & Sons, Inc., New York, 1999.

[34] O. T. von Ramm and S. W. Smith. Beam steering with linear arrays. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, BME-30:438–452, 1983.

[35] B. D. Steinberg. Digital beamforming in ultrasound. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 39:716–721, 1992.

[36] K. E. Thomenius. Evolution of ultrasound beamformers. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 42:1615–1622, 1996.

[37] P. Munk. *Estimation of blood velocity vectors using ultrasound*. PhD thesis, Technical University of Denmark. Lyngby., 2000.

[38] S. I. Nikolov. User's guide to the Beamformation Toolbox. release 1.3. Technical report, Dept. of Information Technology, Technical University of Denmark, 2000.

[39] S. K. Jespersen. *Tools for improving the diagnosis of atherosclerotic plaque using ultrasound. PhD thesis*. Department of Information Technology, Technical University of Denmark, 1997.

[40] J. A. Jensen. Specification of a multi-channel sampling and processing system for ultrasound imaging. Technical report, Dept. of Information Technology, Technical University of Denmark, 1998.

[41] J. A. Jensen, O. Holm, L. J. Jensen, H. Bendsen, H. M. Pedersen, K. Salomonsen, J. Hansen, and S. Nikolov. Experimental ultrasound system for real-time synthetic imaging. In *Proc. IEEE Ultrason. Symp.*, volume 2, pages 1595–1599, 1999.

[42] L. J. Jensen. Receiver unit design specification 100030 ds. Technical report, I/O Consulting Aps., 1998.

[43] J. A. Jensen and S. I. Nikolov. Fast simulations of ultrasound images. In *Proc. IEEE Ultrason. Symp.*, volume 2, pages 1721–1724, 2000.

[44] T. I. Laakso, V. Välimäki, M. Karjalainen, and U. K. Laine. Splitting the unit delay. *IEEE Sig. Proc. Mag.*, pages 30–60, 1996.

[45] S. Freeman, P. Li, and M. O'Donnel. Retrospective dynamic transmit focusing. *Ultrason. Imaging*, 17:173–196, 1995.

[46] J. A. Jensen and P. Gori. Improved accuracy in the estimation of blood velocity vectors using matched filtering. In *Proc. IEEE Ultrason. Symp.*, volume 2, pages 1525 –1530, 2000.

[47] A. W. Rihaczek. *Principles of high-resolution radar*. McGaw-Hill Book Company, New York, 1969.

[48] R. Y. Chiao, K. W. Rigby, and D. G. Wildes. Optimization of 1.5D arrays. In *Proc. IEEE Ultrason. Symp.*, pages 1835–1838, 1998.

[49] S. M. Gehlbach and R. E. Alvarez. Digital ultrasound imaging techniques using vector sampling and raster line reconstruction. *Ultrason. Imaging*, 2:83–107, 1981.

[50] G. R. Lockwood and F.S. Foster. Optimizing sparse two-dimensional arrays using an effective aperture approach. In *Proc. IEEE Ultrason. Symp.*, pages 1497–1501, 1994.

[51] G. R. Lockwood, P. Li, M. O'Donnel, and F. S. Foster. Optimizing the radiation pattern of sparse periodic linear arrays. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 43:7–14, 1996.

[52] R. T. Hoctor and S. A. Kassam. The unifying role of the coarray in aperture synthesis for coherent and incoherent imaging. In *IEEE Proc.*, volume 78, pages 735–752, 1990.

[53] A. V. Openheim and R. W. Schafer. *Discrete-time signal processing*. Prentice Hall International, Inc., New Jersey, 1989.

[54] M. E. Anderson and G. E. Trahey. A seminar on *k*-space applied to medical ultrasound. Technical report, Dept. of Biomedical Engineering, Duke University, 1999.

[55] W. F. Walker and G. E. Trahey. The application of *k*-space in pulse-echo ultrasound. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 45:541–558, 1998.

[56] S. Bennett, D. K. Peterson, D. Corl, and G. S. Kino. A real-time synthetic aperture digital acoustic imaging system. In P. Alais and A. F. Metherell, editors, *Acoust. Imaging*, volume 10, pages 669–692, 1982.

[57] K. Mayer, R. Marklein, K. J. Langenberg, and T. Kreutter. Three-dimensional imaging system based on Fourier transform synthetic aperture focusing technique. *Ultrasonics*, 28:241–255, 1990.

[58] L. J. Busse. Three-dimensional imaging using a frequency-domain synthetic aperture focusing technique. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 39:174–179, 1992.

[59] M. O'Donnell and L. J. Thomas. Efficient synthetic aperture imaging from a circular aperture with possible application to catheter-based imaging. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 39:366–380, 1992.

[60] J. T. Ylitalo. In-vitro study of computed ultrasound imaging method. In *Proc. IEEE Ultrason. Symp.*, pages 1577–1580, 1994.

[61] J. T. Ylitalo and H. Ermert. Ultrasound synthetic aperture imaging: monostatic approach. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 41:333–339, 1994.

[62] J. T. Ylitalo. Synthetic aperture ultrasound imaging using a convex array. In *Proc. IEEE Ultrason. Symp.*, pages 1337–1340, 1995.

[63] Juha Ylitalo. A fast ultrasonic synthetic aperture imaging method: application to NDT. *Ultrasonics*, pages 331–333, 1996.

[64] J. T. Ylitalo. On the signal-to-noise ratio of a synthetic aperture ultrasound imaging method. *Eur. J. Ultrasound*, pages 277–281, 1996.

[65] M. Karaman, P. C. Li, and M. O'Donnel. Synthetic aperture imaging for small scale systems. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 42:429–442, 1995.

[66] T. Robini, T.Rastello, D. Vray, and I. E. Magnin. Space-variant deconvolution for synthetic aperture imaging using simulated annealing. In *International Conference on Image Processing*, pages 432–435, 1997.

[67] D. Vray, T. Rastello, F. Schoofs, and P. Delachartre. Improving the spatial resolution of intravascular ultrasound imaging with synthetic aperture Fourier-based methods. In *Proc. IEEE Ultrason. Symp.*, pages 1531–1534, 1997.

[68] S. Watts. *Lecture 7, Synthetic aperture techniques*. The Institution of Electrical Engineers, Thorn EMI Electronics Ltd. Bayes, Middlesex, UK, 1975.

[69] L. F. Nock and G. E. Trahey. Synthetic aperture imaging in medical ultrasound with correction for motion artifacts. In *Proc. IEEE Ultrason. Symp.*, pages 1597–1601, 1990.

[70] L. F. Nock and G. E. Trahey. Synthetic receive aperture imaging with phase correction for motion and for tissue inhomogenities - part I: basic principles. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 39:489–495, 1992.

[71] L. F. Nock and G. E. Trahey. Synthetic receive aperture imaging with phase correction for motion and for tissue inhomogenities - part II: effects of and correction for motion. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 39:496–501, 1992.

[72] W. F. Walker and G. E. Trahey. Real-time synthetic receive aperture imaging: experimental results. In *Proc. IEEE Ultrason. Symp.*, pages 1657–1660, 1994.

[73] R. Y. Chiao and L. J. Thomas. Analytic evaluation of sampled aperture ultrasonic imaging techniques for NDE. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 41:484–493, 1994.

[74] S. Holm and H. Yao. Improved framerate with synthetic transmit aperture imaging using prefocused subapertures. In *Proc. IEEE Ultrason. Symp.*, pages 1535–1538, 1997.

[75] C. R. Hazard and G. R. Lockwood. Theoretical assessment of a synthetic aperture beamformer for real-time 3-D imaging. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 46:972–980, 1999.

[76] M. H. Bae, M. K. Jeong, T. K. Song, and Y. B. Ahn. Experimental study of transmit synthetic focusing combined with receive dynamic focusing in B-mode ultrasound imaging systems. In *Proc. IEEE Ultrason. Symp.*, pages 1261–1264, 1999.

[77] F. N. Ucar and Mustafa Karaman. Beam space processing for low-cost scanners. In *Proc. IEEE Ultrason. Symp.*, pages 1349–1352, 1996.

[78] M. Karaman and M. O'Donnel. Subaperture processing for ultrasonic imaging. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 45:126–135, 1998.

[79] M. Karaman, H. Ş. Bilge, and M. O'Donnel. Adaptive multi-element synthetic aperture imaging with motion and phase aberation correction. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 42:1077–1087, 1998.

[80] S. I. Nikolov, K. Gammelmark, and J. A. Jensen. Recursive ultrasound imaging. In *Proc. IEEE Ultrason. Symp.*, volume 2, pages 1621–1625, 1999.

[81] J. T. Ylitalo. Improving contrast resolution in ultrasound holographic B-imaging: comparison between simulation and experiment. In *Proc. IEEE Ultrason. Symp.*, pages 1011–1014, 1993.

[82] C. Passmann and H. Ermert. Adaptive 150 MHz ultasound imaging of the skin and the eye using an optimal combination of short pulse mode and pulse compression mode. In *Proc. IEEE Ultrason. Symp.*, pages 1291–1294, 1995.

[83] C. Passmann and H. Ermert. A 100-MHz ultrasound imaging system for dermatologic and ophtalmologic diagnostics. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 43:545–551, 1996.

[84] C. H. Frazier and W. D. O'Brien. Synthetic aperture techniques with a virtual source element. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 45:196–207, 1998.

[85] T. Mochizuki C. Kasai. High speed acquisition scanner and three-dimensional display system using ultrasonic echo data. In *Proc. IEEE Ultrason. Symp.*, volume 2, pages 1459–1462, 1990.

[86] L. Gatzoulis, T. Anderson, S. D. Pye, R. O'Donnel, C. C. McLean, and W. N. McDicken. Scanning techniques for three-dimensional forward looking intravascular ultrasound imaging. *Ultrasound Med. Biol.*, 26:1461–1474, 2000.

[87] P. H. Mills and H. Fuchs. 3D ultrasound display using optical tracking. In *Proc. Visualization in Biomed. Comp.*, pages 490 –497, 1990.

[88] D. F. Leota, P. R. Detmer, O. H. Gilja, J. Jong, R. W. Martin, J. F. Primovich, K. W. Beach, and D. E. Strandness. Three-dimensional ultrasound imaging using multiple magnetic tracking systems and miniature magnetic sensors. In *Proc. IEEE Ultrason. Symp.*, pages 1415–1418, 1995.

[89] D. H. Turnbull and F. S. Foster. Simulation of b-scan images from two-dimensional transducer arrays: part II - comparison between linear and two-dimensional phased arrays. *Ultrason. Imaging*, 14:344–353, 1992.

[90] X. Wang, C. J. Ritchie, and Y. Kim. Elevation direction deconvolution in three-dimensional ultrasound imaging. *IEEE Trans. Med. Imag.*, 15:389–394, 1996.

[91] S. I. Nikolov and J. A. Jensen. 3D synthetic aperture imaging using a virtual source element in the elevation plane. In *Proc. IEEE Ultrason. Symp.*, volume 2, pages 1743–1747, 2000.

[92] S. W. Smith, H. G. Pavy, and O. T. von Ramm. High-speed ultrasound volumetric imaging system – Part I: Transducer design and beam steering. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 38:100–108, 1991.

[93] C. R. Cooley and B. S. Robinson. Synthetic aperture imaging using partial datasets. In *Proc. IEEE Ultrason. Symp.*, pages 1539–1542, 1994.

[94] G. R. Lockwood and F.S. Foster. Design of sparse array imaging systems. In *Proc. IEEE Ultrason. Symp.*, pages 1237–1243, 1995.

[95] R. Y. Chiao and L. J. Thomas. Aperture formation on reduced-channel arrays using the transmit-receive apodization matrix. In *Proc. IEEE Ultrason. Symp.*, pages 1567–1571, 1996.

[96] T. Rastello, D. Vray, and J. Chatillon. Spatial under-sampling of ultrasound images using Fourier-based synthetic aperture focusing technique. In *Proc. IEEE Ultrason. Symp.*, pages 823–826, 1997.

[97] C. R. Hazard and G. R. Lockwood. Effects of motion artifacts on a synthetic aperture beamformer for real-time 3D ultrasound. In *Proc. IEEE Ultrason. Symp.*, pages 1221–1224, 1999.

[98] D. H. Turnbull and F. S. Foster. Beam steering with pulsed two-dimensional transducer arrays. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 38:320–333, 1991.

[99] W. J. Hendricks. The totally random versus the bin approach for random arrays. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 39:1757–1762, 1995.

[100] J. O. Erstad and S. Holm. An approach to the design of sparse array systems. In *Proc. IEEE Ultrason. Symp.*, pages 1507–1510, 1994.

[101] C. Boni, M. Richard, and S. Barbarossa. Optimal configuration and weighting of nonuniform arrays according to a maximum islr criterion. In *Proc. Int. Conf. Acous., Speech, Sig. Pro.*, volume 5, pages V/157–V/160, 1994.

[102] P. K. Weber, R. M. Schmitt, B. D. Tylkowski, and J. Steck. Optimization of random sparse 2-D transducer arrays for 3-D electronic beam steering and focusing. In *Proc. IEEE Ultrason. Symp.*, pages 1503–1506, 1994.

[103] A. Austeng, S. Holm, P. K. Weber, and K. Iranpour. 1D and 2D algorithmically optimized sparse arrays. In *Proc. IEEE Ultrason. Symp.*, volume 2, pages 1683–1686, 1997.

[104] S. Holm and B. Elgetun. Optimization of the beampattern of 2D sparse arrays by weighting. In *Proc. IEEE Ultrason. Symp.*, pages 1345–1348, 1995.

[105] S. Holm, B. Elgetun, and G. Dahl. Properties of the beampattern of weight- and layout-optimized sparse arrays. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 44:983–991, 1997.

[106] A. Trucco. Thinning and weighting of large planar arrays by simulated annealing. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 46:347–355, 1999.

[107] J. L. Schwartz and B. D. Steinberg. Ultrasparse, ultrawideband arrays. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 45:376–393, 1998.

[108] T. S. Sumanaweera, J. Schwartz, and D. Napolitano. A spiral 2D phased array for 3D imaging. In *Proc. IEEE Ultrason. Symp.*, pages 1271–1274, 1999.

[109] E. D. Light, J. O. Fiering, P. A. Hultman, W. Lee, and S. W. Smith. Optimizing sparse two-dimensional arrays using an effective aperture approach. In *Proc. IEEE Ultrason. Symp.*, pages 1497–1501, 1994.

[110] O. T. von Ramm. 2-D arrays. *Ultrasound Med. Biol.*, 26:S10–S12, 2000.

[111] G. R. Lockwood and F. S. Foster. Optimizing the radiation pattern of sparse periodic two-dimensional arrays. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 43:15–19, 1996.

[112] S. I. Nikolov and J. A. Jensen. Application of different spatial sampling patterns for sparse array transducer design. *Ultrasonics*, 37:667–671, 2000.

[113] A. Austeng and S. Holm. Simple weighting to enhance sparse periodic arrays. In *Proc. Int. Conf. Acous., Speech, Sig. Pro.*, volume 5, pages 3109–3112, 2000.

[114] A. Austeng and S. Holm. Sparse arrays for real-time 3D imaging, simulated and experimental results. In *Proc. IEEE Ultrason. Symp.*, volume 2, pages 1187–1190, 2000.

[115] M. O'Donnell. Coded excitation system for improving the penetration of real-time phased-array imaging systems. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 39:341–351, 1992.

[116] S. Qian and D. Chen. *Joint time-frequency analysis*. Prentice Hall PTR, New Jersey, 1996.

[117] Y. Takeuchi. Chirped excitation for ¡ -100 db time sidelobe echo sounding. In *Proc. IEEE Ultrason. Symp.*, volume 2, pages 1309–1314, 1995.

[118] Y. Takeuchi. Coded excitation for harmonics imaging. In *Proc. IEEE Ultrason. Symp.*, volume 2, pages 1433–1436, 1996.

[119] T. X. Misaridis and J. A. Jensen. An effective coded excitation scheme based on a predistorted fm signal and an optimized digital filter. In *Proc. IEEE Ultrason. Symp.*, volume 2, pages 1589–1593, 1999.

[120] T. X. Misaridis, M. H. Pedersen, and J. A. Jensen. Clinical use and evaluation of coded excitation in b-mode images. In *Proc. IEEE Ultrason. Symp.*, volume 2, pages 1689–1693, 2000.

[121] L. R. Welch. Pulse compression ultrasound for minimization of transmitted peak power,. In *Proceedings of the 1994 20th Annual Northeast Bioengineering Conference*, pages 88–91, 1994.

[122] L. R. Welch and M. D. Fox. Practical spread spectrum pulse compression for ultrasonic tissue imaging. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 45:349 –355, 1998.

[123] S. Jian and E. S. Ebbini. A new coded-excitation ultrasound imaging system. i. basic principles. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 43:131–140, 1996.

[124] S. Jian and E. S. Ebbini. A new coded-excitation ultrasound imaging system. ii. operator design. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 43:141–148, 1996.

[125] Y. Takeuchi. An investigation of a spread energy method for medical ultrasound systems. part one: theory and investigation. *Ultrasonics*, July:175–182, 1979.

[126] Y. Takeuchi. An investigation of a spread energy method for medical ultrasound systems. part two: proposed system and possible problems. *Ultrasonics*, September:219–224, 1979.

[127] R. Y. Chiao and L. J. Thomas. Synthetic transmit aperture using orthogonal golay coded excitation. In *Proc. IEEE Ultrason. Symp.*, pages 1677–1680, 2000.

[128] R. Y. Chiao, L. Y. Mo, A. L. Hall, S. C. Miller, and K. E. Thomenius. B-mode blood flow (b-flow) imaging. In *Proc. IEEE Ultrason. Symp.*, pages 1469–1472, 2000.

[129] D. Ž. Đoković. Equivalence classes and representatives of golay sequences. *Discrete Mathematics*, 189:79–93, 1998.

[130] S. D. Silverstein. Application of orthogonal codes to the calibration of active phased array antennas for communication sarellites. *IEEE Trans. Sig. Pro.*, 45:206–218, 1997.

[131] R. C. Gonzales and R. E. Woods. *Digital image processing.* Addison-Wesley Publishing Company, New York, 1992.

[132] J. A. Jensen. *Estimation of Blood Velocities Using Ultrasound: A Signal Processing Approach.* Cambridge University Press, New York, 1996.

[133] O. Bonnefous and P. Pesqué. Time domain formulation of pulse-doppler ultrasound and blood velocity estimation by cross correlation. *Ultrason. Imaging*, 8:73–85, 1986.

[134] S. G. Foster. *A pulsed ultrasonic flowmeter employing time domain methods*. PhD thesis, Dept. Elec. Eng., University of Illinois, Urbana, Ill., 1985.

[135] S. G. Foster, P. M. Embree, and W. D. O'Brien. Flow velocity profile via time-domain correlation: Error analysis and computer simulation. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 37:164–175, 1990.

[136] Paul Michael Embree. *The accurate ultrasonic measurement of the volume flow of blood by time domain correlation*. PhD thesis, Dept. Elec. Eng., University of Illinois, Urbana, Ill., 1986.

[137] P. M. Embree and W. D. O'Brien. The accurate ultrasonic measurement of volume flow of blood by time-domain correlation. In *Proc. IEEE Ultrason. Symp.*, pages 963–966, 1985.

[138] P. M. Embree and W. D. O'Brien. Volumetric blood flow via time-domain correlation: Experimental verification. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 37:176–189, 1990a.

[139] B. A. J. Angelsen. A theoretical study of the scattering of ultrasound from blood. *IEEE Trans. Biomed. Eng.*, BME-27:61–67, 1980a.

[140] O. Bonnefous. Statistical analysis and time processes applied to velocity measurement. In *Proc. IEEE Ultrason. Symp.*, pages 887–892, 1989.

[141] I. A. Hein, J. T. Chen, W. K. Jenkins, and W. D. O'Brien. A real-time ultrasound time-domain correlation blood flowmeter, part I: Theory and design. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 40:768–775, 1993.

[142] I. A. Hein and W. D. O'Brien. A real-time ultrasound time-domain correlation blood flowmeter, part II: Performance and experimental verification. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 40:776–785, 1993b.

[143] J. A. Jensen. Stationary echo canceling in velocity estimation by time-domain cross-correlation. *IEEE Trans. Med. Imag.*, 12:471–477, 1993b.

[144] Jens E. Wilhjelm. *Ultrasound FM Doppler systems for flow profiling with comparison to PW Doppler systems*. PhD thesis, Worcester Polytechnic Institute, Mass., 1991a.

[145] J. E. Wilhjelm and P. C. Pedersen. Target velocity estimation with FM and PW echo ranging Doppler systems – part I: Signal analysis. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 40:366–372, 1993a.

[146] J. E. Wilhjelm and P. C. Pedersen. Target velocity estimation with FM and PW echo ranging Doppler systems – part II: System analysis. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 40:373–380, 1993b.

[147] R. L. Maurice and M. Bertrand. Speckle-motion artifact under tissue shearing. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 46:584–594, 1999.

[148] B. H. Friemel, L.N. Bohs, K. R. Nightingale, and G. E. Trahey. Speckle decorrelation due to two-dimensional flow gradients. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 45:317–327, 1998.

[149] H. Ş. Bilge and M. Karaman. Correlation analysis of ultrasonic speckle for motion estimation. In *Dig. Sig. Proc. Workshop Proceedings IEEE*, pages 211–214, 1996.

[150] S. I. Nikolov, K. Gammelmark, and J. A. Jensen. Velocity estimation using recursive ultrasound imaging and spatially encoded signals. In *Proc. IEEE Ultrason. Symp.*, volume 2, pages 1473–1477, 2000.

[151] C. G. Caro, T. J. Pedley, R. C. Schroter, and W. A. Seed. Mechanics of the circulation. In A.C. Guyton, editor, *Cardiovascular Physiology*. Medical and Technical Publishers, 1974.

[152] O. Bonnefous. Measurement of the complete (3d) velocity vector of blood flows. In *Proc. IEEE Ultrason. Symp.*, pages 795–799, 1988.

[153] J. A. Jensen and I. R. Lacasa. Estimation of blood velocity vectors using transverse ultrasound beam focusing and cross-correlation. In *Proc. IEEE Ultrason. Symp.*, pages 1493–1497, 1999.

[154] I. R. Lacasa. Estimation of blood velocities with ultrasound. Master's thesis, Department of Information Technology, Technical University of Denmark, 1999.

[155] G. E. Trahey, J. W. Allison, and O. T. von Ramm. Angle independent ultrasonic detection of blood flow. *IEEE Trans. Biomed. Eng.*, BME-34:965–967, 1987.

[156] Gregg E. Trahey, S. M. Hubbard, and Olaf T. von Ramm. Angle independent ultrasonic blood flow detection by frame-to-frame correlation of B-mode images. *Ultrasonics*, 26:271–276, 1988.

[157] L. N. Bohs and G. E. Trahey. A novel method for angle independent ultrasonic imaging of blood flow and tissue motion. *IEEE Trans. Biomed. Eng.*, 38:280–286, 1991.

[158] L. N. Bohs, B. J. Geiman, K. R. Nightingale, C. D. Choi, B. H. Friemel, and G. E. Trahey. Ensemble tracking: a new method for 2D vector velocity measurement. In *Proc. IEEE Ultrason. Symp.*, pages 1485 – 1488, 1995.

[159] B. Geiman, L. Bohs, M. Anderson, and G. Trahey. Initial experimental results using ensemble tracking for 2D vector velocity measurement. In *Proc. IEEE Ultrason. Symp.*, pages 1241–1244, 1996.

[160] L. N. Bohs, B. J. Geiman, S. M. Breit, and G. E. Trahey. Ensemble tracking for 2D vector velocity measurement: experimental and initial clinical results. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 45:912–924, 1998.

[161] S. K. Jespersen, J. E. Wilhjelm, and H. Sillesen. Multi-angle compound imaging. *Ultrason. Imaging*, 20:81–102, 1998.

[162] J. A. Jensen and P. Munk. Computer phantoms for simulating ultrasound B-mode and CFM images. In S. Lees and L. A. Ferrari, editors, *Acoustical Imaging*, volume 23, pages 75–80, 1997.

[163] A. T. Kerr and J. W. Hunt. A method for computer simulation of ultrasound Doppler color flow images – II. Simulation results. *Ultrasound Med. Biol.*, 18:873–879, 1992b.

[164] F. Forsberg, H. Oung, and H. V. Ortega. Doppler simulation for pulsatile flow having nonaxial components. In *Proc. IEEE Ultrason. Symp.*, pages 1713–1716, 1994.

[165] L. E. Kinsler, A. R. Frey, A. B. Coppens, and J. V. Sanders. *Fundamentals of Acoustics*. John Wiley & Sons, New York, third edition, 1982.

[166] J. A. Jensen and N. B. Svendsen. Calculation of pressure fields from arbitrarily shaped, apodized, and excited ultrasound transducers. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 39:262–267, 1992.

[167] P. Li, E. Ebbini, and M. O'Donnell. A new filter design technique for coded excitation systems. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 39:693–699, 1992.

[168] C. Fritsch, M. Parilla, O. Martinéz, and D. Jimenéz. A multirate scan conversion method. *Ultrason. Imaging*, 38:172–182, 2000.