NOTE TO USERS

Page(s) not included in the original manuscript and are unavailable from the author or university. The manuscript was scanned as received.

This reproduction is the best copy available.

UMI[®]

Beamforming for Multisensor Stethoscopes

Submitted by

Anita M. McKee, B. Eng., B. Sc.

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment
of the requirements for the degree of

Master of Applied Science

Ottawa-Carleton Institute of Electrical and Computer Engineering
Department of Systems and Computer Engineering

Carleton University
Ottawa Ontario, K1S 5B6
Canada
August 2004

©Copyright Anita M. McKee, 2004

Library and
Archives Canada

Bibliothèque et
Archives Canada

Published Heritage
Branch

Direction du
Patrimoine de l'édition

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

ISBN: 0-612-97426-X

Our file Notre référence

ISBN: 0-612-97426-X

The author has granted a non-exclusive license allowing the Library and Archives Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

The undersigned recommend to the
Faculty of Graduate Studies and Research

Acceptance of the thesis:

Beamforming for Multisensor Stethoscopes

Submitted by

Anita M. McKee, B. Eng., B. Sc.

In partial fulfillment of the requirements of the degree of
Master of Applied Science

Thesis Supervisor, Professor Rafik A. Goubran

Chair, Department of Systems and Computer Engineering

Carleton University

August 2004

Abstract

This thesis examines beamforming using a circular array of stethoscopes. The thesis provides background information on thoracic anatomy and heart and lung sounds as well as microphone arrays and beamforming. Following this, simulation and verification results are presented and analyzed.

In simulation, some parameters affecting beamformer performance are evaluated (signal frequency, array radius, speed of sound, number of microphones, and distance of the source from the center of the array). The performance of the beamformer is evaluated using its beam shape as well as the proposed new performance metric, the Focus Index. The thesis also suggests two methods for source localization. These methods are tested in simulation.

In addition to simulation, verification is done using two experimental models. The experimental results support the conclusions drawn from the simulations. The results demonstrate the expected relationship between beam shape, Focus Index, and localization error.

Acknowledgements

I would like to thank my supervisor, Dr. Rafik Goubran, Chair of the Department of Systems and Computer Engineering at Carleton University, for his guidance and encouragement.

OGS-ST (Ontario Graduate Scholarship-in Science and Technology), NSERC (the Natural Sciences and Engineering Research Council of Canada), and Carleton University provided financial support for this research. 3M Canada Company donated the stethoscope chest pieces used in this experiment.

Finally, a special thanks to my family and friends for their enthusiasm, patience, and support. Thanks to Matthew McKee, my editor.

Table of Contents

Abstract	iii
Acknowledgements.....	iv
Table of Contents	v
List of Tables.....	viii
List of Figures	ix
Chapter 1 Introduction.....	1
1.1 Overview	1
1.2 Problem Statement	3
1.3 Thesis Objectives	4
1.4 Thesis Contributions	5
1.5 Thesis Organization.....	7
Chapter 2 Background Review	8
2.1 Overview of Human Heart and Lung Anatomy	8
2.2 Human Heart and Lung Sounds	11
2.2.1 Heart Sounds	11
2.2.2 Lung Sounds.....	13
2.2.3 Acoustic Properties of Heart, Lungs, and Chest Cavity.....	14
2.2.4 The Stethoscope	16
2.3 Microphone Arrays	17
2.3.1 Delay-and-Sum Beamforming	18
2.3.2 Near-Field Versus Far-Field.....	20
2.4 Multiple Stethoscope Systems	22
2.4.1 Stethographics Multichannel Stethoscope	22
2.4.2 Acoustic Imaging of the Human Chest	23
2.4.3 Comparison of Sounds Recorded over Carotid Artery and Chest	26

2.4.4 Multichannel Heart Sound Recording.....	26
Chapter 3 Simulation Environment and Experimental Methods.....	28
3.1 Computer Simulation Set-Up	28
3.2 In Vitro Set-Up.....	31
3.2.1 Experimental Model I.....	33
3.2.2 Experimental Model II	34
3.3 In Vivo Set-Up	35
3.4 Chapter Summary.....	36
Chapter 4 Evaluation of Beamformer Performance – Beam Shape and Focus Index	37
4.1 Beam Shape.....	37
4.1.1 Different Representations of Beam Shape	38
4.1.2 Beam Shape for Circular Microphone Array with Source Inside	44
4.2 Focus Index	45
4.2.1 Other Beamforming Metrics	46
4.2.2 Proposed Focus Index	47
4.2.3 Calculation of Focus Index in Simulations	49
4.2.4 Empirical Focus Index Results.....	50
4.2.5 Focus Index Results for Simulations.....	52
4.3 Chapter Summary.....	57
Chapter 5 Source Localization.....	58
5.1 Far-Field Source Localization	58
5.2 Near-Field Source Localization	60
5.2.1 Localization Method A – Maximum Power in Beam Shape	61
5.2.2 Localization Method B – Solving Systems of Equations.....	62
5.3 Near-field Localization Simulations	64
5.3.1 Description of Simulations.....	64
5.3.2 Simulation Results.....	66
5.3.3 Effectiveness of Localization	73
5.4 Chapter Summary.....	75
Chapter 6 Verification Through Experiments	76
6.1 Experimental Model I.....	77
6.1.1 Description of Experiments.....	77

6.2 Experimental Model II	78
6.2.1 Foamed Gelatin	78
6.2.2 Description of Experiments.....	79
6.3 Analysis of Sound Recordings	80
6.3.1 Calibration and Speed of Sound.....	81
6.3.2 Beam Shape.....	82
6.3.3 Localization Method A – Maximum Power in Beam Shape	82
6.3.4 Localization Method B - Solving Systems of Equations	82
6.3.5 Focus Index	83
6.4 Experimental Results.....	83
6.5 Discussion of Experimental Results.....	100
6.5.1 Comparison of Experimental Model I and II	100
6.5.2 Effect of Source Location	102
6.5.3 Effect of Different Sound Inputs.....	102
6.5.4 Comparison of Localization Methods A and B.....	103
6.5.5 Validity of Focus Index.....	105
6.5.6 Relationship Between Focus Index and Localization Error.....	106
6.5.7 Sources of Error During In Vitro Trials	106
6.6 In Vivo Experiments	108
6.7 Chapter Summary.....	108
Chapter 7 Conclusions and Future Work.....	110
7.1 Challenges in Human Application	110
7.2 Overall Conclusions	111
7.3 Future Work	112
References	114

List of Tables

Table 4.1 Focus Indices Resulting from Selected P_i	51
Table 5.1: Summary of Simulation Results	67
Table 5.2 Legend for Figures 5.6 to 5.11	67
Table 6.1: Summary of Experimental Results	84
Table 6.2 Legend for Figures 6.3 to 6.16	85

List of Figures

Figure 2.1 Anatomy of the human heart	9
Figure 2.2 Anatomy of human lungs.....	10
Figure 2.3 Phonocardiogram of normal heart sounds	12
Figure 2.4 Phonocardiogram of pansystolic murmur.....	12
Figure 2.5 Delay-and-sum beamformer	19
Figure 2.6 Far-field and near-field wave fronts	21
Figure 3.1 Screen shot of a GUI showing the beam shape	30
Figure 3.2 Stethoscope chest piece-microphone set-up	32
Figure 3.3 Microphone pre-amps connected to M-Audio Delta 1010	32
Figure 3.4 Experimental Model I with stethoscope chest pieces mounted, being inserted into sound-proof box	33
Figure 3.5 Experimental Model II with stethoscope chest pieces mounted on foamed gelatin	35
Figure 3.6 Stethoscope placement for in vivo experiments	36
Figure 4.1 Beam shape for linear microphone array (3 microphones, 0.03m apart); focus located 0.5 m from microphone array, input is 4.4 kHz sinusoid; power measured at indicated angles along half-circle of radius=0.05m	40
Figure 4.2 Beam shape for linear microphone array (3 microphones, 0.03m apart); focus located 0.5 m from microphone array; input is sinusoid of indicated signal frequencies (2, 2.8, 3.6, 4.4, 5.2, 6, 6.8, 7.6 kHz); power measured at indicated angles along half-circle of radius=0.05m.....	41
Figure 4.3 Beam shape for linear microphone array placed along horizontal axis (3 microphones, 0.03m apart); focus located 0.5 m from microphone array; input is 4.4 kHz sinusoid; power measured at indicated angles along full circle of radius=0.05m	42
Figure 4.4 Beam shape for linear microphone array (5 microphones, 0.08m apart) indicated by blue squares); focus designated by red 'x' and located 1.5 m from microphone array; input is 2kHz sinusoid.....	43

Figure 4.5 Beam shape for circular microphone array with source outside array (5 microphones indicated by blue squares, array radius=0.1 m); focus designated by red 'x' and located 1.5 m from microphone array; input is 2kHz sinusoid.....	43
Figure 4.6 Screen shots from GUI showing beam shapes: a) shows position of focus at centre (red) and 8 microphones encircling it (blue), b)- f) show beam shapes with radius=0.15 m, speed of sound= 300 m/s and signal frequencies of b) 300 Hz, c) 700 Hz, d) 1100 Hz, e) 1500 Hz, f) 1900 Hz. White areas have higher power, black areas have lower power.	45
Figure 4.7 Diagram of concentric circles where average power is calculated relative to the power at the ideal focus, indicated by a red "x".....	49
Figure 4.8 Sector (in gray) which repeats eight times in a four microphone circular array if the focus is at the center of the array	50
Figure 4.9 Focus index versus signal frequency, where radius=0.15 m, speed of sound=340 m/s, and number of microphones=8	54
Figure 4.10 Focus index versus array radius, where signal frequency=1000 Hz, speed of sound=340 m/s, and number of microphones=8	54
Figure 4.11 Focus index versus speed of sound, where signal frequency=1000 Hz, radius=0.15 m, and number of microphones=8	55
Figure 4.12 Focus index versus number of microphones, where signal frequency=1000 Hz, radius=0.15 m, and speed of sound=340 m/s	55
Figure 4.13 Diagram showing location "a", at the center of the array and focus location "b", 0.05 m from it (as shown in Figure 4.14)	56
Figure 4.14 Focus index versus distance between focus and center of array, where number of microphones=8, signal frequency=1000 Hz, radius=0.15 m, and speed of sound=340 m/s (Figure 4.13 shows location of "a" and "b")	56
Figure 5.1 Far-field source localization	59
Figure 5.2 Far-field localization results with two microphones and three microphones ..	60
Figure 5.3 Near-field source localization, the red "x" indicates the sound source	61
Figure 5.4 Near-field source Localization Method B, the red "s" indicates the sound source	63
Figure 5.5 Layout of microphones and positions of center of sources (indicated by a red "x") in simulation trials	66
Figure 5.6 Beam shape and localization results for 1 kHz tone source located at center in simulation (the red X disappears because it is directly under the green X)	68
Figure 5.7 Beam shape and localization results for heart sounds source located at center in simulation (the red X disappears because it is directly under the green X)	69

Figure 5.8 Beam shape and localization results for 1 kHz tone source located near mic 1 in simulation.....	70
Figure 5.9 Beam shape and localization results for heart sounds source located near mic 1 in simulation (the red X is partially covered by the green X).....	71
Figure 5.10 Beam shape and localization results for 1 kHz tone source located between mics 3 and 4 in simulation.....	72
Figure 5.11 Beam shape and localization results for heart sounds source located between mics 3 and 4 in simulation (the red X is partially covered by the green X).....	73
Figure 6.1 Layout of microphones and positions of center of sources (indicated by a red “x”) in Experimental Model I.....	78
Figure 6.2 Layout of microphones and positions of center of sources (indicated by a red “x”) in Experimental Model II trials	80
Figure 6.3 Beam shape and localization results for 1 kHz tone source located at center of Experimental Model I.....	86
Figure 6.4 Beam shape and localization results for 100 Hz tone source located at center of Experimental Model I.....	87
Figure 6.5 Beam shape and localization results for heart sounds source located at center of Experimental Model I	88
Figure 6.6 Beam shape and localization results for white noise source located at center of Experimental Model I (the red X disappears because it is directly under the green X).....	89
Figure 6.7 Beam shape and localization results for 1 kHz tone source located near mic 1 of Experimental Model I	90
Figure 6.8 Beam shape and localization results for heart sounds source located near mic 1 of Experimental Model I	91
Figure 6.9 Beam shape and localization results for 1 kHz tone source located between mics 5 and 6 of Experimental Model I.....	92
Figure 6.10 Beam shape and localization results for heart sounds source located between mics 5 and 6 of Experimental Model I.....	93
Figure 6.11 Beam shape and localization results for 1 kHz tone source located at center of Experimental Model II	94
Figure 6.12 Beam shape and localization results for heart sounds source located at center of Experimental Model II.....	95
Figure 6.13 Beam shape and localization results for 1 kHz tone source located near mic 1 of Experimental Model II.....	96
Figure 6.14 Beam shape and localization results for heart sounds source located near mic 1 of Experimental Model II.....	97

Figure 6.15 Beam shape and localization results for 1 kHz tone source located between mics 3 and 4 of Experimental Model II.....	98
Figure 6.16 Beam shape and localization results for heart sounds source located between mics 3 and 4 of Experimental Model II.....	99

Chapter 1 Introduction

1.1 Overview

From its invention by Laennec in the early 1800's¹ until the present time, the stethoscope has been an invaluable tool in the medical profession for patient diagnosis and assessment. This primitive instrument provides the physician with acoustic information which tells of the structure, the physiology, and the pathology under the skin. The traditional stethoscope used in practice today has changed relatively little in the past 200 years when compared to other technological and medical advances that have taken place in the same period of time. Although the stethoscope is a useful tool for non-invasive preliminary diagnosis, it is very limited in its capabilities. It is only able to detect sounds on the surface of the body and diagnoses are highly subjective and can vary greatly between physicians^{2,3}.

By combining the stethoscope with the accessible, compact, powerful, and reasonably priced computers it is possible to achieve something better. Signal processing techniques can be used to amplify the signal and to reduce noise. Simple programs can display the acoustic signal visually which can provide physicians with a visual profile of pathology. In the future, the computer could analyze the sounds and suggest a diagnosis

based on a database of previous results. The acoustic data can be saved on the computer as an electronic record for the patient. This data could then be sent off to other physicians for assistance in diagnosis or referred to later to determine if there is a change in the status of the patient⁴.

In addition to combining the stethoscope with the computer, another useful advancement would be the simultaneous measurement of acoustic data from more than one location. A multisensor system, consisting of several stethoscope chest pieces, could be used to simultaneously record many channels of audio data and to store them for computer-aided analysis. This will provide physicians with more information for diagnosis compared to the traditional stethoscope. The physician will be able to listen to each channel separately, varying amplitude and play-back speed as many times as necessary. Slowing down the playback can be helpful when listening to chest sounds of infants who have faster heart beats and respiration. Additionally, the physician can change stethoscope listening positions while listening to the same breath or heart beat, enabling him or her to hear the same sound at the same point in time, but from a different position. The physician can also amplify certain portions of the signal which may contain valuable diagnostic information but would often be missed with the traditional stethoscope^{3, 4}.

In addition to sequential playback, signal processing algorithms, such as beamforming, can be used to combine the multi-sensor inputs. Beamforming will amplify sounds coming from a particular region inside the chest and attenuate all other regions. Beamforming could allow physicians to hear signals below the surface of the skin by focusing the beam in a particular location. Third, beamforming can also be used for source localization; that is, determining from where a particular sound originated. The information obtained by combining multiple signals can provide a non-invasive, inexpensive acoustic image of the chest cavity for diagnostic and therapeutic purposes^{3,4}.

1.2 Problem Statement

This thesis describes the issues involved in designing a multisensor stethoscope for clinical use. The goal of the multisensor stethoscope is to improve the sound quality and to provide additional diagnostic information compared to that which is provided by the contemporary stethoscope. In aid of this pursuit, this thesis analyses the properties of circular microphone array beamformers where the source is placed inside the array. This investigation begins with simulations in free-field followed by in vitro experiments to bridge the gap from free-field to reality, the human body. This thesis addresses localization and signal-combining and identifies issues in applying the multimicrophone system to human testing.

1.3 Thesis Objectives

The purpose of this thesis is to study the localization of a sound source using multiple microphones and to evaluate the potential for heart and lung sound localization. Sounds are localized by spatially directing the beam towards a source located inside an array of microphones. The objective is to determine the factors affecting the quality of localization and beamforming. To accomplish this objective, the following steps are taken: First, a MATLAB model simulates the situation, then, some physical models are constructed to verify the simulation results. A graphical user interface (GUI) is created in Matlab to display the simulation and experimental results.

The main emphasis of this thesis is the area of beamforming and localization. It does not deal with the design of acoustic sensors, the detection of heart and lung disease, or the analysis of chest sounds for the purpose of diagnosis.

1.4 Thesis Contributions

The following are the major contributions demonstrated in this thesis:

1. A new beamformer performance metric, the Focus Index, was proposed and tested. This contribution was described in a paper entitled “Beam Shape, Focus Index, and Localization Error for Performance Evaluation of a Multisensor Stethoscope Beamformer” that was accepted for presentation at the *International Conference of the IEEE Engineering in Medicine and Biology Society* held in San Fransisco, California in September of 2004³.
2. Beam shape and Focus Index were used to evaluate the parameters that affect beamformer performance, i.e. the signal frequency, array radius, speed of sound, number of microphones, and distance of the source from the center of the array. This contribution was also addressed in the above-mentioned paper presented in the *International Conference of the IEEE Engineering in Medicine and Biology Society*³.

3. Two near-field source localization algorithms were designed and compared: Localization Method A, which is based on the maximum power in the beam shape and Localization Method B, which is based on differences in arrival time and involves solving systems of equations. This contribution was described in a paper entitled "Multisensor Stethoscope for Chest Sound Localization" that was accepted for presentation at the *Canadian Medical and Biological Engineering Society Conference* held in Quebec City, Quebec in September of 2004⁴.
4. Two experimental models were built to verify simulation results. Experimental Model I consisted of a hollow plastic cylinder lined with foam. Experimental Model II was a solid cylinder of foamed gelatin which had a speed of sound similar to that of the human body. This work will be submitted to the *2005 IEEE Instrumentation and Measurement Technology Conference*.
5. A flexible GUI (graphical user interface) was constructed to illustrate the beam shape under a variety of acoustic conditions. This GUI was modified to perform source localization and calculate localization error and Focus Index.

1.5 Thesis Organization

Chapter 1 introduces the problem addressed in this thesis.

Chapter 2 provides background information about heart and lung anatomy and sounds as well as some detail on stethoscopes, microphone arrays, and beamforming.

Chapter 3 describes the simulation environment, experimental set-up, and evaluation methods.

Chapter 4 defines the proposed Focus Index and shows how it relates to beam shape under various conditions in a simulation environment.

Chapter 5 describes two methods of source localization and some related simulation results.

Chapter 6 demonstrates the results of the in vitro verification using two experimental models. Focus Index, beam shape, and localization error are addressed.

Chapter 7 concludes the thesis, summarizes the main findings, and suggests future work.

Chapter 2 Background Review

In this chapter, some background information about heart and lung anatomy and sounds as well as some detail on stethoscopes, microphone arrays, and beamforming is provided.

2.1 Overview of Human Heart and Lung Anatomy

The heart is a muscular pump that enables the blood to circulate throughout the body⁵. The heart consists of four chambers: the left and right atria and ventricles (as shown in Figure 2.1). The chambers are separated by valves which control the movement of blood through the heart by opening to allow flow into the next chamber or closing to prevent backwards flow⁶. The tricuspid valve separates the right atrium and right ventricle. The bicuspid or mitral valve separates the atrium and ventricle on the left. The pulmonary valve is between the right ventricle and the pulmonary trunk which leads out to the lungs. The aortic valve is between the left ventricle and the aorta which delivers blood to the body. Blood flows from the body, through the vena cava, into the right side of the heart where it is pumped to the lungs for gas exchange. The blood is then returned from the lungs to the left side of the heart which pumps blood to the rest of the body^{7,8}.

The cardiac cycle is composed of systole and diastole. In the systolic phase the heart contracts and in the diastolic phase the heart relaxes and fills with blood^{9, 10}.

Figure 2.1 Anatomy of the human heart¹¹

The respiratory tract consists of the vocal tract (trachea) and the subglottal airways¹. The trachea divides into two main branches (bronchi) which lead to smaller and smaller branches into the lungs which are composed of tiny sacs (alveoli) through

which the gas exchange into the blood stream occurs. The lungs are divided into several lobes. On the right there are three lobes (upper or superior, middle, and lower or inferior), and on the left there are only two lobes (upper and lower) in order to make room for the heart^{7, 8} (as shown in Figure 2.2).

Figure 2.2 Anatomy of human lungs¹²

2.2 Human Heart and Lung Sounds

2.2.1 Heart Sounds

The normal heart sounds are called S1 (first heart sound) and S2 (second heart sound). S1-S2 is often described as “lub-dub”⁹ to mimic the sound heard on auscultation (listening through a stethoscope). S1 is caused by the closing of the tricuspid and mitral valves whereas S2 is caused by the closing of the aortic and pulmonary valves^{9, 6}. Sometimes, a third and fourth heart sound (S3 and S4) can be heard at lower energy⁵.

Murmurs are created by a variety of medical conditions or can be normal in certain conditions. One possible cause of murmurs is damage to the valve leaflets⁵. If a valve does not open completely (stenosis), the blood flow will be more turbulent. On the other hand, if a valve does not completely close (valvular insufficiency) some blood may flow in the wrong direction^{6, 10}.

The phonocardiogram is a plot of the heart sounds in terms of amplitude versus time¹³. Figures 2.3 and 2.4 are examples of phonocardiograms. Figure 2.3 shows normal heart sounds, where the peaks are S1 and S2. Figure 2.4 shows a phonocardiogram of a pansystolic murmur which means that a murmur is heard during the entire duration of

time between S1 and S2. These plots are partial screen shots from a heart sounds GUI that was created during the course of the research.

Figure 2.3 Phonocardiogram of normal heart sounds

Figure 2.4 Phonocardiogram of pansystolic murmur

2.2.2 Lung Sounds

The breathing sound heard on auscultation of the chest of a healthy person is called normal lung sounds. These sounds result from air turbulence. Tracheal sounds (sometimes distinguished from lung sounds) are heard at the suprasternal notch or on the neck¹⁴. They are louder and cover a wider range of frequencies than lung sounds heard at the chest wall¹. Over the chest, breath sounds are more muffled, especially the higher frequencies¹⁴. Lung sounds heard on the chest surface are filtered versions of those that can be heard on the neck¹. Diminished lung sound intensity is correlated with poor ventilation. In a healthy person, there is relatively little bilateral asymmetry of lung sounds and asymmetry indicates disease¹⁰. Examples of adventitious lung sounds are: wheeze, rhonchus, and crackle (also called rale)¹.

The wheeze is a continuous musical sound generated in the lower airways and is due to airway obstruction. Its quality is used to indicate the severity of asthma, for instance. The rhonchus is coarse rattling sound somewhat like snoring. It is generated in the larger airways and is due to excess fluid. The crackle is a discontinuous sound, like small explosions, heard most often during inspiration in the lower lungs. It is caused by airway closure and secretions found, for instance, in fibrotic lung disease and pneumonia¹.

2.2.3 Acoustic Properties of Heart, Lungs, and Chest Cavity

The chest consists of many components with substantially different acoustic properties. The lungs, air, blood, muscles, bones, and connective tissues all propagate sound in the chest. The chest wall and the heart transmit sound at about 1500 m/s, with little damping. The large airways transmit at 222 to 312 m/s. The lung parenchyma transmits at between 23 and 60 m/s¹⁵. Speed of sound in the chest varies with inspiration and expiration.

The mixture of tissues in the human chest results in a relatively low speed of sound, on the order of 50 m/s¹. A number of research groups reported results between 23 and 70 m/s¹⁶. This is a low speed of sound given that the propagation speeds in air or tissue alone is much higher. Compare this also to transmission speeds of 340 m/s in air and 1500 m/s in water¹⁷.

Since the chest is considered as a whole and not by its individual components, the speed of sound in the chest will be assumed as around 40 m/s following the results of Kompis *et al.*¹⁵. This low speed of sound means that the wavelength will also be relatively short for any given frequency. The short wavelength means that more information concerning internal structure can be obtained from low-frequency measurements¹. The absorption of sound energy in the chest is highly frequency dependent. Sound absorption is high at higher frequencies where wavelength approaches

alveolar size¹. Additionally, speed of sound in the chest is also frequency dependent. At low frequencies, sounds travel more slowly. In pathology amplitude can change as well as speed of sound.

To measure the response of the chest cavity to acoustic perturbations, many researchers have introduced sounds at the mouth and taken measurements on the chest surface^{1, 16}. Transit time of sounds has been measured using cross-correlation and spectral techniques to estimate phase delay. This is complicated by resonance, internal reflections, and multiple propagation pathways¹. The chest wall is also susceptible to surface waves¹. It was determined that transit times decrease as lung volumes increase¹⁶. It was also shown that phase delays exhibit bilateral asymmetry with faster transit times on the left side, which means that the heart transmits sounds faster than lung parenchyma¹⁶.

Normal heart and lung sounds are in the ranges of 20 to 150 Hz and 40 to 600 Hz, respectively. These low frequencies pose a challenge for delay-and-sum beamformers which focus better at higher frequencies. Adventitious sounds are present at frequencies up to 1200 Hz for the heart and 2000 Hz for the lungs^{18, 19, 20}. Damping also increases with frequency¹⁵.

2.2.4 The Stethoscope

The stethoscope is the most widely used tool for auscultation⁹. Auscultation of the heart and lungs is a low-cost and nonintrusive method for the diagnosis of a variety of pathologies. Alterations in the structure and function of thoracic organs caused by disease often result in changes in sound production and transmission²¹. Medical professionals are trained to recognize and distinguish between normal and abnormal sounds.

Stethoscopes modify the sounds coming from the chest and therefore do not provide a faithful representation of the sound. Stethoscopes amplify sounds with a frequency lower than 112 Hz and attenuate high frequency sounds¹. Incidentally, this is advantageous in a clinical sense for cardiologists since lower frequencies are poorly perceived by the human ear and heart sounds occur in this range¹. However, adventitious lung sounds occur in the higher frequencies and some of their details may be masked by the contemporary stethoscope.

The usefulness of the stethoscope depends on the user's hearing and interpretative ability⁹. It would be helpful to create medical diagnostic aids since this interpretation is subjective⁵. A number of research groups are developing expert systems to capture the knowledge of an expert and automate his or her problem solving process in the form of a rules-based system^{5,6}.

2.3 Microphone Arrays

An array is a group of sensors at distinct spatial locations used to measure a propagating field. Information about the propagating signal is obtained by combining sensor inputs²⁴. Microphone arrays are used in a number of applications such as teleconferencing, speech recognition, speech enhancement, and hearing aids. The goals of array processing can be summarized as follows²⁴:

- 1) To enhance the signal to noise ratio (SNR) beyond that of a unique sensor (if the noise is uncorrelated).
- 2) To characterize the field (determining number of sources, locations, and waveforms).
- 3) To track sources as they move (electronic steering rather than manual).

By using multiple microphones, sounds can be sampled spatially and the direction of arrival can be used to discriminate the signal of interest from interfering signals²². Beamforming is an algorithm applied to the signals from multiple microphones whose goal is to amplify sounds coming from a selected location and attenuate sounds coming from all other locations. It is a method of spatial filtering²⁴.

Beamforming can be classified as either fixed or adaptive. In fixed beamforming the locations of the source and the interfering sounds are known and assumed to be constant. On the other hand, in adaptive beamforming, the locations of the source and interfering sounds are not known and may be changing. The adaptive algorithm depends on the characteristics of the data that it receives²⁴. This thesis focuses entirely on fixed beamforming, namely a method called “delay-and-sum”.

2.3.1 Delay-and-Sum Beamforming

In delay-and-sum beamforming, input signals of a microphone array are delayed to set a specified focus location and then summed (as shown in Figure 2.5 below). Relative delays between sensor signals are set so that the waves propagating in a particular direction and at a particular speed add constructively; noise and all other waves add destructively²⁴. In free-field, the delay is based on time of arrival, whereas in a more realistic situation of a nonhomogeneous medium or with reflections, the delay would be the conjugate of the transfer function between the source and the microphones²⁴. Beamforming algorithms vary according to whether the sources are located in the near-field or the far-field²⁴.

$$w_i(t) = A_i \sin \left[2\pi f \left(t + \frac{dist_i}{v_{sound}} + delay_i \right) \right] \quad (2.2)$$

The $delay_i$ is the delay required to focus the beam on a particular location. The output of the beamformer, $q(t)$, is found by adding all the $w_i(t)$'s.

$$q(t) = \sum_{i=1}^N w_i(t) \quad (2.3)$$

2.3.2 Near-Field Versus Far-Field

In near-field, the source is located close to the microphones and the wave front appears curved with respect to the dimensions of the array²³. In far-field, the source is so far from the microphones that the wave front appears planar and the direction of propagation is approximately equal at each sensor^{24,25}. Figure 2.6 below illustrates the wave fronts. The distinction between near-field and far-field depends on the level of acceptable error²⁴. In near-field beamforming, the summation yields a sound signal focused on a particular point in the field, compared to far-field which can focus on a particular direction but not depth²⁴.

Figure 2.6 Far-field and near-field wave fronts

2.4 Multiple Stethoscope Systems

This section examines several research projects found in the literature that involve multiple stethoscope systems.

2.4.1 Stethographics Multichannel Stethoscope

An American company called Stethographics has created a device called the Multichannel STG for listening to lung sounds²⁶. The patient lies on his or her back on a foam pad with 16 embedded microphones. A computer monitor displays all the recordings in a time-expanded waveform for visual inspection and allows playback. The channels can be visually examined to detect abnormalities and irregular patterns. A PC analyses acoustic energy versus time and detects crackles, wheezes, and rhonchi. It provides the user with information such as the amplitude of breath sounds and the timing and spatial location of abnormalities. The Stethographics system also enables rudimentary three-dimensional localization and modeling of intrathoracic sounds. The generated image can be rotated and sectional views can be extracted and compared. The Multichannel STG is mainly used for research and educational purposes at this point²⁶.

The Multichannel STG was used by Berstesser *et al.* to examine sound transmission in the lungs as a function of lung volume¹⁶. A sound was introduced at the

mouth and recordings on the thoracic surface were measured by the Stethographics device. Cross-correlation was used to calculate transthoracic sound transit times. Recordings were performed on twelve healthy subjects¹⁶.

Transit times varied from 1 to 5 ms depending on microphone location. Transit times were greatest in the lower part of the lungs. In addition, it was found that the transit time at all locations decreased with increasing lung volume and decreasing input sound frequency. At all lung volumes, the transit time on the left was shorter than the right. This is thought to be due to sounds traveling faster through the heart than the lungs¹⁶.

Berstresser *et al.* developed a model to describe sound transmission in the chest. Transit times were used to estimate the speed of sound. Their model was comprised of sound propagation in free-field through the trachea, then through a two-phased system in the parenchyma (in air and in tissue). The speed of sound this research group estimated for sound transmission in the chest was 37 m/s¹⁶.

2.4.2 Acoustic Imaging of the Human Chest

According to Kompis *et al.*, “simultaneous multi-microphone recordings of thoracic sounds on the chest wall contain spatial information of potential diagnostic value”²¹. This research group developed an algorithm to test the hypothesis that, for a given point inside the chest, it contained the only relevant acoustic source. A hypothetical

source signal was calculated by least-squares estimation to explain a maximum of the signal variance in all microphones. The sound propagation was assumed to be uniform with a constant speed of sound and damping factor per unit length. The thorax was scanned using a 1 cm step size. For each location in the thorax, the proportion of all microphone signals which could be explained by the hypothetical source only was given a number between 0 and 1. A three dimensional table represented the thoracic volume. Three-dimensional gray-scale images were used to represent the spatial positions of the sources identified by the algorithm^{15, 21}. Trials were done in simulation, in vitro and in vivo.

In simulation, the microphone signals were simulated by a computer program. Sixteen microphones were arranged in the same positions as were used in the human trials: two rows of four microphones on the front and on the back. The speed of sound was assumed to be 40 m/s. Simulations were done with two point sources emitting white noise. Trials were done with only one point source active as well as with both active at the same time. Three-dimensional gray-scale images demonstrated the expected spatial positions of the sources in all cases¹⁵.

The in vitro trials were performed on an unpublished gelatin model of the human thorax. The model consisted of a chest wall of unfoamed gelatin, foamed gelatin interior, and a rubber tube in the center to represent the trachea. Inside the rubber tube was placed

an acoustic compression driver to introduce white noise. Recordings were made with 4 cm spacing over the entire surface of the model. This was equivalent to 105 microphone recordings. Acoustic images were constructed assuming different speeds of sound in various frequency bands. Broadband images showed the correct model trachea length and placement to within 2 cm. The researchers hypothesized from their results in various frequency bands that low frequency sounds exit the model trachea closer to the sound source whereas higher frequency sounds travel further¹⁵.

Their human trials consisted of using sixteen microphones arranged in two rows of four microphones on the front and back of the thorax. Their subjects were four healthy males and a child with pneumonia. Microphone signals were amplified and bandpass filtered between 100 and 1000 Hz. The subjects breathed through a pneumotachograph (an apparatus for recording the rate of airflow to and from the lungs) to maintain constant breathing patterns. For heart sound imaging the subjects held their breath. The results were displayed in a three-dimensional gray-scale image. Patterns for inspiration, expiration and first and second heart sounds were identified²¹. Inspiratory sounds were produced predominantly in the periphery of the lungs whereas expiratory sounds originated more centrally. Heart sounds were seen to be originating from the lower left thorax. In the subject with pneumonia the area of consolidation evident in x-rays images corresponded well with an area with high data values in the acoustic images¹⁵. According

to Kompis *et al.*, the spatial resolution of an acoustic imaging system cannot resolve differences below about 2 cm¹⁵.

2.4.3 Comparison of Sounds Recorded over Carotid Artery and Chest

Another research group compared the phonocardiograms and spectral content of heart sounds recorded from the chest to those recorded over the carotid artery¹³. Hui *et al.* designed a system that simultaneously recorded heart sounds from two different sites. Recordings were performed on a healthy female. The variations in intensity and frequency content were analyzed to determine if these factors were location-dependent. It was shown that the relative intensities of S1 and S2 vary with location of recording. Recordings from the chest showed that the amplitude of S1 was greater than that of S2 and the reverse was true of recordings from the carotid artery. The frequency content of S1 and S2 were computed. The frequencies corresponding to peak intensities did not vary significantly between recording locations. Spectral power did, however, vary more significantly for S2. Hui *et al.* showed that intensity and frequency content depend on location of recording and also that variations are greater in the second heart sound¹³.

2.4.4 Multichannel Heart Sound Recording

Research by Haghighi-Mood *et al.* addressed the need for a multichannel approach for spectral analysis of heart sounds. An important aspect of auscultation is finding the appropriate location on the chest wall where different components of the heart

sound can be heard clearly²⁷. This suggests a multichannel approach may be useful but according to Haghighi-Mood *et al.*, this may be a limitation of the human auditory system. This study investigated the similarity of cardiac sounds across channels²⁷.

Sounds were recorded simultaneously from four auscultatory areas – those best for listening to the aortic, pulmonary, mitral, and tricuspid valves. The recordings were partitioned into first and second heart sounds. Then, a method developed using cross-spectral and coherence analysis in the frequency domain was applied to compare the spectrum of recordings from different areas on the chest. A high degree of coherence among dominant frequency components was found in all channels²⁷. These findings indicated that the spectral details of the heart sounds can be extracted from any single channel. If only the spectral details are of interest, this knowledge could be used to reduce the complexity of data acquisition and decision making systems²⁷.

Chapter 3 Simulation Environment and Experimental Methods

This chapter describes the simulation environment and experimental set-up.

The simulations and the experiments consider only a two-dimensional circular microphone array. All of the microphones are arranged in one plane and are evenly spaced with the source located inside the array. This method of sound capture was also used by Tamai *et al.*, but for the purposes of a meeting system²⁸. Previous research by Kompis *et al.* used 16 microphones arranged in 2 rows of 4 on the front and back of the chest for a three-dimensional view¹⁵.

3.1 Computer Simulation Set-Up

The computer simulations are conducted using Matlab Version 6.5. In the simulations, the sound source is either a sine wave or a recording of normal heart sounds. Near-field free-field propagation with a non-directional point source is assumed.

Several GUIs were created in Matlab to display the simulation results and to allow the user to easily modify the parameters. One GUI allows the user to select the number of microphones, the radius of the microphone array, the source location, the signal frequency, the sampling frequency, and the speed of sound. Matlab then determines the delays required for the beamformer, and displays the beam shape (as shown in Figure 3.1). Another GUI (a modified version similar to Figure 3.1) allows the user to enter the microphone locations, the source location, the signal frequency, the sampling frequency, and the speed of sound. Matlab then returns a figure illustrating the microphone positions, the true source location, the estimated source location, and it calculates the localization error. Further explanations about the simulations and their results are provided in Chapter 4 and 5.

Figure 3.1 Screen shot of a GUI showing the beam shape

3.2 In Vitro Set-Up

Two models were built to verify the simulation results; these are referred to as Experimental Model I and Experimental Model II. These models are further described in sections 3.2.1 and 3.2.2 as well as in Chapter 6.

The models are placed in sound-proofed boxes and recordings are done in a relatively quiet room. Sounds are picked up using eight equally-spaced Littmann Master Classic II chest pieces attached in one plane along the outside of the cylindrical models. These chest pieces are attached to ATM 803-b Audio-Technica lavalier microphones using a short piece of rubber tubing (as shown in Figure 3.3). The microphones are then connected to an M-Audio Delta 1010, an 8-channel synchronous sound amplifier and digitizer (as shown in Figure 3.4). The microphone levels are amplified by a proprietary 8-channel pre-amplifier, based on the Texas Instruments INA171 low-noise instrumentation amplifier. The data is fed into a Dell Dimension 4600 PC. Data is collected by Adobe Audition using a 16 kHz sampling frequency.

Figure 3.2 Stethoscope chest piece-microphone set-up

Figure 3.3 Microphone pre-amps connected to M-Audio Delta 1010

3.2.1 Experimental Model I

A sound source is placed in a plastic container filled with air and lined with foam to reduce reflections. The stethoscope chest pieces are adhered to the outer wall of the container using ultrasound gel to create a better seal (as shown in Figure 3.5). The sound source is a 2" speaker transducer, baffled in order to create a more omni-directional radiation pattern. Baffling is done with bubble wrap and tape. The speaker is hung inside the container so that it faces the bottom of the container and in the same plane as all the microphones outside the container.

Figure 3.4 Experimental Model I with stethoscope chest pieces mounted, being inserted into sound-proof box

3.2.2 Experimental Model II

The purpose of this model is to lower the speed of sound to a value closer to that in the actual human torso. This model is based on that used in previous experiments by Kompis *et al.* which consisted of a cylindrical solid gelatin wall filled with foamed gelatin in the center of which was a rubber tube to represent the trachea¹⁵. Experimental Model II uses foamed gelatin (as shown in Figure 3.6) alone in order to maintain a homogeneous transmission medium to perform testing without adding in contradictory elements. Foamed gelatin is used because its speed of sound is in the range of that of the human torso, approximately 40 m/s¹⁵. The sound source is a SCG-16A 16 mm speaker from Star Micronics, baffled in the same way as described in section 3.2.1. The speaker is placed into the gelatin in the same plane as the microphones.

Figure 3.5 Experimental Model II with stethoscope chest pieces mounted on foamed gelatin

3.3 In Vivo Set-Up

The in vivo trials are conducted on a human subject. The sound collection set-up is the same as is described in section 3.2 less the sound-proof box. Eight stethoscope chest pieces are attached to the skin, evenly spaced in one plane (as shown in Figure 3.6). The trials are done with the subject breathing normally as well as holding its breath.

Figure 3.6 Stethoscope placement for in vivo experiments

3.4 Chapter Summary

This chapter describes the simulation environment as well as the experimental procedure. In all cases, the source is in the same plane and encircled by the microphone array. Further details about simulation methods as well as the results follow in Chapters 4 and 5, and experimental analysis and results are in Chapter 6.

Chapter 4 Evaluation of Beamformer Performance – Beam Shape and Focus Index

In this chapter, the proposed Focus Index is defined. The relationship between beam shape and Focus Index is examined under various conditions in a simulation environment. The beamformer performance is evaluated by modifying the following parameters: signal frequency, number of microphones, radius of microphone array, and speed of sound. Some of the work presented in this chapter was published in the *Proceedings of the International Conference of the IEEE Engineering in Medicine and Biology Society* in September of 2004⁴.

4.1 Beam Shape

One way in which a beamformer can be evaluated is by observing the beam shape. The beam shape (or beam pattern) usually represents the power in other parts of the field relative to the desired beam focus location²⁹. There are many different representations of beam shape, but, in general, the beam shape is determined based on the positions of the microphones, the desired focus location, and the speed of sound. The

beamformer delays are designed to align the received microphone signals to achieve the maximum power at the desired focus location. In delay-and-sum beamforming, the beamformer delays are derived from the differences in expected time of arrival of the signal at different microphones. The relative power of signals originating from all relevant locations in the field is calculated using the same beamformer delays (in addition to the transit time delays) in order to determine how much sound coming from another location would be present in the output signal⁴.

4.1.1 Different Representations of Beam Shape

Beam shape is represented in many different ways, some of which are presented here. Much research has been done on beam shapes for linear arrays. Figures 4.1 to 4.4 show beam shapes for linear microphone arrays. Figure 4.5 illustrates the beam shape for a source placed outside of a circular microphone array. Figure 4.6 illustrates the beam shape for a source placed inside of a circular microphone array (these results are addressed in section 4.1.2).

Figures 4.1 and 4.2 were generated in Matlab based on reference ²⁹ and they illustrate the relative power obtained at various angles along a half-circle of set radius around a linear array. Figure 4.1 is a two-dimensional slice showing only one of the frequencies of Figure 4.2. Figure 4.3 illustrates the same information as in Figure 4.1 using a different way to illustrate the beam shape. In Figure 4.3, the distance from the

origin represents the relative power at that angle. Figures 4.4 and 4.5 are screen shots from a GUI created in Matlab in order to easily modify parameters and display the relative power at different locations in the field on a gray scale (white indicates areas of maximum power and black indicates minimum power). A figure similar to Figure 4.1 can be obtained from Figure 4.4 by taking power measurements along a half-circle of set radius around the microphone array.

In Figures 4.1 to 4.5, the ideal focus is chosen to be at 90 degrees to the microphone array axis, thus the maximum power occurs at that angle in all the figures. The power at all other points in the field is attenuated relative to the power at the ideal focus. Generally, attenuation increases with the distance from the ideal focus. Spatial aliasing is defined as areas of high intensity away from the ideal focus. This usually occurs when the microphone spacing is greater than half the incoming signal's wavelength²⁹. Spatial aliasing can be seen in Figures 4.1 to 4.5 near 0 and 180 degrees, and in Figure 4.4 near 45 and 135 degrees. Figure 4.2 also illustrates the relationship between frequency and beam shape. For low signal frequencies the main lobe is broad and there is no spatial aliasing. As the signal frequency increases, the main lobe becomes narrower and spatial aliasing appears.

Figure 4.1 Beam shape for linear microphone array (3 microphones, 0.03m apart); focus located 0.5 m from microphone array, input is 4.4 kHz sinusoid; power measured at indicated angles along half-circle of radius=0.05m²⁹

Figure 4.2 Beam shape for linear microphone array (3 microphones, 0.03m apart); focus located 0.5 m from microphone array; input is sinusoid of indicated signal frequencies (2, 2.8, 3.6, 4.4, 5.2, 6, 6.8, 7.6 kHz); power measured at indicated angles along half-circle of radius=0.05m²⁹

Figure 4.3 Beam shape for linear microphone array placed along horizontal axis (3 microphones, 0.03m apart); focus located 0.5 m from microphone array; input is 4.4 kHz sinusoid; power measured at indicated angles along full circle of radius=0.05m

Figure 4.4 Beam shape for linear microphone array (5 microphones, 0.08m apart) indicated by blue squares); focus designated by red 'x' and located 1.5 m from microphone array; input is 2kHz sinusoid

Figure 4.5 Beam shape for circular microphone array with source outside array (5 microphones indicated by blue squares, array radius=0.1 m); focus designated by red 'x' and located 1.5 m from microphone array; input is 2kHz sinusoid

4.1.2 Beam Shape for Circular Microphone Array with Source Inside

The remainder of this thesis focuses on circular microphone arrays with the source or the array focus inside the array. Figure 4.6 contains screen shots of the beam shapes obtained by changing the signal frequency in a Matlab GUI (the entire GUI is shown in Figure 3.1). Figure 4.6 demonstrates the effect of increasing frequency on beam shape. Figure 4.6 (from b to e) shows the continuum of beamforming from non-directional (wide area of high intensity), to focused (small area of high intensity), to aliased (regions of high intensity far from the focus). Observing the beam shapes next to one another, as in Figure 4.6, allows one to express the degree of beam focus by comparing one image to another. For example, the beam shape in Figure 4.6d is more focused than that in Figure 4.6c. It is evident from Figure 4.6 that as frequency increases, the main lobe becomes smaller. However, at 1500 Hz (Figure 4.6e), aliasing starts to interfere with the beam's focus. A similar continuum of beam shapes could be obtained by increasing the array radius or decreasing the speed of sound⁴.

Figure 4.6 Screen shots from GUI showing beam shapes: a) shows position of focus at centre (red) and 8 microphones encircling it (blue), b)- f) show beam shapes with radius=0.15 m, speed of sound= 300 m/s and signal frequencies of b) 300 Hz, c) 700 Hz, d) 1100 Hz, e) 1500 Hz, f) 1900 Hz. White areas have higher power, black areas have lower power⁴.

4.2 Focus Index

Looking at the beam shape, one is able to get a visual subjective impression of the degree of focus of the beam, although this did not seem sufficient. It was determined that a simple numerical objective measure of beam focus was required. It is generally preferable to have an objective measure to predict subjective performance³⁰. An extensive

literature review did not turn up a useful metric for this purpose, and so a “Focus Index” is proposed⁴.

4.2.1 Other Beamforming Metrics

Other beamforming metrics were found in the literature, but these are not the most ideal for the environment addressed in this thesis. Some examples of these metrics are: Directivity Index, Beam Width, Front-to-Back Ratio, Array Gain, and Asymmetry Index.

Directivity Index is used for far-field (whereas this thesis emphasizes near-field). It describes the ability of the array to suppress a diffuse noise field. It is the ratio of the transfer function in the look direction of the array to the spatial integration over all directions²⁹. Directivity Index is not suitable because it is intended for far-field beamforming and it is computationally complex compared to the index proposed by this thesis.

Beam Width measures the main lobe size and it is defined as the angle enclosed by zeros. This is suitable for beamforming with a linear array or for a source located outside a circular array. Beam Width is not suitable for evaluating beamforming within a circular microphone array because the beam is shaped differently.

Front-to-Back Ratio is useful in applications with no specific look direction, where it is only necessary to distinguish between desired sources in front of the array and unwanted noise behind the array²⁹ and this is not the aim of this thesis.

Array Gain is the ratio of the SNR (signal to noise ratio) between one sensor and the array^{24, 29}. This metric is somewhat similar to the proposed Focus Index, however, though it is very effective at comparing the improvement of signal quality using different arrays, it is dependent upon the position of the noise.

Asymmetry Index is a measure of the symmetry of the beam pattern³¹, which again does not describe the degree of focus of the beam pattern.

Although these are important metrics, they do not reflect the beam's focus in the same way as the proposed Focus Index which was deemed to be more suitable for evaluating beamformer performance for the purpose of this thesis.

4.2.2 Proposed Focus Index

This thesis defines the "Focus Index", a new performance metric for evaluating the focus achieved by circular microphone array near-field beamformers where the focus point is contained within the array. The Focus Index compares the average power of signals received from regions at different distances from the selected focus point. A

narrow main lobe (a high intensity region close to the focus point) results in a high Focus Index, whereas a larger main lobe or aliasing (high intensity regions farther from the focus point) will reduce the Focus Index. The validity of the Focus Index is demonstrated by using empirical experiments, by comparison with known beam shapes, and by simulating changes in parameters with a known effect on beam focus⁴.

After investigating several possibilities for Focus Index, the following index is deemed to be the most accurate at reflecting the beam's focus⁴.

$$\text{Focus Index} = \frac{P_0}{\sum_{k=0}^M (k+1)^2 P_k} \quad (4.1)$$

P_k indicates the average power of the signal received from a particular region of the field after beamforming. The sound field is divided into regions delineated by concentric circles around the selected focus location. P_0 indicates the power at the focus (at one point), whereas the remaining P_k 's indicate the average power within rings surrounding the focus. Figure 4.7 illustrates how the concentric rings might be arranged around the desired focus location. For the purposes of these experiments, the sound field is considered to be the same size as the area within the array, and M had a value of 10.

Figure 4.7 Diagram of concentric circles where average power is calculated relative to the power at the ideal focus, indicated by a red “x”.

4.2.3 Calculation of Focus Index in Simulations

Focus Index is found by dividing the field into concentric circles and computing the average power between concentric circles. It is important to calculate power at enough points in order to make evident any major features of the beam shape (this could also be achieved by integration over the areas). However, calculating power at a larger number of points increases the time required to compute Focus Index.

In simulations of homogeneous media, it is possible to explore the symmetry of the beam shape. If the desired focus is at the center of a circular microphone array, then the beam shape is symmetrical. For instance, if there are 4 microphones, then the sound field consists of 8 identical sectors. Adjacent sectors are mirrored over the dashed lines

shown in Figure 4.8. Therefore, in simulations, this symmetry is used to reduce the complexity of computations. In each sector, an odd number of rays with a large number of points are taken. Note also that if the source is located anywhere along a line passing through a microphone and the center of the circular array, then bilateral symmetry occurs.

Figure 4.8 Sector (in gray) which repeats eight times in a four microphone circular array if the focus is at the center of the array

4.2.4 Empirical Focus Index Results

Focus Index must not only account for the size of the main lobe (surrounding the red “x”), but also for any aliasing. A small main lobe should mean a larger (better) Focus Index, but the appearance of aliasing should reduce the Focus Index. The proposed Focus Index does just that by increasing the weight for power appearing further away from the main lobe.

A few hypothetical situations are used to test out the validity of the Focus Index at accurately reflecting the effects of main lobe size and aliasing on beam focus. It is assumed that $M=3$ in equation 4.1. The trials illustrated in Table 4.1 show that the Focus Index chosen does accurately reflect the effects of main lobe size and aliasing.

Trial 1 indicates the perfect situation: All the power is contained in the focus location and there is no power anywhere else. This receives the “perfect” Focus Index of 1. Comparing trials 2 to 4, as the power moves further from the focus location, the Focus Index suffers (decreases). This reflects the effects of aliasing. Comparing trials 2, 5, and 6 the effect of a widening main lobe is apparent. This also results in a lowering of the Focus Index, as expected. Trials 7, 8, and 9 show the combined impact on Focus Index of both a wide main lobe and aliasing. These results are as expected.

Table 4.1 Focus Indices Resulting from Selected P_i

Trial number	P_0	P_1	P_2	P_3	Focus Index
1	1	0	0	0	1
2	1	0.5	0	0	0.3333
3	1	0	0.5	0	0.1818
4	1	0	0	0.5	0.1111
5	1	0.7	0.5	0	0.1205
6	1	0.7	0.5	0.2	0.0870
7	1	0.5	0.7	0.2	0.0800
8	1	0.2	0.7	0.5	0.0621
9	1	0.2	0.5	0.7	0.0571

A few general observations relating Focus Index and beam shape are⁴:

- 1) As the main lobe becomes smaller, the Focus Index increases. This is because the Focus Index puts a larger weight on power present farther from the focus location.
- 2) When aliasing appears, this creates areas of high intensity farther from the focus location (with higher weighting) which causes the Focus Index to decrease.
- 3) When both occur at the same time, that is, the main lobe becomes smaller but aliasing is more prevalent, then these two factors will compete in how they affect the Focus Index since they cause opposite trends.
- 4) The maximum Focus Index is 1 (when power is only present at the focus location) and the minimum is 0.

4.2.5 Focus Index Results for Simulations

The results shown in Figures 4.9 to 4.14 are based on simulations. They show the influence of various parameters on the Focus Index. The effect of changing these parameters is well known and can also be assessed from beam shape. The benefit of Focus Index compared to beam shape is that it allows one to express the degree of focus based on a number, the Focus Index⁴.

Figures 4.9 and 4.10 illustrate that the Focus Index increases with signal frequency (in agreement with Figure 4.6) and array radius. Figure 4.11 shows that increasing the speed of sound leads to a decline in the Focus Index. This is because as the

speed of sound increases, the differences in arrival times between two different points in the field will be very small. Hence, the main lobe will be wider and a larger area will be amplified. Figures 4.9, 4.10, and 4.11 have plateaus centered at about 1500 Hz, 0.22 m, and 250 m/s, respectively. The plateaus are due to the appearance of aliasing counteracting the effect of the reduced main lobe area. Figure 4.12 illustrates that as the number of microphones increases, the beam focus improves until there are 7 microphones. The plateau in Figure 4.12 occurs when aliasing disappears and the main lobe size becomes constant⁴. Figure 4.13 is used to illustrate the parameter that is being changed in Figure 4.14. Figure 4.14 shows that as the distance between the focus and the center of the array increases, the Focus Index decreases.

Figure 4.9 Focus index versus signal frequency, where radius=0.15 m, speed of sound=340 m/s, and number of microphones= 8^4

Figure 4.10 Focus index versus array radius, where signal frequency=1000 Hz, speed of sound=340 m/s, and number of microphones= 8^4

Figure 4.11 Focus index versus speed of sound, where signal frequency=1000 Hz, radius=0.15 m, and number of microphones=8⁴

Figure 4.12 Focus index versus number of microphones, where signal frequency=1000 Hz, radius=0.15 m, and speed of sound=340 m/s⁴

Figure 4.13 Diagram showing location “a”, at the center of the array and focus location “b”, 0.05 m from it (as shown in Figure 4.14)

Figure 4.14 Focus index versus distance between focus and center of array, where number of microphones=8, signal frequency=1000 Hz, radius=0.15 m, and speed of sound=340 m/s (Figure 4.13 shows location of “a” and “b”)

4.3 Chapter Summary

This chapter addresses two methods of evaluating beamformer performance: beam shape and Focus Index. Beam shape is explained and different representations and simulation results are illustrated. The new metric, Focus Index is defined and described. The validity of Focus Index is shown through empirical trials and simulations. The effect on beam focus caused by varying signal frequency, radius of the array, speed of sound, number of microphones, and position of focus is demonstrated using Focus Index.

Chapter 5 Source Localization

This chapter describes some methods of source localization and related simulation results. Some of the work presented in this chapter was published in the *Proceedings of the International Conference of the IEEE Engineering in Medicine and Biology Society* as well as the *Canadian Medical and Biological Engineering Society Conference* in September of 2004^{3,4}.

5.1 Far-Field Source Localization

As described in section 2.4.2, in far-field, it is assumed that the wave fronts are planar because the sound source is so far away that the curvature is so broad it appears flat. Localization in far-field can be done using only two microphones. Using both the *separation* distance between the microphones as well as the *delay* between time of arrival, it is possible to determine the direction (the angle, α or β) from which the source originated²⁴ (as in Figures 5.1 and 5.2). This is much like the process that takes place in humans: With two ears humans are able to approximate the direction of an incoming sound. The angle of the direction from which the sound originates is calculated as in

Equation 5.1³. In the far-field, it is not possible to determine the depth of the source location. Adding a third microphone, as in Figure 5.2, will only serve to verify that the first calculation is correct as it will result in a look direction parallel to that determined in the first calculation.

$$\alpha = \cos^{-1} \left(\frac{\text{delay}}{\text{separation}} \right) \quad (5.1)$$

Figure 5.1 Far-field source localization³

Figure 5.2 Far-field localization results with two microphones and three microphones

5.2 Near-Field Source Localization

Using the near-field assumption, as described in section 2.4.2, permits one to determine the direction as well as the depth of the sound source relative to the microphone positions. Figure 5.3 illustrates a possible arrangement of three microphones around a source in the near-field. Near-field is more appropriate than far-field for measurements on the scale of the human chest³.

Figure 5.3 Near-field source localization, the red “x” indicates the sound source³

Two methods of near-field source localization are described. Section 5.2.1 describes Localization Method A which is based on the maximum power in the beam shape. Section 5.2.2 describes Localization Method B which is based on differences in arrival time and involves solving systems of equations.

5.2.1 Localization Method A – Maximum Power in Beam Shape

Beam shape is a way of illustrating the degree of attenuation of sound in all areas of the field when the microphone array is tuned to focus on one particular location. Section 4.1 explains how beam shapes were formed by computing the power for a source originating at each point in the field subject to the beamformer delays designed for a specific focus location.

Beam shape can also be used for the purpose of source localization. If the sound is originating from one point in space, then it is delayed in a predictable manner upon arrival at the microphones. Assuming that the location of the source is unknown, if the beamformer is used to try to focus at various points in the field, the point at which the power is found to be maximal will be the source location. This is called Localization Method A. This procedure requires at least three microphones to find the source location. The more microphones that are used, the more the beam becomes focused⁴.

5.2.2 Localization Method B – Solving Systems of Equations

In the near-field, the estimated source location is calculated based on the difference in sound arrival time at each microphone (based on ¹⁵ and ³²). This delay is found by determining the cross-correlation between the signals received at two microphones. For instance, $delay_{1,2}$ indicates the delay in sample time units between signals received at microphones 1 and 2. Then, the delay is converted from sampled time units to distance units using the speed of sound, v_{sound} , and the sampling frequency, f_s , as in Equation 5.2. This new delay, $delta_{1,2}$, now expressed in distance units, represents the difference between the distance from microphone 1 to the source and from microphone 2 to the source (this is expressed in Equation 5.3 and illustrated in Figure 5.4). Then, using the standard distance equation (Equation 5.4), one is able to construct a system of two equations, Equation 5.5 and Equation 5.6³. The coordinates of microphones 1, 2, and 3 are (x_1, y_1) , (x_2, y_2) , and (x_3, y_3) , respectively. The coordinates of the source (x_s, y_s) , are

unknown. The solution of the system of equations (Equations 5.5 and 5.6) reveals the coordinates of the source. With more microphones it is possible to set up many similar sets of systems of two equations to solve. Each of these solutions estimates the location of the source.

Figure 5.4 Near-field source Localization Method B, the red “s” indicates the sound source³

$$\delta_{1,2} = (\text{delay}_{1,2}) (v_{\text{sound}}) / f_s \quad (5.2)$$

$$\delta_{1,2} = \text{dist}_{M1_to_S} - \text{dist}_{M2_to_S} \quad (5.3)$$

$$\text{dist}_{M1_to_S} = \sqrt{(x_1 - x_s)^2 + (y_1 - y_s)^2} \quad (5.4)$$

$$\delta_{1,2} = \sqrt{(x_1 - x_s)^2 + (y_1 - y_s)^2} - \sqrt{(x_2 - x_s)^2 + (y_2 - y_s)^2} \quad (5.5)$$

$$\delta_{1,3} = \sqrt{(x_1 - x_s)^2 + (y_1 - y_s)^2} - \sqrt{(x_3 - x_s)^2 + (y_3 - y_s)^2} \quad (5.6)$$

For the purposes of this thesis, this is called Localization Method B. This procedure requires at least three microphones to find the source location. If more microphones are used, the average of all solutions is returned as the estimated source location³.

5.3 Near-field Localization Simulations

5.3.1 Description of Simulations

The validity of Localization Method A and B are first examined using Matlab to conduct the simulations. Later, experimental trials are conducted to verify the results (Chapter 6 contains the experimental trials). Both localization methods are tested in a GUI (graphical user interface) created in Matlab for this purpose. The GUI used is a slightly modified version of the GUI shown in Figure 3.1. Additional functions that are added to the GUI for this purpose include the calculation and display of the Focus Index and the results of source Localization Method A and B.

In the simulations, eight microphones are used to localize one source. The source is either a 1 kHz tone or a recording of normal heart sounds. The microphones and the source are assigned specific coordinates. The radius of the microphone array is 15 cm.

The source is assigned three different positions for simulations (shown in Figure 5.5): in the center of the array, near mic 1, and between mics 3 and 4. When placed near mic 1 and between mics 3 and 4, the source is at 5 cm from the edge of the array. The source is assumed to be a point source.

The microphone inputs are simulated by delaying the source signal as much as it would have been delayed had it been coming from the assigned position. The speed of sound transmission used in the simulations is 40 m/s, the approximate speed of sound propagation in the chest¹⁵. Matlab performs near-field source localization, as described in sections 5.2.1 and 5.2.2, and returns the estimated source location³. Trials were done using many different values for the parameters and a subset of these results are presented in section 5.3.2.

Figure 5.5 Layout of microphones and positions of center of sources (indicated by a red “x”) in simulation trials

5.3.2 Simulation Results

Table 5.1 shows the results of the simulations using either a 1 kHz tone or heart sounds located at the center, near microphone 1, and between microphones 3 and 4. The estimated source locations (using Method A and B) are given as well as the Focus Index. Table 5.2 is the legend for Figures 5.6 to 5.11. Figures 5.6 to 5.11 show the beam shapes as well as the localization results associated with each line in Table 5.1. A detailed discussion of these results can be found in section 5.3.3.

Table 5.1: Summary of Simulation Results

Location of Source Center (cm)	Input	Estimated Source Location (cm)		Beam Shape	Focus Index
		Method A	Method B
Center (0, 0)	1 kHz tone	(0, 0.79)	(0, 0)	Fig. 5.6	0.0170
	Heart sounds	(0, -0.79)	(0, 0)	Fig. 5.7	0.0100
Near mic 1 (10, 0)	1 kHz tone	(9.38, 0.79)	(3.02, 0.15)	Fig. 5.8	0.0166
	Heart sounds	(9.38, -0.79)	(10.29, 0)	Fig. 5.9	0.0069
Between mic 3 and 4 (-3.83, 9.24)	1 kHz tone	(-3.13, 8.68)	(-0.41, 10.0)	Fig. 5.10	0.0151
	Heart sounds	(-3.13, 8.68)	(-3.86, 9.20)	Fig. 5.11	0.0070

Table 5.2 Legend for Figures 5.6 to 5.11

Symbol	Meaning
Microphone locations
Location of point source
Estimated source location found using Method A - maximum power in beam shape
Estimated source location found using Method B – solving systems of equations
0 1	Relative power in the sound field, where 1 is maximum power and 0 is minimum power

Figure 5.6 Beam shape and localization results for 1 kHz tone source located at center in simulation (the red X disappears because it is directly under the green X)

Figure 5.7 Beam shape and localization results for heart sounds source located at center in simulation (the red X disappears because it is directly under the green X)

Figure 5.8 Beam shape and localization results for 1 kHz tone source located near mic 1 in simulation

Figure 5.9 Beam shape and localization results for heart sounds source located near mic 1 in simulation (the red X is partially covered by the green X)

Figure 5.10 Beam shape and localization results for 1 kHz tone source located between mics 3 and 4 in simulation

Figure 5.11 Beam shape and localization results for heart sounds source located between mics 3 and 4 in simulation (the red X is partially covered by the green X)

5.3.3 Effectiveness of Localization

Effectiveness of localization is measured by source localization error. Source localization error is considered to be the distance between the true source location and the estimated source location⁴. In all the simulations, localization error is minimal for both Localization Methods A and B. In many cases, the source location is identified exactly.

According to Kompis *et al.*, the spatial resolution of an acoustic imaging system cannot resolve differences below about 2 cm¹⁵.

Figures 5.6 and 5.7 show the results for 1 kHz and heart sounds when the source is located in the center of the array. In these figures, Method B identifies the source location exactly and Method A is very close (less than 1 cm away). Figures 5.9 and 5.11 demonstrate results for heart sound recordings located away from the center of the array. In these figures, both localization methods are very effective and both have very low localization errors. Figures 5.8 and 5.10 show the results for a 1 kHz source located away from the center of the array. These results are not as precise as those presented in the other figures.

Using Localization Method A, it is not possible to measure the power at every location in the field since there are an infinite number of points in the field, so only a limited number of points are calculated (500 points, in this case). This is the main cause of localization error in Method A. In Localization Method B, the localization error is due mainly to the sampling frequency reducing the resolution of the delay³. Using cross-correlation, one cannot detect a delay which is less than one sampling period. Overall, the results show that both Localization Method A and B are valid.

5.4 Chapter Summary

This chapter describes methods of localizing a single source in the far-field and in the near-field. Localization Method A determines the location by maximum power in the beam shape. Localization Method B determines the location by solving a system of equations based on differences in arrival times at specific microphone locations. This chapter also covers the simulations which are used to test out Localization Methods A and B and analyzes the results.

Chapter 6 Verification Through Experiments

This chapter demonstrates the results of the in vitro verification using two experimental models. Focus Index, beam shape, and localization error are addressed.

These experiments described in this chapter are done in order to test the results obtained in simulation (in Chapters 4 and 5) and also in an attempt to bridge the gap between the properties of sound propagation in free-field and in the human body. The human body is a non-homogeneous transmission medium. At the boundaries between media with differing speeds of sound, there are reflections and scattering of sound in addition to attenuation and refraction^{15,24}. Modeling of the complex system of the human body is outside the scope of this thesis (this is further discussed in section 7.1). The human body can be adequately represented for the purposes of this thesis by a homogeneous medium. This thesis describes the implementation and testing of two simple experimental models.

6.1 Experimental Model I

Experimental Model I is described in section 3.2.1. It consists of a plastic container lined with foam to reduce reflections.

6.1.1 Description of Experiments

In these experiments, eight microphones are used to localize one source. The source is either a 1 kHz tone, a 100 Hz tone, white noise, or a recording of normal heart sounds. The source is placed in three different positions for recordings: in the center of the array, as close to mic 1 as possible and between mics 5 and 6 (as shown in Figure 6.1). The closest distance possible between the center of the source and the mics is 11 cm. The radius of the container is 14.3 cm but due to the lining of the plastic container and the size of the sound source this is the smallest distance. The source is a speaker, that when baffled has dimensions 9 cm x 10 cm x 4 cm.

Figure 6.1 Layout of microphones and positions of center of sources (indicated by a red “x”) in Experimental Model I

6.2 Experimental Model II

Experimental Model II is described in section 3.2.2. It consists of a block of foamed gelatin used to model the speed of sound in the chest.

6.2.1 Foamed Gelatin

Gelatin is an incomplete protein extracted from collagen connective tissues, tendons, bones, ligaments, and skin of pigs. This glutinous material is obtained from animal tissues by prolonged boiling. Foamed gelatin has a speed of sound similar to that

in the human torso¹⁵. Through Internet research, it was found that foamed gelatin is used in theatre for creating such things as elf-ears or wounds. Additionally, it is used for ballistics tests. Gelatin, once set, is Jello-like in texture. Foaming involves adding air bubbles into the gelatin. If it is done properly, the bubbles remain in the gelatin when it has set. Several recipes for foamed gelatin were attempted. Many Websites suggest using bicarbonate and vinegar (or wine acid or ascorbic acid) to create the bubbles. It was found that the most effective and consistent method was to combine 0.8 kg unflavoured gelatin powder with 2.5 L of water, whip it with a hand mixer to add air bubbles, and leave it to set overnight. The foamed gelatin was made in a cylindrical container and popped out for testing. Figure 3.6 shows a sample of the foamed gelatin used in this experiment.

6.2.2 Description of Experiments

In these experiments, eight microphones are used to localize one source. The source is either a 1 kHz tone or a recording of normal heart sounds. The source is placed in three different positions for recordings: in the center of the array, as close to mic 1 as possible and between mics 3 and 4 (as shown in Figure 6.2). In this model, it is possible to bring the center of the source 3 cm from the edge of the gelatin. The radius of the gelatin is 13.4 cm but due to the size of the sound source this is the smallest distance. The source is a speaker, that when baffled has dimensions 3 cm x 2 cm x 2 cm.

Figure 6.2 Layout of microphones and positions of center of sources (indicated by a red “x”) in Experimental Model II trials

6.3 Analysis of Sound Recordings

This section describes the post-processing that is done once the microphone inputs are stored. Analysis of the data is done using Matlab, not in real-time. Yet another GUI was created in order to perform all the computations and display the results. Screen shots are shown in Figures 6.3 to 6.16.

6.3.1 Calibration and Speed of Sound

Microphones are tested to ensure synchronicity of recording. Calibration of sample delays is done using a reference at the center of the microphone array. The recordings are also calibrated for amplitude, again using a reference at the center of the microphone array. Calibrations are repeated throughout the experiments to detect subtle changes which can be caused by loss of battery power in the microphone amplifiers or by moving the microphones.

The speed of sound in air was 340 m/s under the experimental conditions¹⁷. This is the speed of sound which is used in the later calculations for Experimental Model I.

According to Kompis *et al.*, foamed gelatin should have a speed of sound around 40 m/s¹⁵. However, this may be dependent on the particular recipe used to make the foamed gelatin. Some testing was done to ensure that this speed of sound would apply in Experimental Model II. A microphone is placed immediately beside a source in the gelatin. The delay between this microphone and the other microphones is determined. By combining the sample delay, the sampling frequency, and the measured distance between the microphones, it is possible to estimate speed of sound. The estimated speed of sound was found to be approximately 40 m/s. This is the speed of sound used in further computations for Experimental Model II.

6.3.2 Beam Shape

Beam shape is calculated as described in section 4.1 and 5.2.1 and is displayed below in Figures 6.3 to 6.16.

6.3.3 Localization Method A – Maximum Power in Beam Shape

Section 5.2.1 describes this localization method. The power is computed at 500 locations in the field. These locations correspond with those already calculated and displayed in the boxes in the beam shape figures (Figures 6.3 to 6.16). The localization results are displayed in these figures as well as in Table 6.1.

6.3.4 Localization Method B - Solving Systems of Equations

Section 5.2.2 describes this localization method. With eight microphones, a large number of systems of equations are possible, so only a random subset of these is considered. Estimated source locations outside the container (Experimental Model I) or outside the foamed gelatin (Experimental Model II) are disregarded and if an imaginary part comes with the result, it is dropped. Otherwise, estimated source locations are averaged. The localization results are displayed in Figures 6.3 to 6.16 as well as in Table 6.1.

6.3.5 Focus Index

Focus Index is calculated as described in section 4.2.2. It is not possible to use symmetry here to reduce computational complexity. The power at 180 points per power level is calculated and 10 power levels are defined. These results are displayed in Table 6.1.

6.4 Experimental Results

Table 6.1 shows the results of the Experimental Model I and II using either a 1 kHz tone, 100 Hz tone, heart sounds, or white noise as input. The source is located at the center, near microphone 1, and between microphones 5 and 6 or 3 and 4. The estimated source locations (using Method A and B) are given as well as the Focus Index. Table 6.2 is the legend for Figures 6.3 to 6.16. Figures 6.3 to 6.16 show the beam shapes as well as the localization results associated with each line in Table 6.1. A detailed discussion of these results can be found in section 6.5.

Table 6.1: Summary of Experimental Results

Location of Source Center (cm)	Input	Estimated Source Location (cm)		Beam Shape	Focus Index
		Method A	Method B		
		$\times$	$\times$		
Experimental Model I					
Center (0, 0)	1 kHz tone	(1.49, 0.75)	(2.73, 0)	Fig. 6.3	0.0063
	100 Hz tone	(-1.49, 3.76)	(0.83, -3.31)	Fig. 6.4	0.0021
	Heart sounds	(-2.98, -3.76)	(0.92, 2.31)	Fig. 6.5	0.0021
	White noise	(0, 2.26)	(0, 0)	Fig. 6.6	0.0035
Near mic 1 (3.3, 0)	1 kHz tone	(1.49, -0.75)	(3.26, -0.73)	Fig. 6.7	0.0047
	Heart sounds	(0, 0.75)	(1.57, 2.99)	Fig. 6.8	0.0021
Between mic 5 and 6 (-3.0, -1.3)	1 kHz tone	(-1.49, -0.75)	(1.48, -1.63)	Fig. 6.9	0.0034
	Heart sounds	(-4.47, -2.26)	(0, 0)	Fig. 6.10	0.0021
Experimental Model II					
Center (0, 0)	1 kHz tone	(0, -0.71)	(0.48, -0.73)	Fig. 6.11	0.0099
	Heart sounds	(1.40, 0.71)	(0.22, 0.46)	Fig. 6.12	0.0058
Near mic 1 (10.4, 0)	1 kHz tone	(4.19, 0.71)	(5.56, 1.87)	Fig. 6.13	0.0019
	Heart sounds	(2.79, 0.71)	(4.95, 1.40)	Fig. 6.14	0.0023
Between mic 3 and 4 (-4.0, 9.6)	1 kHz tone	(0, 7.76)	(0, 6.23)	Fig. 6.15	0.0026
	Heart sounds	(1.40, 10.6)	(0, 6.23)	Fig. 6.16	0.0027

Table 6.2 Legend for Figures 6.3 to 6.16

Symbol	Meaning
Microphone locations
Outline of the container (Experimental Model I) or of the foamed gelatin (Experimental Model II)
Location of source center
Outline of the source
Estimated source location found using Method A - maximum power in beam shape
Estimated source location found using Method B – solving systems of equations
Relative power in the sound field, where 1 is maximum power and 0 is minimum power

Figure 6.3 Beam shape and localization results for 1 kHz tone source located at center of Experimental Model I

Figure 6.4 Beam shape and localization results for 100 Hz tone source located at center of Experimental Model I

Figure 6.5 Beam shape and localization results for heart sounds source located at center of Experimental Model I

Figure 6.6 Beam shape and localization results for white noise source located at center of Experimental Model I (the red X disappears because it is directly under the green X)

Figure 6.7 Beam shape and localization results for 1 kHz tone source located near mic 1 of Experimental Model I

Figure 6.8 Beam shape and localization results for heart sounds source located near mic1 of Experimental Model I

Figure 6.9 Beam shape and localization results for 1 kHz tone source located between mics 5 and 6 of Experimental Model I

Figure 6.10 Beam shape and localization results for heart sounds source located between mics 5 and 6 of Experimental Model I

Figure 6.11 Beam shape and localization results for 1 kHz tone source located at center of Experimental Model II

Figure 6.12 Beam shape and localization results for heart sounds source located at center of Experimental Model II

Figure 6.13 Beam shape and localization results for 1 kHz tone source located near mic 1 of Experimental Model II

Figure 6.14 Beam shape and localization results for heart sounds source located near mic1 of Experimental Model II

Figure 6.15 Beam shape and localization results for 1 kHz tone source located between mics 3 and 4 of Experimental Model II

Figure 6.16 Beam shape and localization results for heart sounds source located between mics 3 and 4 of Experimental Model II

6.5 Discussion of Experimental Results

6.5.1 Comparison of Experimental Model I and II

There are a few main differences between Experimental Model I and II:

- Model I has a higher speed of sound than Experimental Model II
- Model I has a larger sound source than Model II
- Model I has less variation in source positions than Model II

These differences are well reflected in the results. As demonstrated in section 4.2.5, increasing speed of sound leads to a decrease in Focus Index. This means that it can be expected that beam shapes of Model II will have smaller main lobes and a lower Focus Index than Model I, if all other parameters remain the same. However, it must be noted that the difference in the source sizes which are shown in the figures (Figures 6.3 to 6.16) offset this phenomenon. Additionally, as indicated in section 4.2.5, as the distance between the source and the center of the array increases, the Focus Index decreases. This means that it should be expected that more focused beam shapes occur when the source is located in the center of the array, and worse focus results when the source is located further from the center.

Figure 6.3 (Model I) and Figure 6.11 (Model II) represent the results for a 1 kHz tone input located at the center of the array. Figure 6.11 shows a much smaller and more focused main lobe than Figure 6.3. Table 6.1 also indicates that the Focus Index corresponding to Figure 6.11 is higher than that for Figure 6.3, as expected. The same comparison can be made between Figures 6.5 (Model I) and 6.12 (Model II), which represent the results for the heart sounds input.

By observing the results of Model I (Figures 6.3 to 6.10) in comparison to Model II (Figures 6.11 to 6.16), it can be seen that Model I had less source location variability than Model II due to the restrictions of the experimental set-up. Thus, Model I demonstrates little variability in beam shape for the different source positions (because they do not differ very much), whereas the opposite is true for Model II. In particular, the change in beam shape in Model II from Figure 6.12 to Figure 6.15 is evident. The equivalent comparison for Model I is Figures 6.5 and 6.10. The Focus Index also reflects this conclusion. From Table 6.1, it is apparent that there is little variability in Focus Index within Model I compared to Model II. In Model I for a heart sounds input, the Focus Index remains constant at 0.0021 regardless of source position. In Model II for a heart sounds input, the Focus Index changes from 0.0058 to 0.0023 to 0.0027. The lower Focus Indices are for source positions farther from the center of the array, as expected.

The results for Experimental Model II can be compared with the simulation results presented in section 5.3, as the parameters are very similar. The beam shapes and effectiveness of localization are comparable.

6.5.2 Effect of Source Location

According to section 4.2.5, as the focus location moves away from the center of the array, the Focus Index decreases. Therefore, it can be expected that beam shapes will be more focused the closer the source is to the center of the array. As described in section 6.5.1, for Model I, the location of the source varies very little and so the beam shapes do not change much with source location. On the other hand, for Model II, much larger changes in beam shapes can be seen (greater detail appears in section 6.5.1).

Additionally, it can be observed that the best source localization occurs when the source is located at the center of the array (e.g. Figure 6.12). The worst source localization occurs when the source is farthest from the center of the array (e.g. Figures 6.13 to 6.16).

6.5.3 Effect of Different Sound Inputs

This comparison is best made by looking at the results of Model I, when the source is located at the center and the input is changed. Figures 6.3, 6.4, 6.5, and 6.6 show the results of the 1 kHz tone, the 100 Hz tone, the heart sounds input, and the white

noise input, respectively. According to section 4.2.5, increasing the frequency leads to an increase in Focus Index. In other words, a larger frequency leads to a smaller main lobe and better focus. Figure 6.3 and 6.6 are very similar, as are Figures 6.4 and 6.5. This is due to their frequency content. Figure 6.3 has a more focused beam shape compared to Figure 6.4 which has a lower signal frequency. Figures 6.4 and 6.5 demonstrate similar beam shapes because their signal frequencies are in the same range. These same observations are true for Models I and II when comparing different inputs located in the same positions.

6.5.4 Comparison of Localization Methods A and B

Localization Method A finds the source location by measuring the power at different points in the beam shape and identifying the location of maximum power. In Figures 6.3 to 6.16, the location identified by Method A is marked with a yellow “X”.

Localization Method B finds the source location by solving a system of equations based on the difference in time of arrival of signals at different microphones and the measured microphone positions. In Figures 6.3 to 6.16, the location identified by Method B is marked with a green “X”.

Occasionally, Localization Method A identifies more than one location with the same maximum value. In this case, the location closest to the one identified by

Localization Method B is chosen. Also, Localization Method B gives different answers based on which random set of systems of equations is selected to solve. As the number systems of equations solved increases, the average of their solutions stabilizes around one point. The average solution will converge to this point faster if the data from microphones closer to the suspected (or known) source location is selected.

Both methods are the most effective when the source is located in the center of the array and when the beam is well focused and the Focus Index is high. Figures 6.11 and 6.12 are examples of the precision that is possible with these methods. Both methods are the least effective when the source is far away from the center and the Focus Index is low (Figures 6.13 to 6.16). In most cases (Figures 6.3 to 6.12), both methods perform adequately and identify a source location which is within the red line indicating the boundaries of the source itself (which is not actually a point source). Localization Method A almost always identifies the center of the main lobe of the beam shape. Since only 500 points in the field are tested, the location of maximum power may be missed. For example, in Figure 6.15, it is believed that Method A identifies a point of aliasing and misses the ideal solution because it is between two points that are calculated.

6.5.5 Validity of Focus Index

The goal of the Focus Index proposed in section 4.2.2 is to obtain a numerical representation of the degree of beam focus. A high Focus Index corresponds to a narrow main lobe with very little power in regions far from the desired focus location. A low Focus Index corresponds to a broader main lobe or areas of high power far from the desired focus location. The results demonstrate good correspondence between the Focus Index reported in Table 6.1 and the corresponding beam shapes illustrated in Figures 6.3 to 6.16.

According to Table 6.1, the highest Focus Indices are found in Figures 6.11, 6.12, and 6.3. Figures 6.3 and 6.12 show a narrow and focused main lobe with good attenuation throughout the remainder of the field. Figure 6.12 has a narrower main lobe than Figure 6.3 thus its Focus Index is higher, as expected. Figure 6.11 shows a very small main lobe as well as some aliasing in the remainder of the field. The average power in areas distant from the desired focus is sufficiently low that Figure 6.11 has the highest Focus Index of all the experimental results shown here.

In Figures 6.4, 6.5, 6.8, and 6.10 the Focus Index remains constant at 0.0021. Upon observation, these figures have relatively the same beam shape and focus and the Focus Index supports this observation.

In Figures 6.14 and 6.16 the Focus Index is much lower than would be expected from looking at the beam patterns because the beams are focused but off-center. This suggests that there is the need for another Focus Index-type metric which measures how focused the beam shape is relative to the point where the power is maximal. This would differ from the current Focus Index which is a measure of how focused the beam shape is relative to the intended focus location.

6.5.6 Relationship Between Focus Index and Localization Error

It has been observed that the higher the Focus Index, the lower the localization error. For instance, Figures 6.11, 6.3 and 6.12 have the three highest Focus Indices and Localization Methods A and B give results very close to the center of the sound source. As the Focus Index decreases, Localization Methods A and B tend to give results farther away from each other and farther from the center of the sound source.

6.5.7 Sources of Error During In Vitro Trials

In vitro, the results were encouraging although there are sources of error present in the experiments. The coordinates of the microphones as well as the placement of the source are used in all the computations and these are measured values and thus subject to measurement error.

Many assumptions are used to simplify the computations. However, since in most cases the results are adequate (within 2 cm), it can be concluded that these assumptions were reasonable. The source is assumed to be an omnidirectional point source, which it is not (sections 3.2.1 and 3.2.2 describe the experimental set-up). In the calculations for Experimental Model I, the sound is assumed to have traveled only through air before reaching the microphones when in fact sound travels through air, then through a plastic container to reach the stethoscope chest pieces which transmits the sound to the microphones. Another problem with Experimental Model I is that the sounds recorded at each microphone are not only sound that arrives in a straight line directly from the source, but also sound that is transmitted along the wall of the plastic container (multipath). In Experimental Model II, the speed of sound is assumed to be constant throughout the foamed gelatin. The stethoscope chest pieces also modify the sounds in all experiments. Background noise could somewhat contaminate the received signals but it is eliminated as much as possible by using sound proofing materials. The sampling frequency is 16 kHz, which reduces some of the accuracy, but including more samples would also increase the computation time. Section 6.5.4 indicates errors introduced by the localization methods and ways to combat them.

According to Kompis et al., acoustic sources are localized most accurately if the assumed speed of sound is correct. If the speed of sound is underestimated, artifacts appear as scattered weak sound sources. If the speed of sound is overestimated, areas near

the microphones become amplified. They concluded that overestimation is better than underestimation¹⁵.

6.6 In Vivo Experiments

Some preliminary trials were conducted on humans. It is found that the sound collection system created is suitable for collecting human heart and lung sounds. Individual microphone recordings are very clear. The recordings differ depending on the position of the chest piece, demonstrating obvious spatial differentiation. The spatial differences are due to the underlying anatomy (described in section 2.1). Heart sounds are definitely amplified more on the left side of the body, as expected. Combining several microphone inputs through beamforming improves the quality of heart sounds somewhat. Further experiments and analysis in this area will require ethics approval.

6.7 Chapter Summary

This chapter describes the experiments used to verify the simulation results presented earlier. Two in vitro experimental models are used. Experiments are done using different sound inputs in different locations. Calculations are done to determine

beam shape, location of the source (using two methods), and Focus Index. Results are compared and analyzed. In vivo experiments are briefly addressed.

Chapter 7 Conclusions and Future Work

This chapter concludes the thesis, summarizes the main findings, and suggests future work.

7.1 Challenges in Human Application

The proposed multi-sensor stethoscope for chest sound localization presents many challenges. The first is that the chest cavity is nonhomogeneous. It is composed of bone, air, blood, lung parenchyma, muscles, etc. all of which have different acoustic properties¹⁵. At the boundaries between tissues a change in the speed of sound, signal distortion, dispersion, scattering, and refraction can be found²⁴.

The speed of sound transmission in the chest is quite diverse. The chest wall and the heart transmit sound at about 1500 m/s. The large airways transmit at 222 to 312 m/s. The lung parenchyma transmits at between 23 and 60 m/s¹⁵. Variations in the speed of sound affect the arrival time of the sounds and also prevent the sound signals from traveling in a straight line from the source to the skin surface.

Normal heart and lung sounds are in the ranges of 20 to 150 Hz and 40 to 600 Hz, respectively. Adventitious sounds are present at frequencies up to 1200 Hz for the heart and 2000 Hz for the lungs^{18, 19, 20}. Low frequencies pose a challenge for delay-and-sum beamformers which focus better at higher frequencies. It is also difficult to obtain sound capture systems with a flat response over this frequency range.

Another challenge is that it is very difficult to form a sound transmission model for the human body. There is great variation from person to person and even over time. A different model would be required for different sizes of males, females, adults, and children. This would also be affected by the percentage of body fat and the presence of disease affecting function or structure.

7.2 Overall Conclusions

This thesis examined beamforming using a circular array of stethoscopes. The thesis provided background information on thoracic anatomy and heart and lung sounds as well as microphone arrays and beamforming. Following this, simulation and verification results were presented and analyzed.

In simulation, some parameters affecting beamformer performance were evaluated (signal frequency, array radius, speed of sound, number of microphones, and

distance of the source from the center of the array). The performance of the beamformer was evaluated using its beam shape as well as the proposed Focus Index. The proposed Focus Index is a new performance metric that indicates the degree of focus of the beam shape. The thesis also suggested two methods for source localization. These methods were tested in simulations.

In addition to simulations, verification was done using two experimental models. The experimental results supported the conclusions drawn from the simulations. The results demonstrated the expected relationship between beam shape, Focus Index, and localization error.

7.3 Future Work

An area of microphone arrays that is still in need of more research is the use of microphone arrays in nonhomogenous media. Reflections, delays, interference, and asymmetry must be considered. But first, one must determine the characteristics that affect the performance of microphone arrays in nonhomogeneous environments.

In vivo trials with healthy and diseased persons could be conducted using the experimental set up presented in this thesis. In the future, this kind of application could allow a physician to listen to sounds emanating from a particular area of the heart, such

as a heart valve, to improve medical diagnosis. Additionally, a microphone array placed around the human chest could be used to provide an acoustic scan of the heart and lungs, which could then be analyzed by a computer.

It is also possible to extend the work presented in this thesis to arrays of different shapes (in two or three dimensions) or used for other applications. For instance, a rectangular shaped array may be used for audio conferencing. The proposed Focus Index could equally be applied to a rectangular array, a linear array, or to a circular array with the focus outside of the array.

References

-
- ¹ H. Pasterkamp, S. S. Kraman, G. R. Wodicka, "Respiratory sounds – advances beyond the stethoscope," *American Journal of Respiratory and Critical Care Medicine*, Vol. 156, pp. 974-987, 1997.
- ² S. Mangione, L. Z. Nieman, "Cardiac auscultatory skills of internal medicine and family practice trainees," *Journal of the American Medical Association*, vol. 278, pp. 717-722, Sept. 1997.
- ³ A. M. McKee, R. A. Goubran, "Multisensor Stethoscope for Chest Sound Localization," *Proceedings of the Canadian Medical and Biological Engineering Society Conference*, Sept. 2004.
- ⁴ A. M. McKee, R. A. Goubran, "Beam Shape, Focus Index, and Localization Error for Performance Evaluation of a Multisensor Stethoscope Beamformer," *Proceedings of the International Conference of the IEEE Engineering in Medicine and Biology Society*, Sept. 2004.
- ⁵ Z. Sharif, S. Daliman, A. Z. Sha'ameri, S. H. S. Salleh, "An expert system approach for classification of heart sounds and murmurs," *Sixth International Symposium on Signal Processing and its Applications*, vol. 2, pp.739-740, Aug. 2001.
- ⁶ W. W. Myint, B. Dillard, "An electronic stethoscope with diagnosis capability," *Proceedings of the 33rd Southeastern Symposium on System Theory*, 18-20, pp.133-137, March 2001.
- ⁷ R. Rhoades, R. Pflanzner, *Human Physiology*. Montreal: Saunders College Publishing, 1996.
- ⁸ E. J. Akesson, J. A. Loeb, L. Wilson-Pauwels, *Thompson's core Textbook of Anatomy*. New York: J. B. Lippincott Company, 1990.

-
- ⁹ Z. M. Zin, S. H. S. Salleh, A. Z. Sha'ameri, "The design of heart classification system," *Sixth International Symposium on Signal Processing and its Applications*, vol. 2, pp.741-742, Aug. 2001.
- ¹⁰ M. H. Swartz, *Textbook of Physical Diagnosis- History and Examination*. Toronto: W. B. Saunders Company, 1998.
- ¹¹ Texas Heart Institute, "Anatomy of the Human Heart". [online] Available at: <http://www.tmc.edu/thi/anatomy2.html>.
- ¹² Enchanted Learning, "Answers: Label the Lungs Diagram". [online] Available at: <http://www.enchantedlearning.com/subjects/anatomy/lungs/label/labelanswers.shtml>.
- ¹³ W. W. Hui, R. A. Pitt, J. P. Matonick, J. K.-J. Li, "Comparison of heart sounds recorded at the chest and a remote arterial site," *Proceedings of the IEEE 28th Annual Northeast Bioengineering Conference*, 20-21, pp. 61-62, April 2002.
- ¹⁴ The R.A.L.E. Repository, "R.A.L.E Lung Sounds Demo". [online] Available at: <http://www.rale.ca/LungSounds.htm>.
- ¹⁵ M. Kompis, H. Pasterkamp, G. R. Wodicka, "Acoustic imaging of the human chest," *Chest Journal*, vol. 120, pp. 1309-1321, Oct. 2001.
- ¹⁶ T. Bergstresser, D. Ofengeim, A. Vyshedskiy, J. Shane, R. Murphy, "Sound transmission as a function of lung volume," *Journal of Applied Physiology*, Vol. 93, pp. 667-674, Aug. 2002.
- ¹⁷ D. E. Hall, *Basic Acoustics*. New York, NY: Harper and Row Publishers, Inc., 1987, p. 139.
- ¹⁸ M. T. Pourazad, Z. K. Moussavi, G. Thomas, "Heart sound cancellation from lung sound recordings using adaptive threshold and 2D interpolation in time-frequency domain," *Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2003, Vol. 3, 17-21, pp. 2586-2589, Sept. 2003.
- ¹⁹ E. C. Guler, T. Elgin, B. Sankur, Y. P. Kahya, "Spectral analysis of respiratory sounds in healthy and pathological subjects," *Proceedings of the 1992 International Biomedical Engineering Days*, 18-20, pp. 70-74, Aug. 1992.

-
- ²⁰ V. Padmanabhan, J. L. Semmlow, and W. Welkowitz, "Accelerometer type cardiac transducer for detection of low-level heart sounds," *IEEE Transactions on Biomedical Engineering*, vol. 40, issue: 1, pp. 21-28, Jan. 1993.
- ²¹ M. Kompis, H. Pasterkamp, Y. Oh, Y. Motai, G. R. Wodicka, "Spatial representation of thoracic sounds," *Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 20, No. 3, pp. 1661-1664, Oct. 1998.
- ²² W.H. Neo, B. Farhang-Boroujeny, "Robust microphone arrays using subband adaptive filters," *IEEE Proceedings – Vision, Image and Signal Processing*, vol. 149, issue: 1, pp. 17-25, Feb. 2002.
- ²³ Y. R. Zheng, R. A. Goubran, M. El-Tanany, "A Nested Sensor Array Focusing on Near Field Targets," *Proceedings from IEEE Sensors 2003 Conference*, pp. 834-848, 2003.
- ²⁴ D. H. Johnson, D. E. Dudgeon, *Array Signal Processing: Concepts and Techniques*. Englewood Cliffs, NJ: Prentice-Hall, 1993, pp. 111-190.
- ²⁵ J. G. Ryan, R. A. Goubran, "Array optimization applied in the near field of a microphone array," *IEEE Transactions On Speech and Audio Processing*, vol. 8, issue: 2, pp. 173-176, March 2000.
- ²⁶ Stethographics Inc., Company website. [online] Available at: www.stethographics.com.
- ²⁷ A. Haghighi-Mood, J. N. Torry, "Coherence analysis of multichannel heart sound recording," *Computers in Cardiology*, 8-11, pp. 377-380, Sept. 1996.
- ²⁸ Y. Tamai, S. Kagami, H. Mizoguchi, K. Sakaya, K. Nagashima, T. Takano, "Circular microphone array for meeting system," *Proceedings of IEEE Sensors 2003 Conference*, Toronto, ON, pp. 1100-1105, 2003.
- ²⁹ M. Brandstein, D. Ward (Ed.), *Microphone Arrays*. New York: Springer, 2001.
- ³⁰ J. G. Ryan, "Predicting the quality of sound pickup using microphone arrays," *IEEE ASSP Workshop on Applications of Signal Processing to Audio and Acoustics*, 15-18, pp. 161-164, Oct. 1995.

³¹ P. Moquin, S. Dedieu, R. Goubran, "Beamforming for a microphone array embedded in asymmetrically shaped objects," *Journal of the Acoustic Society of America*, vol. 113, p. 2297, 2003.

³² Y. Huang, J. Benesty, G. W. Elko, "Microphone arrays for video camera steering," *Acoustic Signal Processing for Telecommunication*, S. L. Gay and J. Benesty, Ed. Norwell, MA: Kluwer Academic Publishers, 2000, ch. 11, pp.239-259.